CERGE
Center for Economic Research and Graduate Education
Charles University Prague

Anti-Corruption Mechanisms in Economic Models of Corruption

Jana Krajčová

Dissertation

Prague, November 2009

Jana Krajčová

Anti-Corruption Mechanisms in Economic Models of Corruption

Dissertation

Prague, November 2009

Acknowledgements

My deepest gratitude goes to my advisor, Andreas Ortmann, without whose encouragement, guidance, and commitment this dissertation would not have been possible. It was first of all him who inspired my academic curiosity and helped me to formulate the research ideas. By carefully reading and commenting on countless revisions of my manuscripts, he taught me how to question thoughts and express my ideas. He taught me how to do research.

I am also very thankful to Marco Celentani, who helped me very much to shape Chapter 1 of my dissertation. By asking insightful questions he helped me understand my research area better. He also guided me through the difficult transition from having an economic model to writing a paper about it.

I appreciate the contribution of all the members of my dissertation committee, Peter Katuščák, Libor Dušek, and Junghun Cho, for giving me valuable questions and comments at various stages of my research. My special thanks go to Peter Katuščák, whose probing questions helped me to refine Chapter 1 and make it read much better than it initially did.

For inspirational discussions and numerous brainstorming sessions, I would like to thank my friend, Lenka Drnáková, and my husband, Marian Krajč. I would also like to thank everyone who assisted me in running the experiments: Ondřej Rydval, Marian Krajč, Jan Myslivec̆ek, and Katarína Kálovcová.

Last but not least, I acknowledge financial support from various institutions that funded parts of my research. The general research idea was formulated under the support of the CERGE-EI/World Bank Fellowship. Chapter 1 of this dissertation originated during my mobility stay at the Universidad Carlos III de Madrid. I am grateful to the Mobility Fund of CERGE-EI and Charles University in Prague for financial support and to the Universidad Carlos III de Madrid for kindly hosting me as a visiting student. The experimental research in Chapters 2 and 3 would not have been possible without funding provided by the Grant Agency of the Czech Republic (grant No. 402/04/0167).

Contents

Preface

Corruption remains a serious problem in many modern economies. The severe consequences of corruption have been documented in numerous empirical studies (e.g. Mauro 1995, Tanzi 1998, Hwang 2002, Gupta et al. 2002). Therefore, the design and implementation of effective anticorruption measures remains an important concern.

The problem is that, due to its illegal and secret nature, corruption is hard to detect and observed indictments are likely only the tip of the iceberg. Therefore, authorities should rely on anti-corruption measures that (ex-ante) undermine incentives for bribery by, for example, reducing the expected gains from corruption, and/or by undermining trust between potential illegal partners. The former can typically be ensured by e.g. increased penalties and/or probabilities of detection, if possible, or by increasing the opportunity cost of corruption by offering high-enough wages. The latter can be facilitated by leniency policies.

The first chapter of my dissertation addresses the interplay of monitoring, detection, and wages and is motivated by recent discussions in the Czech Republic about the introduction of new means of monitoring police officers on duty. Leaving aside the question whether such monitoring systems can be effective anti-corruption measures, I conjectured that they might negatively affect officers' work incentives - as suggested by Cremer (1995) in a labor market setting without corruption. Therefore, I analyze the effect of increased monitoring on both the agent's incentives to be corrupt and to exert a high level of effort.

I construct a two period principal-agent model. The agent's type is his predisposition to corruption and is unobserved before signing the contract. I assume two types of agent: honest, who are never corrupt (their exogenously given psychic cost of corruption is too high), and opportunistic, who can be either corrupt or non-corrupt in a given period (their psychic cost of corruption is low enough). The output that the agent produces is altogether affected by three factors: the agent's chosen effort (low or high), exogenous luck (good or bad) and the agent's decision to be corrupt or not in the given period. Only a high level of effort, good luck and non-corruptibility can lead to high produced output. The principal offers a two-period contract in which she commits to wages, monitoring technology and a firing policy. Both monitoring and the possible firing of the agent happen after the first period. If the principal does not monitor the agent, she only observes the produced output. Alternatively she could detect the agent who misbehaved by using one of the monitoring technologies: Monitoring of Effort, or Monitoring of Corruption.

I use my model to illustrate that it might not always be in the principal's best interest to conduct monitoring. The intuition is in line with that in Cremer (1995) - when the principal only observes output she can threaten to fire the agent after producing low output. This threat is an important part of the incentive package (the high quality type has an extra incentive to put in more effort to improve his probability of producing high output). Once the principal observes the reasons underlying low output, this threat is no longer credible (the principal prefers to keep the high quality type to firing him and hiring a new agent of unknown characteristics for the second period) and thus, this part of the incentives disappears.

I find such a "Cremer-like" result in the case when the principal monitors the effort choice of the agent. Monitoring of Effort improves the sorting of types. Therefore, the principal cannot credibly threaten to fire the agent after observing low output, which results in weaker incentives of the agent to "behave." Specifically, I find that the agent might have more incentive to be corrupt when Monitoring of Effort is implemented. Monitoring of Corruption, in contrast, does not improve the sorting of types. Therefore, the principal's firing rule after the first period is exactly the same as with No Monitoring. Consequently, no "Cremer-like" result emerges. Quite on the contrary, the expected penalty for corruption serves as an additional enforcement mechanism and Monitoring of Corruption negatively affects the agent's incentives to be corrupt. As a result, the principal can pay a lower expected wage with Monitoring of Corruption than with No Monitoring (or with Monitoring of Effort).

Importantly, the results suggest that the effect of monitoring one dimension of the agent's strategy profile may spill over to the other dimension. Specifically, my conclusion differs from those in the motivating literature:

the incentives to exert effort are not distorted by monitoring; it is indeed the incentive to be corrupt that might be negatively affected if the principal monitors the agent's effort choice.

In the two chapters that follow I analyze (experimentally) leniency policies as promising anti-corruption measures.

Leniency policies award fine reductions of varying intensities to wrongdoers who spontaneously report an illegal agreement and thereby help to convict their accomplice(s). They serve as both the enforcement mechanism as much as a means of deterrence in that, if appropriately designed and implemented, they have the potential to undermine trust between wrongdoers. Leniency policies have been in use in a number of countries (e.g. Great Britain, France, Germany, the USA, or Canada, also the Czech Republic and Slovakia), mostly recognized as an anti-cartel mechanism. Also in the economic literature, the promising anti-cartel properties of leniency programs have been confirmed (see e.g. Spagnolo 2004, Apesteguia et al. 2004, Bigoni et al. 2008a,b). Corruption also arises from an illegal relationship that relies heavily on trust and therefore leniency policies have a potential to work well in corruption scenarios.

Leniency policies to deter cartels are, however, not directly applicable as anti-corruption measures: cartel deterrence is essentially a simultaneous game while strategies, payoffs, and the move structure of anti-corruption measures are asymmetric. To my knowledge the first theoretical work analyzing the various effects of leniency policies on corruption is Buccirossi and Spagnolo (2006). The authors show that poorly designed moderate policies may have a serious counter-productive effect: they might allow a briber to punish at a relatively low cost a partner who does not respect an illegal agreement. In other words, some leniency policies might actually provide an enforcement mechanism for occasional illegal transactions. Thus, contrary to the intention, they might increase corruption.

Buccirossi and Spagnolo's result, together with the theoretical and experimental evidence from the literature on cartel deterrence, suggests that the potential of leniency policies to undermine trust between wrongdoers hinges upon proper design and implementation (Spagnolo 2004, Apesteguia et al. 2004, Bigoni et al. 2008a,b).

In Chapters 2 and 3 I use a generalized Buccirossi and Spagnolo model (see Richmanová and Ortmann 2008) for the experimental testing of leniency policies as an anti-corruption measure. In Chapter 2, I use this model to study the effects of parameter changes that do not affect the theoretical prediction. In Chapter 3, which has Andreas Ortmann as a co-author, we use this model to study the effects of "loaded" and "neutral" instructions.

It has been documented in the literature that a change in parameterization that does not affect the theoretical prediction might indeed have consequences for the behavior of subjects in the lab (see e.g. Goeree and Holt 2001). Due to their social context, anti-corruption experiments might be particularly sensitive to changes in design. Moreover, in the generalized Buccirossi and Spagnolo game, the action bringing the highest possible payoff is also associated with a risk of considerable loss, therefore, risk or loss attitudes are also likely to play a role. Altogether, I expected that subjects in the lab might not behave in accordance with the theoretical prediction, especially when the prediction is made under an assumption of risk neutrality. It was therefore important to test the sensitivity of subjects to some parametric changes. The question I ask is whether by making corruption more attractive by i) increasing the potential gain and ii) reducing the penalty if bribery is discovered, I can induce more corruption in the lab even if the theoretical prediction suggests no change (high-incentive treatment). I do indeed find a significant effect of parametric change. Even though the change I implemented has no consequences for the theoretical prediction, I observe much more corruption in the high-incentive treatment. Thus, my results suggest that details do matter, and this should be taken into account in law enforcement design.

Following the earlier work of Abbink and Hennig-Schmidt (2006), we study in the final chapter the effect of "loaded" instructions in the generalized Bucccirossi and Spagnolo model. Abbink and Hennig-Schmidt find no significant impact of instructions framing (although this result may be the artifact of too few data points, as the difference seems visible from the graphs). The authors conclude that this result may be caused by the nature of the game: it is very simple, and as it was designed to capture all the basic features of bribery, even with neutral wording, subjects may have deciphered what the experiment was about. Our bribery game includes stages where players can report their opponents and receive leniency, which makes it more complex and also potentially more susceptible to the non-neutral context. Therefore, it calls for a separate analysis. We find a strong gender effect - male and female participants react differently to the non-neutral context. The effect of context becomes significant once we allow for gender-specific coefficients. Thus, in contrast to the results of Abbink and Hennig-Schmidt (2006), we find that a bribery context indeed makes a (significant) difference.

Moreover, our result on gender is interesting also from a policy perspective. It adds to existing evidence (e.g. Swamy et al. 2001) that using women in "dangerous" positions in organizations, like in auditing, might be one way to fight corruption.

In both Chapters 2 and 3 we also study to what extent the subjects' decisions in the experiment can be explained by their basic socio-demographic characteristics – the econometric analysis provides limited evidence on their role. Our data suggest that trust and preferences towards others might also play a role. Results in Chapter 2 suggest that calibration, i.e. parameterization that reflects "real-life" situations reasonably well, might be even more important than in other scenarios. The results in Chapter 3 suggest that subjects in a bribery game engage in all sorts of social considerations, including moral scruples, which should not be dismissed by experimenters looking for relevant policy implications. Both chapters provide a testbed for the experimental testing of anti-corruption measures and add evidence to the on-going discussion of the need for socio-demographic controls. Further experimental testing of leniency policies might have to take these findings into account.

That said, let me stress that the experimental results provide tantalizing evidence that there is something to the Bucccirossi and Spagnolo model. For the first time, a number of real world subjects are shown to understand and use a strategic and ill-designed legal environment to enforce occasional corrupt transactions, as hypothesized by Buccirossi and Spagnolo. Our results indeed provide a confirmation of Buccirossi and Spagnolo's hypothesis.

Chapter 1

On the Hidden Costs of Monitoring Corruption and Effort

Abstract

In this chapter, I analyze the effects of monitoring on an agent's incentives in a two-period principal-agent model in which the agent decides on his effort and corruptibility. The agent's type and strategy are unknown to the principal. I compare incentive-compatible wages under three different scenarios: when the principal does not monitor and only observes output; when she monitors the agent's effort choice; and when she monitors the agent's corruptibility. I find that monitoring of effort improves the sorting of types but it might also give the agent more incentive to be corrupt. Monitoring of corruption does not improve the sorting of types but it negatively affects the agent's incentive to be corrupt.

Keywords: corruption, monitoring, contract, incentive-compatibility
JEL classification: D73,D86,K42

1.1 Introduction

It has been shown in the career concerns literature that the principal's monitoring might harm the agent's incentives (see e.g. Cremer 1995; Holmström 1999; Dewatripont, Jewitt, and Tirole 1999; Prat 2005): if the agent's ability is unknown and his effort cannot be perfectly observed by the principal, the agent might have an incentive to put in more effort in order to signal high ability. Once the principal starts to observe the agent's ability, this incentive disappears.

Similarly, the literature on intrinsic motivation suggests that monitoring, as a display of distrust, is likely to be detrimental to the agent's intrinsic motivation and his dedication to the job (e.g., Kreps 1997 and Benabou and Tirole 2003).

Monitoring thus might not necessarily lead to improved performance of the agent and in some situations the principal might be better off not monitoring.

The present research is motivated by a recent discussion in the Czech Republic. To fight corruption, law enforcement authorities discussed installing cameras and GPS systems into police cars in order to monitor traffic police officers on duty. Leaving aside the question whether such monitoring systems can be effective, an important question is whether they might negatively affect officers' incentives. Some officers, especially those who are honest, may be offended and reduce effort in retaliation (the intrinsic motivation argument); others' incentives to signal high quality may be affected (the career concerns argument). Some officers may simply believe that as long as they are honest, they do not need to work hard because the principal values honesty more than effort. One way or another, the principal may face reduced incentives on the part of the agent to put in effort. In addition, the effect of monitoring systems on individual corruption decisions is in question.

The decision to be corrupt results from officers' attitudes towards corruption, which might be based on their home-grown moral scruples, the perception of attitudes towards corruption in their social context, the perception of risks connected to corrupt behavior, etc. Consequently, some people may be more prone to corruption than others. In the literature this is modelled as the psychic cost of corruption (see, for example, Celentani and Ganuza 2002, or Cule and Fulton 2005). The actual decision to take (or ask for) a bribe then also depends on the value of the bribe. We say that people who are conditionally susceptible to taking (or asking for) a bribe are opportunistic.

When talking about petty police corruption, it is likely that some fraction of people are honest because their psychic cost of corruption is too high to outweigh (relatively low) potential monetary gains.[1]

An authority, or the principal, evaluating the work of police officers, might, in general, observe no more than their daily output which, for example for traffic police, might be measured by the total value of issued tickets. How much an officer collects in fines, however, depends on several factors: the effort he is exerting and his corruptibility, but also on the number of misbehaving drivers in his area or other exogenous factors. Therefore, an officer who collects few fines may not necessarily be corrupt or shirking. Thus, the observed "output" gives the principal only limited information about the actual behavior of the officer and the incentive-compatibility design of a reward and punishment system becomes an issue.

Assume that the principal wants to induce a high level of effort and non-corruptibility. The question I ask is whether, and at what price, the principal can influence the incentives of officers in the desired way by monitoring them and thereby acquiring additional information about their type and/or their action choices.

I examine the effect of the imperfect partial monitoring of an agent (traffic police officer) who has two binary decision margins: (1) unobservable effort (diligence and time spent pursuing misbehaving drivers); and (2) an unobservable decision about how much of the produced output to extract for himself (how many of the violations to fine officially and how many to "fine" in the form of a bribe). Hence the model captures an element of hidden action (effort) with an element of hidden information (the output before any potential extraction). On top of that, there is another source of hidden information in that the agent may be intrinsically honest (will never extract any rents) or opportunistic (will extract rents under the right economic incentives). However, the type is not known before signing the contract and hence this is just another layer of hidden information rather than an element of adverse selection.

I study and compare three monitoring technologies: (1) No Monitoring, (2) Monitoring of Corruption, and (3) Monitoring of Effort. If the principal does not monitor she can only observe output that the agent produces. If

[1] Aleš Pachmann, who was associated with the Police Academy of the Czech Republic, suggested in private communication that surveys on police corruption suggest about 10% of police officers being unconditionally honest.

the principal monitors, the monitoring is partial in that only one decision margin is monitored and it is imperfect in that the probability of detecting a lack of effort or corruptibility is less than one. I am interested in the effects of monitoring on both agent's effort choice and his corruptibility.

The police officers thus affect the quality of enforcement they are expected to provide by deciding on the effort and their corruptibility. These two decisions are not completely independent, though. On the one hand, putting in more effort generates a higher expected output (catching more violating drivers), which gives the agent more opportunities for rent extraction. On the other hand, planning to extract a certain fraction of the rent gives the agent more incentives to exert effort. This generates two key differences from the career concerns literature, where the result is driven by a certain substituability between the *exogenous* ability of the agent and the *endogenous* effort decision. First, in the present model, even though the propensity to corruption is exogenously given to the agent, the actual decision to take bribes is endogenous and therefore it may respond to economic incentives (whereas ability in the career concerns literature is out of the agent's control). Second, the interplay of effort choice and corruptibility is more complicated than in the career concerns literature: high effort increases the probability of high produced output (before any rent extraction) but, at the same time, it creates more opportunities for bribery and thereby might reduce observed output (after any rent extraction). Effort choice and corruptibility are not substitutes. Consequently, the monitoring is likely to affect the incentives in a somewhat different way than in the career concerns literature. In fact, any effect of monitoring on one decision margin may "spill over" to the other margin as well, which is one of the interesting properties of the model.

The punishment for low effort or rent extraction is the termination of the contract with the agent. Since at least two periods are necessary to make monitoring and its consequences for future payoffs an important part of the incentives package, I construct a two-period principal-agent model in which a principal offers a two-period contract to an agent of unknown type. In the contract, the principal commits herself to a monitoring technology and to wages to be paid to the agent at the end of the employment. The principal decides, at the end of the first period, based on her monitoring, whether she will keep her current employee or fire him and hire a new one for the second period. I assume that the principal is also ex-post exogenously committed to monitoring and to the firing rule specified in the contract. Monitoring is costless.

I define the agent's type as his predisposition (or propensity) to corruption. I assume that whether an agent with a given predisposition to corruption takes a bribe in a given period is a decision that he takes in response to the reward and punishment system that is in place. To simplify the analysis, I assume that there are two types of agents in the population: an honest type, with zero utility from corruption, and a corrupt type, with some positive utility from corruption.

I assume that before signing the contract the agent does not know his type.[2] I make this assumption to avoid adverse selection, which would make the agent's strategy space richer and thereby the analysis more complicated. Agents discover their types quickly, though. Imagine, for example, a new officer who has no prior experience either with this kind of a job or with bribery as such. It is likely to take him only a few transactions to find out what his "price" is, or whether there is any price at all. To simplify the analysis, I assume that the agent learns his type the very same moment when he starts the job. Thus when choosing the effort level, he already makes an informed decision.

In order to model the exogenous factors mentioned above that make it impossible for the principal to distinguish the types solely based on the output realization, I assume that luck will affect the output of both types of agent. Imagine days with high traffic and lots of speeding drivers (may be caused by good road conditions) and days when people drive more in compliance with law (may be caused by poor visibility or other poor road conditions). In order to keep things as simple as possible, I assume that luck comes into play only when the agent is exerting a high level of effort.

I assume that two levels of output are possible: low or high. Only a hardworking and lucky agent produces high output and can extract for himself through bribes that part of output that is above the output resulting from low effort.[3] This brings about an interesting conflict on the side of the principal: on the one hand, she prefers high effort as it increases the likelihood of high output yet, on the other hand, high effort implies also more opportunities for opportunistic agents to collect bribes. It is therefore not clear whether it is

[2]In reality, he might have some at least imperfect knowledge but I will abstract from that for simplicity. The adverse selection case might be an interesting extension of the model.

[3]Producing low output and then, in addition, extracting some part of it for himself would automatically reveal the agent's type to the principal and hence eliminate any chance of second-period profits.

in the principal's best interest to automatically induce a high level of effort.

In general, the principal might have different preferences about the agent's actions. In real-life scenarios, the principal might put more weight on how much effort the agent puts in, or, alternatively, non-corruptibility might be a top priority for her. As for the former, imagine the example of traffic police. Even an opportunistic officer, who is working really hard in order to create more opportunities for corruption, might produce some, though a smaller, deterrence effect (compared to a hard-working honest officer). The overall deterrence might, however, be higher than the one produced by the honest, but shirking, officer. Therefore, the principal might prefer to concentrate on the agent's effort choice. As for the latter, imagine for example a question of issuing driver's licences (or some other license or permit for that matter). In this case, with a relatively well-defined set of criteria, the question of effort (or, the amount of work dealt with) might be less important than the question of non-corrupt decision making in order to avoid the possible social costs that might be generated by, for example, unqualified drivers.

In this chapter, I refrain from making specific assumptions about the principal's utility function. Instead, I focus on the decision-making of the agent and on how his incentives are affected by monitoring. This way, the conclusions about the agent's incentives are relatively general, as they do not depend on the specific preferences of the principal or on the parameters of the model such as the proportion of honest types. Specifically, I analyze the case when the principal wants to induce the "most efficient" strategy profile, which includes high effort and non-corruptibility over both periods.

Cremer (1995) is the article that is most closely related to this study. Cremer used a two-period principal-agent model to demonstrate that increased monitoring (and hence more information about the agent's characteristics) may make it more difficult for the principal to commit to some threats, thereby weakening the agent's incentives.

Cremer's main result draws on the idea of renegotiation-proofness introduced by Dewatripont (1988). Cremer shows that with efficient monitoring technology (which allows the principal to learn at no cost everything about the agent's type/quality and action choices), the only renegotiation-proof contract will commit the principal to monitoring and to firing the agent who is found to be of low quality. With an inefficient monitoring technology (for which the cost of additional information is infinite and which is therefore analogical to the No Monitoring case), the principal will not conduct monitoring (and will only observe output) and will rehire the agent only if he produced high first-period output.

The intuition behind Cremer's result is that the additional information prevents the principal from committing to some threats. Assume, for example, a situation in which the principal commits to an efficient monitoring technology and, at the same time, she claims to fire the agent if he produces low output. The efficient monitoring technology, however, enables the principal to observe the reasons underlying low output at no cost. If she observes that the agent is of high quality and exerted high effort but was unlucky, the principal would prefer keeping that agent to having to go to the market and hiring a new agent of unknown characteristics (but of lower average quality). Knowing that no threat of such a kind would be credible when the efficient monitoring technology is employed, the agent's incentives are altered.

In Cremer's model the agent's type is his (exogenously given) suitability for the job which the agent cannot control and, thus, he only decides about the effort level. This is the most important difference from this chapter, in which I study the possible effects of monitoring when the agent decides about both the effort level and corruptibility.

I find a "Cremer-like" result in the case when the principal monitors the effort choice of the agent. Monitoring of Effort improves the sorting of types. Consequently, the principal cannot credibly threaten to fire the agent every time when observing low output. Therefore, the agent might have more incentive to be corrupt. To induce high effort and non-corruptibility over both periods, the principal has to pay a higher expected wage with Monitoring of Effort than with No Monitoring, to compensate for the lost part of incentives due to reduced risk of getting fired after the first period. Monitoring of Corruption, on the other hand, does not improve the sorting of types. Therefore, the principal's firing rule after the first period is exactly the same as with No Monitoring. Consequently, no "Cremer-like" result emerges. Quite on the contrary, the expected penalty for corruption serves as an additional enforcement mechanism and Monitoring of Corruption negatively affects agent's incentives to be corrupt. As a result, the principal can pay a lower expected wage with Monitoring of Corruption than with No Monitoring (or with Monitoring of Effort) to induce high effort and non-corruptibility over both periods.

Importantly, the results suggest that the effect of monitoring one dimension of the agent's strategy profile may spill over to the other dimension. Specifically, my conclusion differs from that in the motivating literature: the incentives to exert effort are not distorted by monitoring; it is indeed the incentive to be corrupt that might be negatively affected if the principal monitors the agent's effort choice. This is the main contribution of this chapter.

The remainder of this chapter is structured as follows. In section 2, I discuss the model. First, the main assumptions for the agent and for the principal are presented. Afterwards, I introduce the three types of monitoring technology. In section 3, I provide a discussion of the main result. All proofs can be found in the appendix. Section 4 concludes.

1.2 The Model

This section summarizes the main assumptions of the proposed model. An overview of the key notation for the model is provided in Table 1.1.

p	-	proportion of honest-type agents in the population
q	-	probability of being lucky
c	-	cost of exerting high effort
δ_c	-	exogenous probability of detecting corruption
δ_e	-	exogenous probability of detecting low effort
F	-	penalty imposed after detecting corruption
B_H/B_L	-	high/low output
A	-	utility from corruption, $A > 0$ for the opportunist
$w_{HH}/w_{HL}/w_{LL}$	-	two-period wage after producing high+high/high+low/low+low output
w_F	-	wage paid to an agent who is fired after the first period
e_H/e_L	-	high/low effort
C/NC	-	corrupt/non-corrupt
$\{e_1CD_1, e_2CD_2\}$	-	agent's strategy; e_iCD_i is effort choice and corruptibility decision (CD) in the period $i, i \in \{1, 2\}$, $e_i \in \{e_H, e_L\}$, $CD_i \in \{C, NC\}$
NM	-	No Monitoring
MC	-	Monitoring of Corruption
ME	-	Monitoring of Effort

Table 1.1: Overview of the key notation.

The main assumptions and the basic structure of the model are summarized in Table 1.2. More details are discussed below.

1.2.1 The Agent

There are two types of agent. As in Cremer (1995), I assume that p of them are good and $(1-p)$ of them are bad, where $0 < p < 1$. The good type, which I will call honest, is constituted of those agents whose psychic cost of engaging in illegal transactions is high enough to outweigh whatever potential benefits there may be. The bad type, which I will call opportunistic, is constituted of

The Agent

is one of two types

- honest type
  - → chooses effort level $\in \{e_H, e_L\}$
  - → is always non-corrupt
- opportunistic type
  - → chooses effort level $\in \{e_H, e_L\}$
  - → decides to be corrupt or non-corrupt

the observed productivity of either type depends on

- chosen effort level $\in \{e_H, e_L\}$
- exogenous realization of luck $\in \{lucky, unlucky\}$
- chosen "corruptibility" $\in \{corrupt, non\text{-}corrupt\}$

The Principal

At the beginning of the 1^{st} period offers a two-period contract in which she

- commits to a monitoring technology
  - → No Monitoring (NM) – the principal can only observe output realization
  - → Monitoring of Corruption (MC) – the principal can observe output and with probability δ_c also detect corruption
  - → Monitoring of Effort (ME) – the principal can observe output and with probability δ_e also detect low effort
- specifies two-period wages (w_{HH}, w_{HL}, w_{LL}) and the "firing" wage (w_F)
- specifies conditions under which the agent's employment continues after the first period

At the end of the 1^{st} period

- implements monitoring technology
- decides whether to keep or fire the agent based on the outcome of the monitoring and conditions stated in the contract
  - → if she keeps the agent, nothing changes and the two-period contract is fulfilled
  - → if she fires the agent, she terminates the two-period contract, pays the agent the "firing" wage w_F, and offers a one-period contract to a new agent.

At the end of the 2^{nd} period

- pays wages according to the applicable contract.

Table 1.2: Summary of the model.

those agents with lower psychic costs, who may be corruptible if the expected benefit of doing so is high enough. An opportunistic agent who decides to take bribes in a given period will be called corrupt; if he decides not to take bribes in a given period, he will be called non-corrupt.[4]

[4]Throughout the text, "honest type"/"honest agent" and "opportunistic type"/"opportunistic agent"/"opportunist" will always refer to an agent's type (his given (non)propensity to corruption). "Corrupt" ("non-corrupt") will refer to the opportunistic agent's decision to (not) take bribes in a given period.

Both types choose an effort level, which can take one of two possible values: high (e_H), or low (e_L). If the agent exerts high effort, he bears a cost of c, $c > 0$. The cost of exerting low effort is normalized to zero.

Two levels of output are possible – high output (B_H) and low output (B_L). B_H is assumed to be strictly greater than B_L. If exerting low effort, the agent of either type automatically produces low output. Both types of agent are capable of producing high output B_H. The only difference is that an opportunist may, through bribery, extract the difference between high and low output, and thus, at the end of the day the principal will observe low output B_L. If the agent exerts high effort, then the output he produces further depends on two things: (exogenous) luck and the agent's corruptibility. Denote the probability of being lucky q, $0 < q < 1$. The unlucky agent produces low output. The lucky agent produces high output.

The decision to be corrupt depends on the agent's utility from corruption. I assume that the agent's utility from corruption is A. A represents the agent's utility from extracting for himself through bribes that part of output that is above the output resulting from low effort;[5] it also factors in the agent's psychic cost of corruption. It is not necessary for the purpose of this chapter to specify how exactly these factors enter A, though.

For the honest type, I assume that A is equal to zero (their psychic cost of corruption is too high). The opportunists have one specific value of $A > 0$ which is common knowledge.

An opportunistic agent can be corrupt only on lucky days, when he can gain A by extracting the difference between B_H and B_L for himself. If the opportunist is corrupt, the principal will observe low output B_L. Thus, in fact, the realization of an agent's type is in fact equivalent with the realization of A; A is a parametric representation of the type.

I assume that the agent does not know "his" A before starting the job. However, I assume that he learns it right after he starts working and thus is able to adjust the effort level instantly.

[5]With the minimum output being B_L, the agent can, essentially, extract part of the difference between B_H and B_L. One could specify A as $A = \alpha(B_H - B_L)$, where $\alpha < 1$ – the agent extracts a linear part of the "extra output" and $\alpha < 1$ accounts for the psychic cost of corruption as well as some cost of bargaining (a bribe is typically lower than an actual penalty would be). For the sake of generality I, however, refrain from specifying such a specific relationship between gain from corruption and output. It is not necessary for the purpose of this paper.

1.2.2 The Principal

A risk-neutral principal offers a two-period contract to an agent of unknown characteristics. Before signing the contract neither the principal nor the agent know the agent's A; only the probability p of the distribution of A in the population is known.

In the two-period contract, the principal specifies the monitoring technology, wages to be paid after the second period, conditions under which the contract continues after the first period and the wage to be paid to the agent in case he is fired after the first period.

After observing the outcome of the first period (how much information is being observed depends on the chosen monitoring technology), the principal decides whether to continue the contract or to fire the agent and to offer a one-period contract to an agent of unknown characteristics. Note that the monitoring, if the principal commits to it, occurs only after the first period, i.e. the principal does not monitor after the second period.[6]

The two-period wages are contingent on observed output and also depend on the rehiring decision of the principal. Wages are paid at the end of the employment. The principal will pay w_{HH} after observing high output in both periods; w_{HL} after observing a combination of high and low output over two periods;[7] w_{LL} after observing low output in both periods;[8] and, finally, she will pay w_F in the case when she fires the agent based on the outcome of the first period (including monitoring in relevant cases).

[6] In this, I follow Cremer's approach. The monitoring, by giving (or not) additional information to the principal after the first period, should in general affect the sorting of the agents after the first period and thereby affect the incentives of the agent. The purpose of this paper is to explore in which direction the incentives are affected. The second period is important ex-ante, so that the agent needs to optimize over two periods and monitoring in between, when deciding about his actions. Basically, the second period captures the lost opportunity of the agent who misbehaves and might be fired afterwards.

[7] Here I implicitly assume symmetry, $w_{HL} = w_{LH}$, as in both cases a total output of $B_H + B_L$ is produced. Note that with some monitoring technologies the agent will be fired after producing B_L in the first period, in which case he will be paid w_F. The details will be discussed later on, as the firing rule is specific to the monitoring technologies chosen and it is part of the results to be shown.

[8] Here, the same comment applies as for w_{HL} - as with some monitoring technologies, the agent will be fired after delivering low output, w_{LL} will not always be relevant. More details follow later on.

1.2.3 Monitoring Technologies

Based on the agent's strategy space and on the preferences of the principal, three monitoring technologies seem relevant. First, the benchmark case, which I will call *No Monitoring* (NM), in which the principal does not monitor the agent. Second, the case in which the principal can monitor the corruptibility of the agent and discover it with some positive probability. I will call this case *Monitoring of Corruption* (MC). Finally, the case in which the principal can monitor the effort the agent puts in and discover shirking with some positive probability. I will call this case *Monitoring of Effort* (ME).

All three monitoring technologies affect the agent's incentives to exert a high level of effort and to be corrupt; how exactly they affect the agent's incentives is my primary focus.

No Monitoring (NM)

In this case the principal can only observe output at the end of the first period.

Table 1.3 summarizes all the possible combinations of the agent's type, his decisions (about effort and corruptibility), nature's moves (luck) and the result observed by the principal (level of output) after the first period. The two-period case is analogous but more complicated, as it involves combinations of the agent's actions and nature's moves over two periods. Moreover, the continuation of the employment into the second period depends on the outcome observed by the principal after the first period.

type	**effort**	**luck**	**corruptibility**	**observed output**
HONEST	high (e_H)	lucky	non-corrupt (NC)	B_H
		NOT lucky	non-corrupt (NC)	B_L
	low (e_L)	–	non-corrupt (NC)	B_L
OPPORTUNIST	high (e_H)	lucky	non-corrupt (NC)	B_H
			corrupt (C)	B_L
		NOT lucky	non-corrupt (NC)	B_L
	low (e_L)	–	non-corrupt (NC)	B_L

Table 1.3: Possible combinations of type, effort, luck, corruptibility and observed output; e_H/e_L stand for high/low effort, NC/C for non-corrupt/corrupt and B_H/B_L for high/low output.

Note that the principal observes high output only if she employs either an honest type who exerts high effort and is lucky, or an opportunist who exerts high effort, is lucky and is non-corrupt. In all other cases, the principal observes low output. Consequently, she is not able to distinguish which type she is currently employing based on observed output.

Monitoring of Corruption (MC)

In this case, the principal has access to a technology that allows her to detect corruption with some nonzero probability. I assume that this technology returns no false positive – it does not detect the agent as corrupt if he has not been corrupt. The agent who has been corrupt is detected with probability δ_c, where $0 < \delta_c < 1$, and detection leads to punishment with certainty. Thus, after being detected, the agent is fired. In addition, he is punished by an external law-enforcing authority and a fine F is imposed on him. The fine F and the detection rate δ_c are exogenous parameters.[9]

The principal monitors the agent after observing low output in the first period.

Monitoring of Effort (ME)

In this case, the principal has a technology that allows her to detect low effort. As in the previous case, I assume that the technology returns no false positive – it does not detect the agent as exerting low effort if he has not exerted low effort. The agent who has exerted low effort is detected with probability δ_e, where $0 < \delta_e < 1$. After being detected, the agent is fired with certainty. The detection rate δ_e is an exogenous parameter.[10]

The principal monitors the agent after observing low output in the first period.

[9] Typically, the penalties are set by law and collected by an external authority. It is, however, not crucial for this paper who in fact collects the fine, as it is not my goal to identify the optimal contract for the principal; rather I look at incentives of the agent and how they are affected by various monitoring technologies. In the case when the principal would collect the fine, F could be, in fact, simply a part of the wage. As regards the detection rate, one can think of it as the quality (or effectiveness) of the feasible monitoring technology.

[10] One can think of δ_e as the quality (or effectiveness) of the feasible monitoring technology. With ME, I do not assume an additional external punishment in the form of a penalty, as exerting low effort is not an illegal action.

1.3 Results

Before discussing the details, Table 1.4 below provides a brief overview of the main results. Specifically, I am looking for the effect that Monitoring of Corruption and Monitoring of Effort (as opposed to No Monitoring) have on: 1) sorting of types ("Does the information from MC/ME help the principal to fine-tune her firing rule?"); and 2) incentives of the agent to be non-corrupt and to exert effort.

	Sorting	**Non-Corruptibility**	**Effort**
MC	–	↑	–
ME	↑	↓	–

Table 1.4: Effects of Monitoring of Corruption (MC) and Monitoring of Effort (ME) on the sorting of types, agents' corruptibility and effort choice. "↑" corresponds to a positive effect, "↓" corresponds to a negative effect, and "–" corresponds to no effect.

Assumption 1 *Throughout the analysis, I assume that the principal prefers the honest type to exert high effort.*

In the simplest case, when the principal would set wages such that the honest type would prefer exerting low effort over both periods (the principal might want to do that to reduce her cost), the opportunist would prefer either low effort (when $qA < c$) or high effort and being corrupt (when $qA > c$). In either case, in the end both types would deliver low output. Consequently, firing and replacing the agent with a new one could not help to improve efficiency. Basically, I assume that the levels of output are such that the improvement in efficiency of the honest type outweighs the cost in wages to the principal. This assumption allows concentrating on (strategically) interesting cases and helps to simplify the analysis.

The principal's Bayesian updating, after observing the first-period output and the result of the monitoring, gives the first result:

Proposition 1 *Monitoring of Corruption does not improve the sorting of types. Monitoring of Effort improves the sorting of types if the probability of detection is high enough.*

Improved sorting means that, based on the result of monitoring, the principal can fine-tune her firing rule (compared to NM) so as to improve her probability of having an honest type for the second period.

The detailed proof can be found in the appendix,; a discussion follows below.

With all three monitoring technologies, there are, in general, three possible first-period strategy profiles for the opportunist: $\{e_H C\}$, $\{e_H NC\}$, $\{e_L NC\}$. According to *Assumption 1*, the principal always prefers the honest type to exert high effort e_H. Altogether, there are three possible strategy profiles that the principal might wish to induce in the first period: $e_H, \{e_H NC\}$, $e_H, \{e_H C\}$, and $e_H\{e_L NC\}$, where the first term denotes the strategy of the honest agent and the second term, in the braces, the strategy of the opportunist. Depending on the particular contract the principal offers, the outcome of the first period can give her more or less information about the type of the agent she is employing. To prove *Proposition 1* I compute for each monitoring technology the updated (a-posteriori) probabilities of having an honest type for all possible first-period strategy profiles and the observed outcomes of the first period. Then I compare the a-posteriori probability of having an honest type with the proportion of honest types on the market.

With No Monitoring, the only information the principal has after the first period is the realization of output. In the first case, when $e_H, \{e_H NC\}$ is induced in the first period, if high output is observed, it could have been produced by either an honest or an opportunistic lucky agent. Similarly, if low output is observed, it could have been produced by either an honest or an opportunistic unlucky agent. After the first period, the principal has no additional information about the type of the agent compared to the start of the employment and thus her best prediction about the probability that she is indeed employing an honest agent is p. The probability that she would hire an honest agent if she fires her current employee and goes to the market again is the same. Therefore she is indifferent between keeping and firing the currently employed agent.

In the other two cases, when $e_H, \{e_H C\}$ or $e_H, \{e_L NC\}$ is induced in the first period, if high output is observed, the principal knows with certainty that she is employing an honest agent, and she prefers to keep this agent to firing him and employing a new agent of unknown characteristics. If low output is observed, than the a-posteriori probability of having an honest type is lower than the proportion of honest types on the market and therefore, the principal would be better off firing her current employee and hiring a new agent.

Thus, in this case the principal will ex-ante commit to the following firing rule: "I will keep the agent who has produced high output and will fire the agent after observing low output." Ex-post, the principal cannot be better off by not keeping her ex-ante firing rule.

With Monitoring of Corruption, the principal can in addition detect a corrupt agent with some probability. After observing low output, detection is a sufficient signal that the current employee is indeed opportunistic. Unfortunately, no detection is not a sufficient signal of having an honest type (the a-posteriori probability of having honest type is lower than the proportion of honest types on the market). So, in the end, the principal will fire the agent after observing low output, no matter what the result of monitoring is. She will keep the agent who has produced high output. Thus, the principal's firing rule will be the same as in the NM case.[11]

With Monitoring of Effort, the principal can detect a shirking agent with some probability. After observing low output, detection is a sufficient signal that the current employee is indeed opportunistic. Unlike with MC, no detection is a sufficient signal of having an honest type (a-posteriori probability of having an honest type is higher than the proportion of the honest type on the market) if the detection probability $\delta_e > q$. When $e_H, \{e_H C\}$ is induced, ME always returns "no detection" as the opportunist is exerting high effort and thus ME cannot help to obtain additional information. Therefore, in this case, the principal will always fire after observing low output. Altogether, with ME, the principal's firing rule will be based directly on the outcome of the monitoring.[12]

This result might seem surprising at first, but it is in fact a consequence of the structure of the model. Recall that only those opportunists who exert high effort and are <u>lucky</u> can collect bribes (as a part of the "above-the-minimum" output) which, as I argued at the beginning, is indeed a realistic assumption. Thus, corruptibility is conditional on good luck. Consequently, with MC, some opportunists are not detected because the detection

[11]Even though the information generated by the monitoring in this case is not sufficient to improve the sorting of the types (and thus the informational value of the test might seem negligible), it will be shown later that Monitoring of Corruption does affect the incentives of the agent in the desired way and therefore the principal might want to commit to conducting it.

[12]This improvement in sorting is a possible value added to the principal. The principal, for whom non-corruptibility of her agents is the top priority, might want to bear the extra cost (it will be shown later that the principal has to pay higher expected wages with ME) connected to ME in order to sort out the opportunists. As I said, it is not the purpose of this paper to specify the optimal contract of the principal but rather to explore the possible effects of monitoring on agents' incentives.

technology has failed, others are not detected because "luck did not bring them enough opportunities" to be corrupt. With ME, every opportunist who chooses to exert low effort can be detected. Thus, with MC a smaller proportion of opportunists is detected and therefore, no detection is not a sufficient signal for the principal to keep the agent.

Assumption 2 *From now on, I will assume that the principal has access to a Monitoring of Effort technology that is successful enough or, that* $\delta_e > q$.

Table 1.5 summarizes the firing rules for all three monitoring technologies. The principal will ex ante commit to these firing rules in the contract. Ex post, the principal cannot be better off not keeping her ex-ante rule.

	$e_H, \{e_H NC\}$		$e_H, \{e_H C\}$			$e_H, \{e_L NC\}$		
	B_H	B_L	B_H	$B_L + ND$	$B_L + D$	B_H	$B_L + ND$	$B_L + D$
MN	keep	fire	keep	fire		keep	fire	
MC	keep	fire	keep	fire	fire	keep	fire	
ME	keep	keep	keep	fire		keep	keep	fire

Table 1.5: Summary of the firing rules with all three monitoring technologies: NM, MC, and ME; e_H/e_L stand for high/low effort, NC/C for non-corrupt/corrupt, B_H/B_L for high/low output, and D/ND for "detection"/"no detection."

Assumption 3 *From now on, I will assume that the expected gain from corruption is greater than the cost of exerting high effort, thus* $qA > c$.

The further results of the model depend on how the expected gain from corruption compares to the cost of exerting high effort. Depending on that, the opportunist might prefer high effort and corruptibility to exerting low effort or vice versa, which has important consequences for the implementability of various strategy profiles. Therefore, I will distinguish two cases, when $qA \leq c$ and when $qA > c$.

The first case is less interesting because in the benchmark case, with No Monitoring, corruption does not exist. Recall that both $\{e_H C\}$ and $\{e_L NC\}$ lead to low observed output and thereby to firing with certainty. As the expected (net) gain of corruption $qA - c$ is negative, the opportunist will be better off exerting low effort than exerting high effort and collecting bribes. Therefore, corruption is not an issue even without monitoring and the principal-agent interaction becomes, in this case, just a simple problem of effort choice.

The second case is more interesting: corruption is not suppressed by the choice of parameters and can occur with all three monitoring technologies. In fact, as the expected (net) gain of corruption $qA - c$ is positive, the opportunist (unless he is provided extra incentives), prefers $e_H C$ to $e_L NC$ in a single period. Given the interdependence of effort choice and corruptibility, this is an interesting case when the interplay of the incentives to extract bribes and to exert effort becomes an issue. As the purpose of this chapter is to examine various effects of monitoring technologies on agent's corruptibility and effort level, I will concentrate on this case.

Proposition 2 summarizes the main result for Monitoring of Corruption assuming that the principal (exogenously) commits to monitoring and to the optimal firing rule as specified above.

Proposition 2 *The wages sufficient to ensure both high effort and non-corruptibility over both periods are lower with Monitoring of Corruption than with No Monitoring.*

The proof can be found in the appendix. Intuitively, the principal has two main channels of influencing the agent's incentives: the threat of firing after the first period and an incentive-compatible payment scheme (including expected penalty). As to the first, MC does not improve the sorting of types and the firing rule of the principal does not change compared to NM: the principal keeps the agent only after observing high output; low observed output leads to firing with certainty. With an unchanged firing rule, the risk of getting fired is exactly the same as with NM. Consequently, there is no "Cremer-like" (negative) effect on the incentives of the agent that would result from the reduced risk of getting fired.[13] As to the second, in the case of detection, the agent is fired and penalized. Therefore, the threat of penalty serves as an additional enforcement mechanism compared to the NM case. Consequently a lower wage compensation is sufficient from the principal to induce non-corruptibility.

To put it differently, with MC, "No Corruption" is the binding constraint. With $qA > c$, with the threat of getting fired after delivering low output, and with binary effort choice, the principal offers a wage premium to induce non-corruptibility as opposed to corruptibility (and not to induce high effort as

[13]Recall that according to Cremer (1995) additional information prevents the principal from committing to some threats and therefore might be detrimental to the agent's incentives.

opposed to low effort). Once the principal bans corruption (imagine $\delta = 1$), she no longer needs to "bribe" the agent not to take bribes and therefore she can lower the wage a little bit and still ensure his non-corruptibility. Thus, MC allows maintaining high effort and non-corruptibility with a lower wage.

The presence of the expected penalty allows the principal to change the wage structure, which results in lower expected wage for the agent but the threat of firing after the first period maintains the incentives to exert effort.[14]

Monitoring of Effort is more interesting, as in this case, the principal's firing rule is based on the result of monitoring (and it is different than with NM), which is why a "Cremer-like" (negative) effect on agent's incentives can be expected. *Proposition 3* summarizes the main result for ME, assuming that the principal (exogenously) commits to monitoring and to the optimal firing rule as specified above.

Proposition 3 *The wages necessary to ensure both high effort and non-corruptibility over both periods are higher with Monitoring of Effort than with No Monitoring.*

A detailed proof can be found in the appendix. When the principal wants to induce high effort and non-corruptibility over two periods from both types of agent, with No Monitoring the wages $w_{HH} = 2A - (qA - c)$, $w_{HL} = A - (qA - c)$, and $w_F = 0$ are sufficient; with Monitoring of Effort, she has to offer $w_{HH} = 2A$, $w_{HL} = A$, $w_{LL} = 0$, and $w_F = 0$ to induce the same strategy profile. It follows immediately from *Assumption 1* that the No-Monitoring wages are lower.

This result has the flavor of Cremer's (1995) main result: additional information to the principal about the agent makes some threats not credible and thereby weakens the incentives of the agent. Consequently, in order to induce the same strategy profile, the principal has to compensate with higher wages. The optimal firing rule for ME is different than for NM, as "no detection"

[14] A simple exercise can be done to illustrate how MC affects the incentives of the agent. Assume that the principal would commit to the NM technology but with (lower) MC-incentive-compatible wages. All the other conditions of the contract would remain unchanged. It is easy to show that in such a case the opportunist would switch to $\{e_H C\}$ in the first period (and get fired afterwards). Importantly, the absence of monitoring and of the expected penalty results in increased corruption. Given that $qA > c$ and that both $\{e_H C\}$ and $\{e_L NC\}$ lead to firing with certainty, the agent prefers $\{e_H C\}$ to $\{e_L NC\}$ in the first period. Either higher wages or the threat of penalty are necessary to ensure the non-corruptibility of the agent.

is a sufficient signal for the principal to keep the agent even after delivering low output. Consequently, as in Cremer (1995), the principal cannot credibly threaten to fire every agent who has delivered low output, which weakens the incentives of the agent.

ME affects the incentives of the honest type to exert high effort and the incentives of the opportunistic type to be non-corrupt. The binding constraint is, again, the one ensuring non-corruptibility of the opportunistic type. Let me illustrate why.

First, assume the extreme case when the monitoring technology is perfect, and thus $\delta = 1$. The honest type will continue exerting high effort, because otherwise he would be detected and fired with certainty. The opportunist, however, even when $\delta = 1$, has incentives to switch to $\{e_H C\}$ and here is why.[15] With NM, going from $\{e_H NC\}$ to $\{e_H C\}$ would increase his probability of getting fired from $(1 - q)$ to 1 as $\{e_H C\}$ leads to low observed output and firing. With ME, on the other hand, his probability of getting fired would drop to zero, as after observing low output the principal would monitor the agent and the test for low effort would return "no detection." Clearly, the incentives to switch to $\{e_H C\}$ are stronger with ME than with NM, and that is why the principal has to "bribe" the agent not to take bribes.

Second, when the monitoring technology is not perfect, and thus $\delta < 1$, the incentives of the honest type will be affected. In this case, the honest type's probability of getting fired after exerting low effort would be 1 with NM (because he delivers low output he is fired), whereas it is only $\delta < 1$ with ME. Therefore, as in Cremer (1995), the principal has to compensate what is lost on incentives due to the "lower threat of firing" by offering a higher wage to the agent.

The fact that the binding constraint is that on non-corruptibility follows from the assumption that $qA > c$ (*Assumption 3*) and thus the opportunist always prefers $\{e_H C\}$ to $\{e_L NC\}$.[16] Consequently, when the constraint on non-corruptibility is binding, the incentive-compatibility constraint of the honest type to exert effort is satisfied automatically. If we were to reduce the

[15] Assume that the principal offers wages to induce high effort and non-corruptibility over both periods. I want to compare the incentives of the opportunist to switch $\{e_H C\}$ in the first period under NM vs. ME. Also, keep in mind *Assumption 3* that $qA > C$, which ensures that the opportunist always prefers high effort and corruption to exerting low effort.

[16] This assumption ensures that the expected gain from corruption is sufficient to compensate for the cost of high effort and therefore corruption is indeed a problem.

wages, such that the effort constraint would be binding, the opportunistic type would switch to corruption (see *Proposition 4* below). Importantly, the higher wages that the principal has to offer with ME to ensure that the opportunistic type is non-corrupt are high enough to also ensure that the honest type exerts high effort.

To support the above arguments and to illustrate how exactly the incentives of the two types are affected by ME, I also did a simple exercise looking at a hybrid monitoring technology which combines some properties of No Monitoring and of Monitoring of Effort. The results are summarized in *Proposition 4*.

Proposition 4 *ME can negatively affect the opportunist's incentives to be non-corrupt.*

a) When the principal maintains the same $\{e_H NC, e_H NC\}$ wages as with NM while introducing ME, the opportunist's strategy profile will involve corruption. The incentives of the honest type are not distorted.

b) When the principal implements ME but cannot offer higher expected wages than with NM,[17] the best strategy profile $\{e_H NC, e_H NC\}$ is no longer implementable.

The detailed proof can be found in the appendix. When the principal offers the NM wages, $w_{HH} = 2A - (qA - c)$, $w_{HL} = A - (qA - c)$, and $w_F = 0$, to induce $\{e_H NC, e_H NC\}$ but introduces ME,[18] the opportunist no longer prefers $\{e_H NC, e_H NC\}$. Instead, he will prefer $\{e_H NC, e_H C\}$. Thus, when the principal introduces ME with (lower) NM wages, the second-period incentives are distorted and, specifically, they invoke corruption. Intuitively, the second period incentives are the first to be affected by lower wages, as there is no monitoring after the second period. Given that $qA > c$ the opportunist always prefers $\{e_H C\}$ to $\{e_L NC\}$ in a one-period horizon. As lower wages are weakening the incentives of the opportunist, he will naturally switch to $e_H C$ rather than to $\{e_L NC\}$ to be able to extract extra $qA - c > 0$. The incentives of the honest type to exert effort are not distorted, which is,

[17]Thus, the contract that the principal offers has all the properties of ME but her budget constraint does not allow her to pay more (in expectations) than she would with NM.

[18]Thus, the contract that the principal offers has all the properties of ME, just the wages that she offers are the NM-incentive-compatible wages.

as I discussed above, a consequence of $qA > c$, an assumption which ensures that when reducing the wages, it is first the constraint on non-corruptibility that breaks down, while the effort constraint is still satisfied.

The proof of part *b)* can be also found in the appendix. When the principal implements Monitoring of Effort but does not want to spend, in expectation, more than with No Monitoring, the best strategy profile $\{e_H NC, e_H NC\}$ is no longer implementable. Therefore, the principal, depending on her preferences, would have to offer a contract that would induce an alternative strategy profile.

1.4 Conclusion

In the two-period principal-agent model in which an agent of unknown propensity to corruption decides about his effort level and corruptibility (both his action choices hidden to the principal), I show that imperfect partial Monitoring of Effort with a high enough detection rate improves the sorting of types. It might, however, also support agent's incentive to be corrupt. In contrast, Monitoring of Corruption does not improve the sorting of types and it negatively affects the incentive of the agent to be corrupt.

In particular, I show that when the expected gain from corruption is high enough and corruption might exist with No Monitoring, it is more expensive for the principal to induce a high level of effort and non-corruptibility over two periods with Monitoring of Effort than with No Monitoring. Monitoring of Effort positively affects the incentive of the agent to be corrupt. This result is in line with earlier findings in the career concerns literature (Cremer 1995; Holmström 1999; Dewatripont, Jewitt, and Tirole 1999; Prat 2005) that additional information to the principal about her agent prevents the principal to credibly commit to some threats and thereby might weaken the incentives of the agent to exert high effort in order to signal high ability. In contrast, Monitoring of Corruption, even though it does not improve the sorting of types – or, in fact, thanks to it – does not have this detrimental effect on the agent's incentives. With Monitoring of Corruption, the principal does not rely on the result of monitoring (specifically, "no detection" is not a sufficient signal to keep the agent for the second period) and therefore, the firing rule is the same as with No Monitoring. Consequently, there is no "Cremer-like" negative impact on the agent's incentives. As the expected penalty serves as an additional enforcement mechanism, lower expected wages are necessary to induce high effort and non-corruptibility over two periods than with No Monitoring (and than with Monitoring of Effort).

A policy implication can be drawn from the results. When the agent's propensity to corruption is hidden to the principal, and when the agent's actions are not directly observable and might also depend on exogenous random realization (luck), the principal might want to rely on an (stricter) output-contingent firing rule and monitor for corruption even though this monitoring technology does not improve the sorting of types. Monitoring of Effort is inferior to both Monitoring of Corruption and No Monitoring, in that it distorts the incentives to be non-corrupt. This conclusion is drawn for the principal who wants to induce the most efficient strategy profile that involves high effort and non-corruptibility over both periods.

A conclusion can be drawn also for the case of Czech traffic police officers that motivated this investigation. The cameras installed in the police cars, even if one would doubt their effectiveness as a mean of detecting corruption and providing admissible evidence, could help to improve the officers' incentives when implemented together with a properly designed reward and punishment system.

References

Benabou, R., Tirole, J. (2003). Intrinsic and Extrinsic Motivation. *Review of Economic Studies, 70*, 489-520.

Celentani, M., Ganuza, J. (2002). Corruption and Competition in Procurement. *European Economic Review, 46*, 1273-1303.

Cremer, J. (1995). Arm's Length Relationships. *The Quarterly Journal of Economics, 110 (2)*, 275-295.

Cule, M., Fulton, M. (2005). Some Implications of the Unofficial Economy–Bureaucratic Corruption Relationship in Transition Countries. *Economics Letters, 89*, 207–211.

Dewatripont, M. (1988). Commitment through Renegotiation-Proof Contracts with Third Parties. *Review of Economic Studies, LV*, 377-90.

Dewatripont, M., Jewitt, I., Tirole, J. (1999). The Economics of Career Concerns, Part I: Comparing Information Structures. *Review of Economic Studies 66(1)*, 183—198.

Holmström, B. (1999). Managerial Incentive Problems: A Dynamic Perspective. *Review of Economic Studies 66(1)*, 169—182.

Kreps, D. (1997). Intrinsic Motivation and Extrinsic Incentives. *American Economic Review, 87*, 359-364.

Prat, A. (2005). The Wrong Kind of Transparency. *American Economic Review, 95(3)*, 862-877.

APPENDIX

Proof of Proposition 1. With all three monitoring technologies, there are three possible first-period strategy profiles for the opportunist: $e_H C$, $e_H NC$, $e_L NC$. According to *Assumption 1*, the principal always prefers the honest type to exert high effort. The principal can offer various contracts (various wages to induce various two-period strategy profiles of the opportunist), each of which would involve one of the above-mentioned first-period strategies.

For each monitoring technology, I compute the probability of having an honest type given the induced first-period strategy profile and observed outcome of the first period (including monitoring, if relevant). Recall the Bayesian updating rule:

$$P(honest/outcome) = \frac{P[honest] \cdot P[outcome/honest]}{P[outcome]}^{19}$$

<u>NM:</u> With NM, the only new information the principal has after the first period is the output realization.

1) If $e_H, \{e_H NC\}$ is induced, then $P[honest/B_L] = \frac{(1-q)p}{(1-q)} = p$ and $P[honest/B_H] = \frac{qp}{q} = p$. Thus the a-posteriori probability of having an honest type is the same as the a-priori probability (and as the probability of the honest type on the market) after both B_H and B_L. Therefore in this case, the principal is indifferent between keeping and firing her current employee. $P[honest/B_H] = \frac{qp}{q} = p$

2) If $e_H, \{e_H C\}$ is induced, then $P[honest/B_L] = \frac{(1-q)p}{(1-pq)} < p$. Thus the the a-posteriori probability of having an honest type is lower than the a-priori probability. More importantly, it is lower than the probability of obtaining an honest type on the market. Therefore in this case, the principal would be better off firing her current employee and hiring a new one for the second period after observing B_L. As $P[honest/B_H] = 1$, the principal wants to keep her current employee after observing B_H.

3) If $e_H, \{e_L NC\}$ is induced, then $P[honest/B_L] = \frac{(1-q)p}{(1-pq)} < p$. Thus the the a-posteriori probability of having an honest type is lower than the a-priori probability (and than the probability of obtaining an honest type

[19] Note that Table 1.3 provides a review of all the possible combinations of the agent's type, decisions, luck, and observed output after the first period, which can be useful for the computations of these probabilities.

on the market). Therefore in this case, the principal would be better off firing her current employee and hiring a new one for the second period after observing B_L. As $P[honest/B_H] = 1$, the principal wants to keep her current employee after observing B_H.

All in all, the principal will fire her current employee after observing low output and keep the agent otherwise.

MC: With MC, the principal observes output and the result of Monitoring of Corruption after the first period. Therefore, the "outcome" has now two components: "observed output" and "detection (D) or no detection (ND)." Note that the honest type is never corrupt and therefore never detected and thus $P[honest/B_L + D] = 0$ always (and $P[opportunistic/B_L + D] = 1$ so detection will always lead to firing). Also, the principal does not monitor after observing B_H as a "misbehaving" agent never produces high output.

1) If $e_H, \{e_H NC\}$ is induced, then $P[honest/B_L + ND] = \frac{(1-q)p}{(1-q)} = p$, and $P[honest/B_H] = p$. Thus in this case, the principal is indifferent between keeping and firing her current employee.

2) If $e_H, \{e_H C\}$ is induced, then $P[honest/B_L + ND] = \frac{(1-q)p}{(1-pq-q\delta_c+pq\delta_c)} < p$, and $P[honest/B_H] = 1$. Thus in this case, the principal would be better off firing her current employee after observing low output (with detection as well as without) and hiring a new one for the second period. She will only keep her current employee after observing high output.

3) If $e_H, \{e_L NC\}$ is induced, then $P[honest/B_L + ND] = \frac{(1-q)p}{(1-pq)} < p$, and $P[honest/B_H] = 1$. Thus in this case, the principal would be better off firing her current employee after observing low output (with detection as well as without) and hiring a new one for the second period. She will only keep her current employee after observing high output.

All in all, the principal's firing rule will be the same as in the NM case. Even though the updated probabilities of having an honest type after the first period are not the same as with NM, the principal does not acquire sufficient information to base the firing upon the result of the monitoring. Or, to say it differently, no detection is not a sufficient signal of having an honest type. Therefore, similarly as with NM, the principal will fire her current employee after observing low output and keep the agent otherwise.

ME: With ME, the principal observes output and the result of Monitoring of Effort after the first period. Therefore, the "outcome" now has two components: "observed output" and "detection (D) or no detection (ND)."

Note that the principal always prefers the honest type to exert high effort and therefore $P[honest/B_L + D] = 0$ always (and $P[opportunistic/B_L + D] = 1$ so detection will always lead to firing). Also, the principal does not monitor after observing B_H as a "misbehaving" agent never produces high output.

1) If $e_H, \{e_H NC\}$ is induced, then $P[honest/B_L + ND] = \frac{(1-q)p}{(1-q)} = p$, and $P[honest/B_H] = p$. Thus in this case, the principal is indifferent between keeping and firing her current employee.

2) If $e_H, \{e_H C\}$ is induced, then $P[honest/B_L + ND] = \frac{(1-q)p}{(1-pq)} < p$, and $P[honest/B_H] = 1$. Thus in this case, the principal would be better off firing her current employee after observing low output (with detection as well as without) and hiring a new one for the second period. She will only keep her current employee after observing high output.

3) If $e_H, \{e_L NC\}$ is induced, then $P[honest/B_L + ND] = \frac{(1-q)p}{(1-\delta_e+p\delta_e-pq)} > p$ if $\delta_e > q$, and $P[honest/B_H] = 1$. Thus in this case, the principal would be better off keeping her current employee after observing low output and no detection as well as after observing high output. She will only fire after detection.

All in all, with ME, the principal can base her firing rule on the outcome of the monitoring. When $e_H C$ is induced then low output is a sufficient signal for firing the agent. When $e_L NC$ is induced, then no detection is a sufficient signal of having an honest type if $\delta_e > q$. Therefore, the principal will fire her current employee in two cases: if a) $e_H C$ is induced and B_L is observed, or b) $e_L NC$ is induced and $B_L + D$ is observed. Thus, ME improves the sorting of types.

Proof of Proposition 2. For both NM and MC I will set up the principal's problems. The principal minimizes the expected wage such that the strategy profile $\{e_H NC, e_H NC\}$ from both types is incentive-compatible.

NM: Recall that with NM after delivering low output in the first period the agent is fired with certainty. The principal maximizes

$$-q[qw_{HH} + (1-q)w_{HL}] - (1-q)w_F.$$

The second-period incentives of the honest type are ensured by

$$qw_{HH} + (1-q)w_{HL} - c \geq w_{HL}.$$

The first-period incentives of the honest type are ensured by

$$q[qw_{HH} + (1-q)w_{HL} - c] + (1-q)w_F - c \geq w_F.$$

The opportunist must, in addition, prefer being non-corrupt to being corrupt. The opportunist's second-period incentives are therefore ensured by

$$qw_{HH} + (1-q)w_{HL} - c \geq w_{HL},$$

$$qw_{HH} + (1-q)w_{HL} - c \geq w_{HL} + qA - c.$$

The first period incentives of the opportunist are ensured by

$$q[qw_{HH} + (1-q)w_{HL} - c] + (1-q)w_F - c \geq w_F,$$

$$q[qw_{HH} + (1-q)w_{HL} - c] + (1-q)w_F - c \geq w_F + qA - c.$$

The agent's (ex-ante) participation constraint is

$$q[qw_{HH} + (1-q)w_{HL} - c] + (1-q)w_F - c \geq 0.$$

The participation constraint does not depend on probability p as both types are induced the same strategy profile. It is easy to show that given the incentive-compatibility and non-negativity constraints on wages, the agents' participation constraint holds as well.

The optimal wages satisfying the incentive-compatibility constraints are

$$w_{HH} = (2-q)A + c, w_{HL} = (1-q)A + c, w_F = 0.$$

MC: Recall that with MC the firing rule is the same as with NM. The principal maximizes

$$-q[qw_{HH} + (1-q)w_{HL}] - (1-q)w_F.$$

The incentives of the honest agent are ensured by

$$qw_{HH} + (1-q)w_{HL} - c \geq w_{HL},$$

$$q^2 w_{HH} + q(1-q)w_{HL} + (1-q)w_F - (1+q)c \geq w_F.$$

The incentives of the opportunist are ensured by

$$qw_{HH} + (1-q)w_{HL} - c \geq w_{HL},$$

$$qw_{HH} + (1-q)w_{HL} - c \geq w_{HL} + qA - c,$$

$$q^2 w_{HH} + q(1-q)w_{HL} + (1-q)w_F - (1+q)c \geq w_F,$$

$$q^2 w_{HH} + q(1-q)w_{HL} + (1-q)w_F - (1+q)c \geq w_F + q(A - \delta_c F) - c.$$

The agent's (ex-ante) participation constraint is

$$q[qw_{HH} + (1-q)w_{HL} - c] + (1-q)w_F - c \geq 0.$$

The participation constraint does not depend on probability p as both types are induced the same strategy profile. It is easy to show that given the incentive-compatibility and non-negativity constraints on wages, the agents' participation constraint holds as well.

The optimal wages satisfying the incentive-compatibility constraints are:

i)for $c \geq q(A - \delta_c F)$ and
ia) for $A \leq \frac{(1+q)c}{q^2} : w_{HH} = (1-q)A + \frac{(1+q)c}{q}, w_{HL} = -qA + \frac{(1+q)c}{q}, w_F = 0,$
ib) for $A > \frac{(1+q)c}{q^2} : w_{HH} = A, w_{HL} = w_F = 0.$

ii) for $c < q(A - \delta_c F)$ and
iia) for $A \leq \frac{(1+q)c}{q^2} or A > \frac{(1+q)c}{q^2}$ & $F \leq \frac{(1-q)A+c}{\delta_c}$:
$w_{HH} = (2-q)A + c - \delta_c F, w_{HL} = (1-q)A + c - \delta_c F, w_F = 0$, or
iib) for $A > \frac{(1+q)c}{q^2}$ & $F > \frac{(1-q)A+c}{\delta_c} : w_{HH} = A, w_{HL} = w_F = 0.$

As $qA > c$, it is easy to show that all the MC wages are lower than wages with MN. Also the ex-ante expected wage cost to the principal is lower with MC for all parameters. Thus, it is more expensive for the principal to induce the same strategy profile with NM than with MC.

Proof of Proposition 3. Now I need to set up the principal's problem for ME. The principal minimizes the expected wage such that the strategy profile $\{e_H NC, e_H NC\}$ from both types is incentive-compatible.

NM: Recall that with NM, the optimal wages satisfying the incentive-compatibility constraints are

$$w_{HH} = (2-q)A + c, w_{HL} = (1-q)A + c, w_F = 0.$$

ME: Recall that with ME the principal fires the agent after he exerted low effort and was detected, or after observing B_L when e_H, $\{e_H C\}$ is induced in the first period. The principal maximizes

$$-q[qw_{HH} + (1-q)w_{HL}] - (1-q)[qw_{HL} + (1-q)w_{LL}].$$

The incentives of the honest agent are ensured by

$$qw_{HH} + (1-q)w_{HL} - c \geq w_{HL},$$

$$qw_{HL} + (1-q)w_{LL} - c \geq w_{LL},$$

$$q^2 w_{HH} + 2q(1-q)w_{HL} + (1-q)^2 w_{LL} - 2c \geq \delta_e w_F + (1-\delta_e)(qw_{HL} + (1-q)w_{LL} - c).$$

The incentives of the opportunist are ensured by

$$qw_{HH} + (1-q)w_{HL} - c \geq w_{HL},$$

$$qw_{HL} + (1-q)w_{LL} - c \geq w_{LL},$$

$$qw_{HH} + (1-q)w_{HL} - c \geq w_{HL} + qA - c,$$
$$qw_{HL} + (1-q)w_{LL} - c \geq w_{LL} + qA - c,$$
$$q^2 w_{HH} + 2q(1-q)w_{HL} + (1-q)^2 w_{LL} - 2c \geq w_F + qA - c,$$
$$q^2 w_{HH} + 2q(1-q)w_{HL} + (1-q)^2 w_{LL} - 2c \geq \delta_e w_F + (1-\delta_e)(qw_{HL} + (1-q)w_{LL} - c).$$

The agent's (ex-ante) participation constraint is

$$q^2 w_{HH} + 2q(1-q)w_{HL} + (1-q)^2 w_{LL} - 2c \geq 0.$$

The participation constraint does not depend on probability p as both types are induced the same strategy profile. It is easy to show that given the incentive-compatibility and non-negativity constraints on wages, the agents' participation constraint holds as well.

The optimal wages satisfying the incentive-compatibility constraints are

$$w_{HH} = 2A, w_{HL} = A, w_{LL} = 0, w_F = 0.$$

As $qA > c$, it is easy to show that wages with NM are lower than wages with ME. Also, the ex-ante expected wage cost to the principal is lower with NM. Thus, it is more expensive for the principal to induce the same strategy profile with ME than with NM.

Proof of Proposition 4. **a)** First, I will show that if the principal offers NM wages but introduces Monitoring of Effort the opportunist's incentives to be non-corrupt are weakened.

Assume that the principal offers NM wages $w_{HH} = (2-q)A + c, w_{HL} = (1-q)A + c$, and $w_{LL} = w_F = 0$.

a) If the principal introduces ME, the opportunist's expected payoffs from all possible strategy profiles will be as in Table 1.6.

strategy profile	expected 2-period payoff	expected 2-period payoff with NM wages
$\{e_H NC, e_H NC\}$	$q^2 w_{HH} + 2q(1-q)w_{HL} + (1-q)^2 w_{LL} - 2c$	$(q^2 - 2q + 2)(qA - c)$
$\{e_H NC, e_H C\}$	$qw_{HL} + (1-q)w_{LL} + qA - 2c$	$(2-q)(qA - c)$
$\{e_H NC, e_L NC\}$	$qw_{HL} + (1-q)w_{LL} - c$	$(1-q)(qA - c)$
$\{e_L NC, e_H NC\}$	$\delta_e w_F + (1-\delta_e)(qw_{HL} + (1-q)w_{LL} - c)$	$(1-\delta_e)(1-q)(qA - c)$
$\{e_L NC, e_H C\}$	$\delta_e w_F + (1-\delta_e)(w_{LL} + qA - c)$	$(1-\delta_e)(qA - c)$
$\{e_L NC, e_L NC\}$	$\delta_e w_F + (1-\delta_e)w_{LL}$	0
$\{e_H C, --\}$	$w_F + qA - c$	$qA - c$

Table 1.6: The opportunist's expected wages for all possible strategy profiles when ME is introduced with NM wages.

It is easy to show that (as $0 < q < 1, 0 < \delta_e < 1, c > 0, A > 0$, and $qA > c$) the opportunist will ex-ante prefer the strategy profile $\{e_H NC, e_H C\}$. The strategy space of the honest type is a subset of the strategy space of the opportunist. For the honest type only those strategy profiles that do not involve corruption are relevant. It is easy to check that the incentives of the honest type are not affected - he will still prefer the strategy profile $\{e_H NC, e_H NC\}$.

Now, we also need to check the ex-post incentives of the agent who is standing at the beginning of the second period. Table 1.7 summarizes the expected (two-period) payoffs contingent on the first-period outcome and on the results of monitoring.[20]

1^{st}-period strategy & outcome	2nd-period strategy	expected 2-period payoff	expected 2-period payoff with NM wages
e_H	$e_H NC$	$qw_{HL} + (1-q)w_{LL} - c$	$(1-q)(qA-c)$
NOT lucky $\Rightarrow B_L$	$e_H C$	$w_{LL} + (qA-c)$	$(qA-c)$
(or e_L)	e_L	w_{LL}	0
e_H	$e_H NC$	$qw_{HH} + (1-q)w_{HL} - c$	A
lucky $\Rightarrow B_H$	$e_H C$	$w_{HL} + (qA-c)$	A
non-corrupt	e_L	w_{HL}	$A-(qA-c)$

Table 1.7: The opportunist's expected wages for all possible strategy profiles when MC is introduced with NM wages.

It is easy to show that the opportunist cannot be better of by not keeping to his ex-ante chosen strategy profile. For the honest type, his strategy space consists of the strategies that do not involve corruption. It is easy to show that the second-period incentives of the honest type are not affected. Thus, with ME the opportunist now prefers the strategy profile $\{e_H NC, e_H C\}$.

b) To prove this part of the proposition, we solve the same problem as in the proof of *Proposition 3*, just with additional budget constraints of the principal. With NM, the principal's expected cost is $q^2 w_{HH} + q(1-q)w_{HL} + (1-q)w_F$, which, with optimal wages $w_{HH} = (2-q)A+c, w_{HL} = (1-q)A+c$, and $w_F = 0$, is equal to $q(A+c)$. Thus, the additional budget constraint for ME is $q^2 w_{HH} + 2q(1-q)w_{HL} + (1-q)^2 w_{LL} \leq q(A+c)$. With ME and the additional budget constraint, the strategy profile $\{e_H NC, e_H NC\}$ is no longer implementable.

[20] Recall that being corrupt in the first period, the opportunist delivers low output and therefore is fired. Therefore, such a first-period strategy is not relevant for an agent deciding about his second-period strategy.

Chapter 2

Testing Leniency Programs Experimentally: The Impact of Change in Parameterization

Abstract

I analyze subjects' sensitivity to parametric change that does not affect the theoretical prediction. I find that increasing the value of an illegal transaction to a briber and reducing the penalties to both culprits leads to more bribes being paid but does not affect the cooperation of the bribee. My data also suggest that trust and preferences towards others might play a role. This chapter provides a testbed for experimental testing of anti-corruption measures and adds evidence to the on-going discussion on the need for socio-demographic controls.

Keywords: corruption, anti-corruption mechanisms, optimal contract, monitoring
JEL classification: C91, D02, D73, K42

2.1 Introduction

The severe consequences of corruption have been documented in numerous empirical studies. For example, Mauro (1995) and Tanzi (1998) have shown a negative effect of corruption on economic growth; Hwang (2002) has demonstrated that corruption, through tax evasion, reduces government revenues; and Gupta, Davoodi and Alonso-Terme (2002) have shown that corruption increases income inequality and poverty. The design and implementation of effective anti-corruption measures therefore remains an important concern.

One promising anti-corruption measure is the leniency policy. Leniency policies award fine reductions of varying intensities to wrongdoers who "spontaneously" report an illegal agreement and thereby help to convict their accomplice(s). They serve as an enforcement mechanism as much as a means of deterrence in that, if appropriately designed and implemented, they have the potential to undermine the trust between wrongdoers. Leniency policies have been analyzed in the literature mostly as an anti-cartel mechanism.

The deterrence effect of leniency policies in the case of cartels has been analyzed – and confirmed – both theoretically (e.g. Spagnolo 2004) and experimentally (e.g. Apesteguia, Dufwenberg and Selten 2004; Bigoni, Fridolfsson, Le Coq and Spagnolo 2008 a,b).

Spagnolo (2004), for example, theoretically examines the effects of leniency policies of various degrees – from moderate (which reduce or cancel the penalty for a criminal who reports) to full (which, in addition, pay a reward). He shows that reward-paying leniency programs provide a (socially) costless[1] and very efficient measure for cartel deterrence.

Drawing on earlier versions of Spagnolo (2004), Apesteguia, Dufwenberg and Selten (2004) conducted an experiment that confirms the promising cartel-deterring properties of leniency policies. Bigoni, Fridolfsson, Le Coq and Spagnolo (2008 a,b) conducted related experiments. In addition to confirming the basic results about the effectiveness of leniency programs, they attempted to acquire a deeper understanding of the driving forces. In several treatments they vary specific features of the game – fine levels, exogenous risk of detection, reward schemes, possibility to communicate, and the eligibility for leniency. They control for past convictions and for subjects' risk attitudes. The experiments are run in Stockholm and in Rome, which allows

[1] This is the case if the rewards are fully financed from fines imposed on other convicted members of the cartel.

assessing potential cultural effects.

Bigoni et al. (2008a) find that leniency leads to higher deterrence but at the same time helps to sustain higher prices. Rewards lead to almost complete deterrence – which is in line with Spagnolo's (2004) result. Past convictions reduce the number of cartels but increase collusive prices. The authors, in addition, distinguish between two types of past convictions: 1) those that occur as a result of reporting and 2) those that occur as a result of external investigation. They find that past convictions after reporting have a much stronger deterrence effect than past convictions after external investigation. The results also confirm a strong cultural effect.

Bigoni et al. (2008b) focuses on the role of risk attitudes. They find that risk aversion and the willingness to form a cartel are negatively correlated. The results suggest that past experience might have more important consequences for the perception of risk than the exogenous probability of detection, and that the strategic risk (the risk of being cheated upon) plays a key role for the effectiveness of a leniency policy.

Both Bigoni et al. (2008 a,b) contribute to a better understanding of the cartel-deterring properties of leniency policies and highlight the importance of proper policy design.

Leniency policies to deter cartels are, however, not directly applicable as anti-corruption measures, since cartel deterrence is essentially a simultaneous game while strategies, payoffs, and the move structure of anti-corruption measures are asymmetric.[2] A proper theoretical and experimental analysis is therefore called for.

To my knowledge the first theoretical work analyzing the various effects of leniency policies on corruption is Buccirossi and Spagnolo (2006). The authors show that poorly designed moderate policies may have a serious counter-productive effect: they might allow a briber to punish at relatively low cost a partner who does not respect an illegal agreement. In other words, some leniency policies might actually provide an enforcement mechanism for occasional illegal transactions.[3] Thus they can, contrary to the intention, increase corruption.

Buccirossi and Spagnolo's result together with the theoretical and experimental evidence from the literature on cartel deterrence suggests that the potential of leniency policies to undermine trust between wrongdoers hinges

[2]For a more detailed discussion see Richmanová (2006).

[3]Occasional illegal transactions are essentially one-shot transactions.

upon proper design and implementation.

Experimental methods have been widely used, albeit rarely, to study corruption (Dušek, Ortmann and Lízal 2005). They become especially useful when counterfactual institutional arrangements such as leniency programs need to be explored: they provide, for example, relatively cheap ways to examine the effects of such arrangements in controlled environments (see Dušek et al. 2005, Apesteguia et al. 2004, Buccirossi and Spagnolo 2006, Bigoni et al. 2008 a,b, Richmanová and Ortmann 2008, and also Roth 2002).

In Richmanová and Ortmann (2008), we proposed a generalization of the Buccirossi and Spagnolo (2006) model by introducing the probabilistic discovery of evidence.[4] Our generalization makes the model more realistic and more readily applicable for experimental testing without changing the qualitative results of Buccirossi and Spagnolo.

We use the generalized Buccirossi and Spagnolo model for the experimental testing of leniency policies as an anti-corruption measure. As we address two different methodological issues that (anti-)corruption experiments are afflicted with, our results are reported in two papers: Krajčová (2008) (also this chapter), and the closely related work reported in Krajčová and Ortmann (2008) (also Chapter 3). Both papers provide a new testbed for anti-corruption programs.

Altogether, we design three experimental treatments: a benchmark, which is common for both studies, Krajčová and Ortmann (2008) and this chapter, and in which all instructions are presented in completely neutral language; a context treatment, in which we use the same parameterization as in the benchmark but instructions are presented in full bribery context (Krajčová and Ortmann 2008); and a high-incentive treatment, which implements a new parameterization within neutral framing (this chapter).[5]

In Krajčová and Ortmann (2008), following the earlier work of Abbink and Hennig-Schmidt (2006), we study the effect of "loaded" instructions in a bribery experiment. Surprisingly, Abbink and Hennig-Schmidt find no signif-

---

[4]In the original model, Buccirossi and Spagnolo assume that the briber and bribee agree to produce hard evidence, which serves as a hostage. Without hard evidence being produced, the occasional illegal transaction is not enforceable. An audit, if it takes place, discovers the evidence with a probability of one. In Richmanová and Ortmann (2008), we argue that instead some evidence is created unintentionally and this can be discovered by an audit with some probability that is less than one.

[5]We have, in addition, designed some additional exploratory treatments which we use for a robustness check of the main results. See the appendix for more details.

icant impact of instructions framing. The authors conclude that this result may be caused by the nature of the game: it is very simple, and as it was designed to capture all the basic features of bribery, even with neutral wording, subjects may have deciphered what the experiment was about. Our bribery game includes stages where players can report their opponents and receive leniency, which makes it more complex and also potentially more susceptible to the non-neutral context. Therefore, it calls for a separate analysis. We find a strong gender effect - male and female participants react differently to the non-neutral context. The effect of context becomes significant once we allow for gender-specific coefficients. Thus, in contrast to the results of Abbink and Hennig-Schmidt (2006), we find that a bribery context indeed makes a (significant) difference.

In this chapter I study the effect of a change in parameterization. It has been documented in the literature that a change in parameterization that does not affect the theoretical prediction might indeed have consequences for the behavior of subjects in the lab (see e.g. Goeree and Holt 2001). Anti-corruption experiments might be particularly tricky, being sensitive to changes in design. In the generalized Buccirossi and Spagnolo game, the action bringing the highest possible payoff is also associated with a risk of considerable loss. Therefore, risk or loss attitudes are also likely to play a role. Altogether, I expect that subjects in the lab might not behave in accordance with the theoretical prediction, especially when the prediction is made under an assumption of risk neutrality.[6] The question I ask is whether by making corruption more attractive by i) increasing the potential gain and ii) reducing the penalty if bribery is discovered, I can induce more corruption in the lab even if the theoretical prediction suggests no change. I also study to what extent the subjects' decisions in the experiment can be explained by their basic socio-demographic characteristics.

I do indeed find a significant effect of parametric change. Even though the change I implemented has no consequences for the theoretical prediction, I observe much more corruption in the high-incentive than in the benchmark treatment. My data suggest that trust and preferences towards others might play a role. The econometric analysis provides limited evidence on the role of basic socio-demographic characteristics. I find no differences in how the parametric change affects the behavior of male and of female participants.

[6] In fact, in our data we observe deviations from the theoretical prediction in all three treatments.

Both Chapters 2 and 3 provide a testbed for the experimental testing of anti-corruption measures and add evidence to the on-going discussion on the need for socio-demographic controls.

The remainder of this chapter is organized as follows. In the next section I discuss the generalized Buccirossi and Spagnolo model in detail, and I also describe and compare two experimental treatments. Section 3 talks about experimental implementation and in Section 4 I review the results. Section 5 concludes.

2.2 Experimental Design

The experiment implements the bribery game in Richmanová and Ortmann (2008). An entrepreneur has an investment possibility of net present value v, if a bureaucrat is willing to perform an illegal action, *Action a*. For doing so, the bureaucrat may require compensation in the form of a bribe, b.

The timing of the game is as follows. First, the entrepreneur decides whether to *Pay* or *Not Pay* a bribe. If she does not pay a bribe, the game ends. If she does, the bureaucrat chooses one of three possible actions: *Denounce*, do *Nothing*,[7] or perform *Action a*.[8]

If the bureaucrat chooses *Denounce*, an audit is carried out. The audit may (with probability β, $\beta \in (0, 1)$) or may not (with probability $1 - \beta$) discover some evidence of bribery. If the bribery attempt is detected, the leniency policy guarantees that the bureaucrat will have to pay only a reduced fine whereas the entrepreneur will have to pay the full fine. In addition, bribe b is confiscated.[9] If the bribery is not detected, the bureaucrat will enjoy bribe b.

If the bureaucrat chooses *Nothing* or *Action a*, the entrepreneur has another move. In both cases, she may choose between *Denounce* and do *Noth-*

[7] *Nothing* denotes a passive action choice. For the bureaucrat, it means that he neither denounces nor respects (by providing the favor) the illegal agreement. For the entrepreneur, it means that she does not denounce in response to the bureaucrat's action.

[8] *Action a* means that the bureaucrat respects the illegal agreement and thus provides an (illegal) favor to the entrepreneur. That is, strictly speaking, not a corrupt action because it does not impose a negative externality on the public. According to Abbink, Irlenbusch and Renner (2002) it is not such a problem since people do not care much about the costs they impose on others.

[9] Note that in this case the illegal transaction has been detected without *Action a* being performed and therefore there is no gain to the entrepreneur to be confiscated.

ing.

If the entrepreneur chooses *Denounce* and the ensuing audit discovers evidence (which, again, happens with probability β), then she will have to pay a reduced fine whereas the bureaucrat will have to pay the full fine and, in addition, their illegal gains will be confiscated. If no evidence is discovered, both the bureaucrat and the entrepreneur will keep their illegal gains.

If the entrepreneur chooses *Nothing*, then an audit may still occur with some nonzero probability α. If the audit detects bribery (which happens with probability β), both parties are subject to a sanction, which consists of the confiscation of the illegal gains plus the full fine. The illegal gains include bribe b in any case and value v only in the case when the bureaucrat has chosen to perform *Action a*.

Figure 2.1 summarizes the extensive form of the game and the expected payoffs.

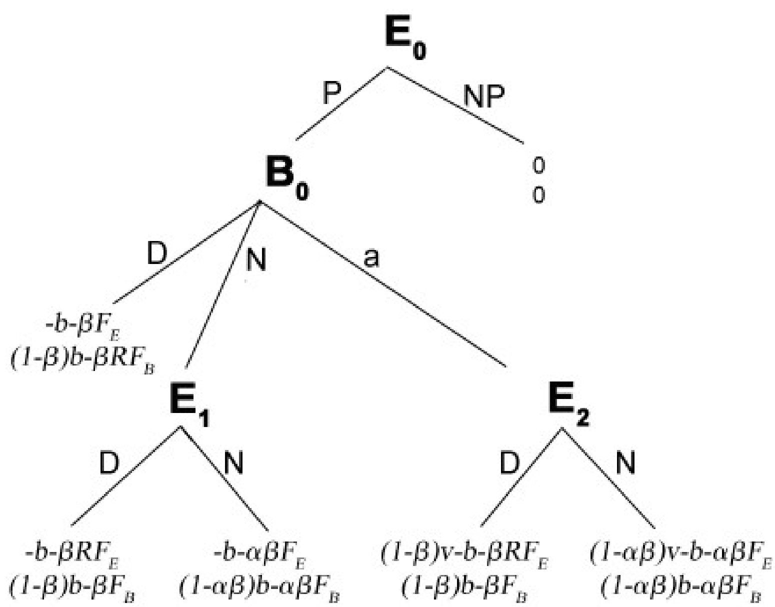

Figure 2.1: Extensive form of the corruption game in the generalized model. P stands for *Pay*, NP for *Not Pay*, D for *Denounce*, N for doing *Nothing*, a for performing *Action a*, b for bribe, v for the value of the project to the entrepreneur, α for the exogenous probability of an audit, β for the probability of conviction, F_E and F_B for full fines and RF_E and RF_B for reduced fines to the entrepreneur and the bureaucrat, respectively.

The contribution of the generalized model lies in the introduction of probability β. In Buccirossi and Spagnolo (2006) it is assumed that, before the illegal transaction takes place, the bureaucrat and the entrepreneur agree on the production of hard evidence. Without hard evidence being voluntarily produced by both of them the illegal transaction is not enforceable. In essence it is assumed that both involved are holding a hostage that commits each other to the desired outcome. It is furthermore assumed that, if an audit takes place, corruption is discovered and both culprits are convicted with a probability of one. Richmanová and Ortmann (2008) assume instead that some hard evidence is created unintentionally along the way and that this evidence may be discovered by an audit with probability $\beta \in (0, 1)$. The basic structure of both the original and the modified game is the same except that in the original version the probability β is set to 1. The generalization makes the model more suitable for experimental testing, as no additional stage is needed in which subjects would have to agree on producing a hostage. In addition, the generalized model arguably resembles real-world situations more closely.[10]

Buccirossi and Spagnolo (2006) show that in the absence of a leniency program, occasional illegal transactions are not implementable.[11] The result carries over into the generalized model. After the introduction of a modest leniency program,[12] occasional illegal transactions are enforceable if the following three conditions are satisfied simultaneously. First, the *no-reporting condition* for the bureaucrat: the reduced fine must be such that the bureaucrat prefers performing *Action a* to *Denouncing* once the bribe has been paid. Second, the *credible-threat condition* for the entrepreneur: the reduced fine and the full fine must be set such that the entrepreneur can credibly threaten to report if the bureaucrat does not deliver. Third, the *credible-promise condition*: the entrepreneur must be able to credibly promise not to report if the bureaucrat respects the illegal agreement.

[10]I realize that in such a game beliefs about the probability of detection might play an important role. However, I believe that the introduction of beliefs would make the game more complex than necessary for experimental testing. Instead, I view probability β as the empirical success rate, or effectiveness, of a detection technology that is known to subjects.

[11]Facing the full fine even after reporting, the entrepreneur cannot credibly threaten to report the bureaucrat in the case when he would not deliver. Therefore, the bureaucrat would keep the bribe and not perform *Action a*, knowing that it is not profitable for the entrepreneur to punish him. Consequently, the entrepreneur would not enter the illegal agreement in the first place.

[12]Similar to Spagnolo (2004), "modest" means that a leniency program does not reward for reporting, at best it cancels the fine.

These three conditions, given the value of the project together with the full and reduced fines, define a bribe range for which the occasional illegal transaction is implementable. Even though these conditions are modified in the generalized model, the qualitative result remains unaffected.

I used the generalized version of the game for experimental testing of the theoretical prediction under two different scenarios: when the occasional illegal transaction is implementable in equilibrium, and when it is not. Implementability is a function of the per-round endowment for the entrepreneur. The per-round endowment exogenously defines the value of the bribe if the entrepreneur decides to pay it.[13] For each treatment I use two possible values of the per-round endowment: a low endowment that theoretically leads to a no-corruption equilibrium, and a high endowment that theoretically leads to a corruption equilibrium.

I want to study whether a change in parameterization that does not affect the theoretical prediction will have an impact on the behavior of subjects in a lab. In a game like this, where an action bringing the highest possible payoff is also associated with the risk of an enormous loss, it is likely that subjects in the lab will not behave in full accordance with the theoretical prediction. I want to see whether by making the risky choice more tempting I can induce more transferring in the lab. I also want to see what the consequences are for later stages of the game, particularly for denouncing. For that purpose I ran two treatments: a benchmark and a high-incentive treatment.

Table 2.1 summarizes the parameterizations chosen for the Benchmark treatment (B) and for the (Benchmark-)High-incentives treatment (BH).

Treatment	α	b	v	RF$_E$	RF$_B$	F$_E$	F$_B$	E$_L$	E$_H$	show-up
B	0.1	0.2	100	0	0	300	300	20	40	300
BH	0.1	0.2	200	0	0	200	200	10	30	200

Table 2.1: Experimental parameterization. α and β denote the probability of an audit and of discovering evidence of bribery, respectively; v denotes the value of the project to the entrepreneur; RF_E and RF_B denote reduced fines and F_E and F_B full fines to the entrepreneur and to the bureaucrat, respectively; E_L and E_H denote low and high per-round endowment, respectively; show-up stands for the show-up fee.

In the B treatment, the probabilities α and β were chosen such that they approximately correspond to real-world exogenous probabilities of audit

[13]This way I reduce the cognitive demand on subjects: the only decision they have to make is whether they want to transfer their per-round endowment or not.

and to real-world conviction rates; at the same time they are intuitively comprehensible for subjects. The value of the project v was chosen together with full fines F_E and F_B such that the subject faces a considerable gain from the investment but also severe punishment in the case of detection. I set reduced fines RF_E and RF_B equal to zero to analyze the case of full leniency programs which, according to Apesteguia et al. (2004), have promising anti-cartel properties. Endowment determines the value of a bribe to (not) be paid. The "low endowment" of 20 leads (theoretically) to no corruption, whereas the "high endowment" of 40 leads to corruption equilibrium. Finally, the show-up fee was set such that I eliminate the possibility of earning a negative total from the experiment.

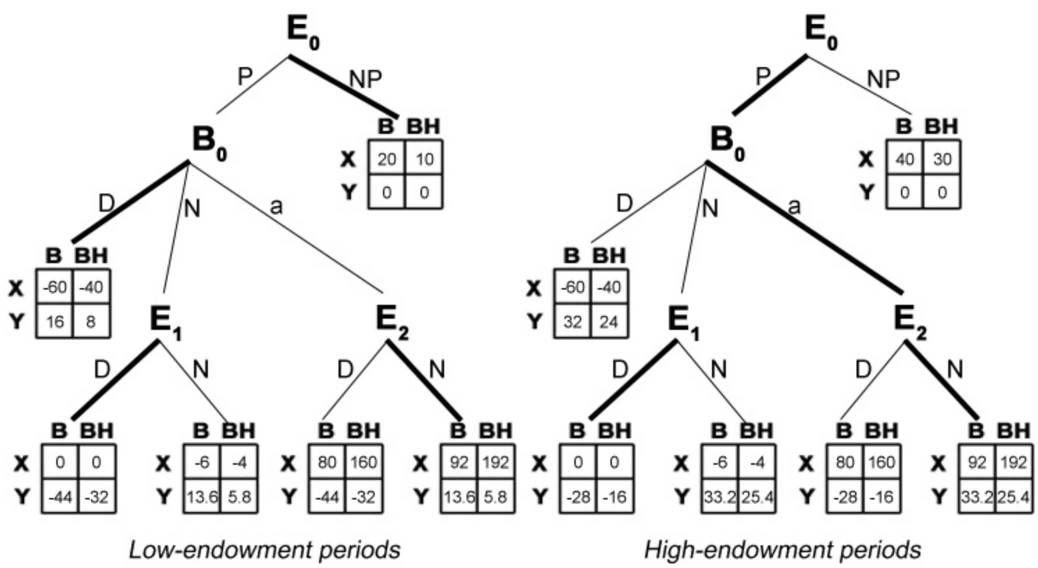

Figure 2.2: Expected payoffs from the corruption game in the B (benchmark) and the BH (high-incentive) treatments, respectively. Rows in the tables correspond to Participant X and Participant Y; columns correspond to the B and BH treatments. The theoretical prediction is the same for both treatments, it only varies with the endowment.

In the BH treatment, in order to make the risky but high-payoff choice more tempting, I increased the value of the project to the entrepreneur and, at the same time, I reduced the fines both agents face in case of detection. I keep the probabilities of detection and of conviction (thus the exogenous risk) unchanged. In order to keep the theoretical prediction for low- and high-endowment periods qualitatively the same as in the benchmark treatment, the per-round endowments were also adjusted. Finally, the show-up fee is set such that subjects cannot end up with a negative final payoff, but there is a chance that they will earn zero.

The extended game forms together with the expected payoffs resulting from the parameterizations for both the B and the BH treatments are illustrated in Figure 2.2 for low- and for high-endowment periods. The branches identifying the equilibrium choices of risk-neutral agents are in bold.

2.3 Implementation

The experiment was conducted in November and December 2006 at CERGE-EI in Prague, using a mobile experimental laboratory.[14]

Participants were recruited from the Faculty of Social Sciences of Charles University in Prague and from various faculties of the Czech Technical University in Prague. Students were approached via posters distributed on campus and via e-mail.[15]

I conducted four sessions of each treatment. Twelve participants, six in the role of Participant X – the entrepreneur – and six in the role of Participant Y – the bureaucrat – interacted in each session. In each session, all subjects participated in six rounds during which they kept the role that was assigned to them at the beginning of the first round.[16] Participants were randomly and anonymously re-matched after each round so that no subject was matched twice with the same co-player. This was common knowledge. The incentive compatibility of this matching scheme is discussed in Kamecke (1997).

[14]http://home.cerge-ei.cz/ortmann/BA-PEL.htm

[15]By email, I also directly invited students who participated earlier in unrelated experiments conducted at CERGE-EI.

[16]After each Participant X interacted exactly once with each Participant Y, the roles were switched for another six rounds. Subjects were not informed about the switch of roles in advance in order to avoid a possible impact on their behavior in the first six rounds. Before the beginning of the seventh round the announcement about the switch of roles appeared on their screens. The decisions in the last six rounds are likely affected by subjects' experience from the first six rounds and therefore I do not report them in the main text. A comparison of the before-switch and after-switch data is provided in the appendix. For the B treatment, I observe more transferring in the after-switch data, and also more denouncing in both the second and the third stage. In the BH treatment, I observe less denouncing in the second stage. The rest of the results seem unaffected.

Treatment	**Subject Source**[17]	**M/F ratio**[18]	**mean (age)**	**mean (RA score)**	**mean (final pay)**[19]	**Irreg**[20]
B	FSS	8/4	20.9	29.7	320	1
B	FSS	10/2	21.75	28.8	330	0
B	CTU	11/1	22.9	34.7	330	0
B	FSS	9/3	22.3	26.4	323.3	0
BH	CTU	9/3	22.6	33	185.8	1
BH	CTU	10/2	22.8	28.9	309.2	0
BH	CTU	10/2	22.5	29.3	241.7	1
BH	FSS	10/2	21.9	24.8	259.2	1

Table 2.2: Summary of the demographic characteristics of subjects for all eight sessions.

Table 2.2 summarizes the demographic characteristics of subjects participating in the experiment. The majority of my subjects were male, reflecting the composition of the subject pools that I drew on. Mean age ranges between 20.9 and 22.9; over all sessions the minimum is 18 and maximum 27. I also measured subjects' risk aversion using a questionnaire based on Holt and Laury (2002). Mean RA score ranged between 24.8 and 34.7; over all sessions the minimum is 15 and maximum 51.[21] Average final payoffs for the B treatment ranged from 320 to 330, with the minimum being 300 and the maximum 400; for the BH treatment it ranged between 185.8 and 309.2,

[17]For each session, subjects were recruited from one source. FSS stands for the Faculty of Social Sciences in Prague, CTU for the Czech Technical University in Prague. I control for imbalance of the subject pool by including the econ and gender dummies in the econometric analysis.

[18]Male/Female ratio in the session.

[19]This is the average final payoff after the computerized part of the experiment.

[20]Irreg stands for a dummy variable for session irregularities. In the first B-treatment session an experimenter effect is possible; in the first BH-treatment session a typo in the Z-tree program caused incorrect payoffs for the two final nodes displayed on the screens, which was pointed out by one of the subjects only after several rounds; in the third BH-treatment session two subjects continued communicating despite several admonitions; and in the fourth BH-treatment session two subjects were reading a newspaper in between making their choices. I do not believe that these irregularities would matter but wanted to control nevertheless. After running the preliminary regressions I concluded that they indeed did not matter.

[21]The higher the score the more risk averse the subject is. The maximum possible RA score is 60 which, using the standard CRRA utility function $x^{(1-r)}$, approximately corresponds to a relative risk aversion coefficient of .17. The minimum possible RA score is 0, which approximately corresponds to a relative risk aversion coefficient of $-.13$. An RA score of 23 corresponds to risk-neutrality.

with the minimum being 0[22] and the maximum 400.[23]

Each session began with general instructions. Afterwards, subjects were asked to fill in Risk-aversion and Demographic questionnaires, for which they earned their show-up fee. Then the instructions to the computerized part of the experiment were distributed. Understanding of the instructions was tested by a brief questionnaire. The computerized part of the experiment started only after every participant answered all testing questions correctly.[24] The session concluded with a final questionnaire asking for the subject's feedback on the experiment.[25]

All instructions were read aloud by the experimenter. As a part of the instructions subjects received a pictorial representation of the game with a minimum use of game-theoretic terminology. Probabilistic outcomes were presented in both probabilistic terms and frequency representation (see e.g. Gigerenzer and Hoffrage 1995, or Hertwig and Ortmann 2004). All instructions were presented in completely neutral language, with no reference to bribery. The roles of the bureaucrat and the entrepreneur were renamed *Participant X* and *Participant Y*, actions were labelled with neutral letters, *Pay/Not Pay* a bribe was replaced with *transfer/not transfer*; and no detection/detection were labelled *outcome A* and *outcome B*, respectively (for an analysis of the impact of loaded instructions see e.g. Abbink 2006 or Krajčová and Ortmann 2008).[26]

The experiment was computerized using Z-tree software (Fischbacher 2007). At the beginning of each round, each participant was notified of her/his role. Participants X also learned their current per-round endowment. Then each pair interacted sequentially.[27] Between the second and the third

[22]At that point 400 CZK corresponded to about 16 USD, in purchasing power up to twice as much. Subjects were informed during recruitment that their final payoff from the experiment might be zero, but could not be negative. The non-negativity of the final payoff was ensured by the show-up fee.

[23]The difference in average payoffs in the B and in the BH treatment results from different parameterization as well as from different behavior of subjects as will be illustrated later.

[24]This was common knowledge.

[25]For filling out this last questionnaire, subjects were paid an additional 50-200 CZK (corresponds to about 2-9 USD) - the amount varied between sessions. This mechanism was used to adjust average earnings for the session to the level promised during the recruitment.

[26]Originals (in Czech) of all materials that subjects received during the experiment are available at http://home.cerge-ei.cz/richmanova/WorkInProgress.html.

[27]Choices were made by clicking the respective buttons on the screen. Subjects were

stage, Participants X were asked about their choices in each node of the third stage if they were to reach it. After they made their conditional choices, they learned the actual decision of their co-player and they were asked to confirm, or to change, their previous choice. This mechanism allowed me to collect some additional data in rounds when the third stage was not reached.

At the end of each round subjects were given feedback about their action(s), the action(s) of the player they were paired with, the realization of the random outcome (A or B) and their resulting payoff. At the end, one round was randomly chosen to determine the final payoff from the computerized part of the experiment. This mechanism was chosen in order to ensure that the decision in every round is made as if in a one-shot game. This payment procedure was common knowledge ex ante.

Participants were paid anonymously in cash right after each session. I used the Czech crown as the currency unit throughout the whole experiment.

2.4 Results

2.4.1 Summary Data

In Figure 2.3, the results from low- and high-endowment periods are presented. Each figure integrates the results from both treatments – the B treatment data in the upper rows, the BH treatment below. The equilibrium choices for each case are in bold face.

For the aggregate first-stage data, a clear treatment effect can be observed – the frequencies of choosing *Pay* are higher in the BH treatment than in the B treatment. In both treatments, the frequencies of choosing *Pay* are higher in the low-endowment periods than in the high-endowment periods, which contradicts the theoretical prediction. Intuitively, subjects seem to be willing to transfer their endowment in order to get a chance of receiving a high payoff, but they are more willing to put at stake a low endowment than a high. Instead of risking the high endowment they seem to prefer choosing the sure outcome.

As for the second-stage data, only relative percentages can be compared across treatments, as different numbers of subjects actually entered this stage of the game. In both low- and high-endowment periods, the results for the two treatments are very similar: it is about an equal split between playing

notified that once they made their choice it would not be possible to take it back.

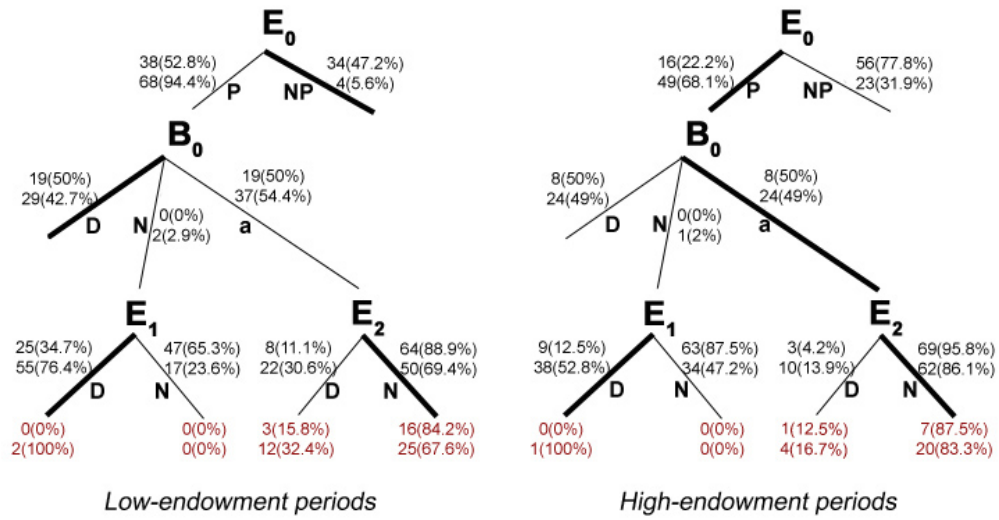

Figure 2.3: Experimental results. For each branch of the extensive form of the game, the upper row always displays the frequency of the action in the B treatment; and the lower row displays the frequency of the action in the BH treatment (both with the corresponding percentage in parentheses). For the nodes E_1 and E_2, above the branches, I present the conditional choices subjects were asked to report before they made their actual choice. Frequencies of real choices, which depend on the preceding decision of Participant Y, are presented at the bottom part of each figure.

Denounce or *Action a*. Only in low-endowment periods of the BH treatment *Action a* slightly dominates. These results are not in line with the theoretical prediction. The difference in expected payoffs resulting from *Denounce* and *Action a* is, however, very small and that may be the reason why I do not observe a stronger inclination to either choice. Also note that in both treatments *Denounce* is the only action through which the bureaucrat can avoid a negative expected round payoff with certainty.[28] In line with theoretical prediction and also intuition, *Nothing* was almost never chosen.

As for the third-stage data, conditional choices provide mixed evidence. In the E_1 node, both conditional and sequential choices in the BH treatment are closer to the theoretical prediction than in the B treatment. In the E_2 node, it is just the opposite: in the BH treatment both conditional and

[28]See Figure 2.2 and Table 2.1 for more details. Even though the subject could possibly earn a negative round payoff, each subject also received a show-up fee that ensured a non-negative total payoff.

sequential[29] choices move further away from the theoretical prediction.

Note that for the second and third stage data I have too few independent observations (especially so for the B treatment and for the high-endowment periods)[30] to perform a reliable formal analysis. Therefore, I only perform a statistical and regression analysis of the first-stage data.

2.4.2 Analysis of the First-Stage Data

In the following two subsections I report the results from the formal analysis of the first-stage data. I conducted standard non-parametric tests identifying differences in the distribution of choices under the two treatments. I also computed the effect size indices to measure the magnitude of the treatment effect. Finally, I report the results from the estimation of a linear probability model in which I control for some demographic characteristics of subjects.

Due to the panel nature of the data, I considered four different approaches to formal regression analysis: 1) clustered data analysis – data from periods $1, 3$, and 5 (low-endowment) and from periods $2, 4$, and 6 (high-endowment) are clustered by subject to correct standard errors for likely within-subject correlation; 2) first-period data analysis – only first-period data (for the low-endowment case) and only second-period data (for the high-endowment case) are analyzed; 3) averaged data analysis – averaged data for periods $1, 3$, and 5 and for periods $2, 4$, and 6 are analyzed; and 4) dominant-choice data analysis – for each endowment value (low or high) each subject makes choices in three periods, and the dominant choice is the one that is played more often.

Clustered data have the advantage of using all the available information, while the other three approaches use only a part of the available information. Therefore, in the main text I discuss the results for clustered data. The analysis of averaged, first-period, and dominant-choice data can be found in the appendix, as a robustness check of the main results. By and large, there are no major findings in these robustness tests.

In addition to the robustness checks based on different "data handling" I also run a few additional exploratory sessions of treatments in which the

[29]When I asked subjects to make their real sequential choices, only one subject in the B treatment changed her/his decision in the E_2 node from *Denounce* to *Nothing* (after observing what Player 2 had chosen) in the low-endowment period. No one changed her/his decision in the high-endowment period or in the BH treatment.

[30]Recall that Figure 2.3 presents the aggregated data from all the relevant periods, therefore it contains repeated observations for individual subjects

experimental conditions are only slightly modified compared to the benchmark and the high-incentive treatments. The results from the analysis on the extended data set is provided in the appendix, as an additional robustness check of the main results. By and large, there are no major findings in these robustness checks. Pooling slightly different treatments leads to noisier results, which is not very surprising.

Statistical Analysis

In Table 2.3 I report the results of three standard non-parametric tests in order to identify the differences in the distributions of choices under the two treatments. Specifically, I test the null hypothesis of no differences between the two treatments using the averages of the binary transfer variable[31] over periods $1, 3$, and 5 and $2, 4$, and 6. According to Wilcoxon rank-sum, Kolmogorov-Smirnov and Fisher's exact tests I reject the hypothesis of no differences in the distribution of choices under two treatments at the 5% significance level.

periods	Ranksum[32]	Ksmirnov[33]	Fisher[34]
1,3,5	-3.632 (.000)	.500 (.002)	(.001)
2,4,6	-3.853 (.000)	.625 (.000)	(.000)

Table 2.3: Non-parametric tests.

To assess the magnitude of the effect for practical purposes, I in addition compute two standardized measures of effect size: Cohen's d and odds ratio, again, using the averages of the binary transfer variable over periods $1, 3$, and 5 and $2, 4$, and 6. The results for the full sample and for the male and female subsamples are reported in Table 2.4.

Cohen (1998) defines effect sizes of $d = 0.2$ as *small*, $d = 0.5$ as *medium*, and $d = 0.8$ as *large*. For the full sample, as well as for the male and female subsamples, the results suggest a large effect – the transfer rates in the BH

[31]Transfer has a value of one if Participant X chooses *Pay* and a value of zero if s/he chooses *Not Pay* in the respective period.

[32]Ranksum stands for the two-sample Wilcoxon rank-sum (or Mann-Whitney) test. I report the normalized z statistic and corresponding p-value below.

[33]Ksmirnov stands for the Kolmogorov-Smirnov test. I report the statistic and below the corresponding p-value from testing the hypothesis that average transfer is lower in the B treatment.

[34]Fisher stands for Fisher's exact test. I report the resulting p-value.

treatment are considerably higher than in the B treatment for both male and female subsamples.

		B			BH			effect size	
Periods	Sample	N	mean	std.dev.	N	mean	std.dev.	odds ratio	Cohen's d
1,3,5	full	24	.528	.4495	24	.944	.2123	1.788	1.1827
	male	18	.519	.4461	19	.930	.2378	1.792	1.150
	female	6	.556	.5018	5	1	0	1.799	1.251
2,4,6	full	24	.222	.3764	24	.681	.3330	3.068	1.2924
	male	18	.296	.4105	19	.719	.3194	2.429	1.150
	female	6	0	0	5	.533	.3801	NA[35]	1.983

Table 2.4: Effect-size indices.

Altogether, both statistical tests and effect-size measures suggest that there are significant differences between the first-stage choices in the BH and B treatments. In the next step I perform a further analysis in which I control for gender and for other subject characteristics.

Econometric Analysis

During the experiment I distributed several questionnaires in order to collect basic demographic data. Specifically, I have information about subjects' age, gender, university and field of study.[36] I also measured each subject's risk aversion.

The dependent variable was defined as a 0/1 dummy variable *translog* identifying *Pay* being chosen (value of 1) or not (value of 0) in a particular period. I estimate a clustered linear probability model. I prefer a linear probability model to other non-linear alternatives, as it does not rely on very specific distributional assumptions, the violation of which leads to inconsis-

[35] A division-by-zero problem occurs, due to no variation in this subsample.

[36] In addition, I collected data on: size of subject's household, number of cars in the household, and whether the subject himself has his own car and what is its approximate value, all of which serve as proxies for income. I also asked the subjects whether they considered themselves as technical types compared to their peers. I recorded the occurrence of any inconsistencies in the after-instructions questionnaire, which served as a simple test of understanding of the basic structure of the game, and in the risk-aversion questionnaire. At the end of the session I asked my subjects whether they understood the experiment. Finally, I recorded some general information about each session – the time of day it started and any session irregularities if they occurred. After running some preliminary regressions I, however, conclude that none of these variables is significant for explaining subjects' decisions. The demographic and the risk-aversion questionnaires are based on Rydval (2007).

tent estimates if non-linear models are employed. Another advantage of the linear probability model lies in the straightforward interpretation of the estimated coefficients. I run clustered robust estimation to correct standard errors for likely within-subject correlation and for heteroskedasticity.

In the appendix, I provide a discussion of the robustness checks I conducted in addition to the clustered regression analysis. As the theoretical prediction differs for low- and high-endowment periods,[37] these two groups were analyzed separately.

I start with a basic minimal model:[38]

$$P(translog = 1|x) = \beta_0 + \beta_1 \cdot age + \beta_2 \cdot male + \beta_3 \cdot econ + \beta_4 \cdot BHtreat,$$

where *age* corresponds to the subject's reported age, *male* is a dummy variable defined based on the subject's reported gender, and *econ* is a dummy variable identifying a subject having (value of 1) or not having (value of 0) an economic background, which is defined based on the subject's reported field of study. As I am mainly interested in the treatment effect, I also include a BH-treatment dummy *BHtreat* in the model.

The results from the estimation are summarized in Table 2.5.

The model is strongly significant for both low- and high-endowment periods. Importantly, also the treatment dummy is significant at any conventional significance level.

The mean predicted probability of transfer in the low-endowment periods is .7361 for the pooled sample. For the B treatment it is .5278, for the BH treatment .9444. In the high-endowment periods, the mean predicted probability of transfer is considerably lower. For the pooled sample it is .4514, for the B treatment .2222, and for the BH treatment .6806. Thus, as I expected, the transfer rate is much higher in the high-incentive treatment. For both

[37] Recall that in periods $1, 3$, and 5 the endowment was low and in periods $2, 4$, and 6 the endowment was high.

[38] The second approach I used was $P(translog = 1|x) = \beta_0 + \beta_1 \cdot ra\_score$, where ra_score is a risk aversion score computed based on data from the risk-aversion questionnaire. Preliminary analysis suggests that age, male and econ predict ra_score well (all three are jointly significant at the 5% level, age and male with a negative sign on the coefficient, age with a positive; my proxy for income appeared insignificant, which is reasonable given my population sample). It was natural to consider these two sets of independent variables - one including ra_score only, and the other including male, age and income - as candidates for minimal models for my analysis. However, in $P(translog = 1|x) = \beta_0 + \beta_1 \cdot ra\_score$ ra_score never appeared significant and only rarely I observed the joint significance of the estimated models. Therefore, I omit a discussion of these results.

	periods 1,3,5	periods 2,4,6
age	**-.0761** (.001)	**-.0452** (.053)
male	**.0335** (.785)	**.3008** (.007)
econ	**-.1990** (.056)	**-.0773** (.498)
BHtreat	**.3019** (.007)	**.3972** (.001)
mean $\widehat{p}$(y=1)	**.7361**	**.4514**
# of obs.	144	144
joint p-value	(.000)	(.000)

Table 2.5: The results from the estimation of the linear probability model. The first row of each cell reports the estimated coefficients. The second row reports the corresponding p-value. *Mean $\widehat{p}$(y=1)* denotes the mean predicted probability of a transfer being made.

treatments the transfer rate is higher in low- than in high-endowment periods. This result contradicts the theoretical prediction[39] (we find the same result in the context treatment, see Krajčová and Ortmann 2008).

Age is significant at the 5% level for low-endowment periods and at the 10% level for high-endowment periods, in both cases with a negative sign on the coefficient. An additional year of age reduces the probability of transfer by .08 with low and by .05 with high endowment.

The *male* dummy is not significant for low-endowment periods, but I get strong significance for high-endowment periods.[40] In both cases, the sign on the coefficient is positive, meaning that men are more likely to transfer – by .03 when the endowment is low and by .30 when it is high – than women.

Econ is significant at the 10% level for low endowment and not significant for high endowment periods. The sign on the coefficient is, in both cases, negative. Thus, subjects with an economic background are less likely to transfer.

The *BHtreat* dummy is significant at the 1% level for both low- and high-endowment periods. The sign on the coefficient is positive meaning that, as I expected, subjects in the high-incentive treatment are more likely to transfer

[39] Recall that in the equilibrium Participant X always transfers with a high endowment and never transfers with a low.

[40] I find no evidence of gender-specific effects such as in Krajčová and Ortmann (2008). In the first stage, both male and female participants transfer more in the BH treatment than in the B treatment. I find some differences in the behavior of men and women – in the second stage with high endowment; and for sequential choices in the E_2 node with both low and high endowment. In all three cases, however, the size of the female subsample is too small to make any plausible inferences.

– by .30 with low and by .40 with high endowment – than subjects in the benchmark treatment.

In general, the main results that can be observed from the descriptive data are also statistically significant.

2.5 Discussion

I expected that subjects in my experiment might not behave in complete accordance with the theoretical predictions made under the assumption of rationality and risk-neutrality. Apart from risk attitudes, phenomena such as altruism, reciprocity (positive or negative) and/or trust might play important roles. In fact, in my data I observe considerable deviations from equilibrium at some stages of the game. The change in parameterization shifts some of the results closer and some further away from the predictions. In this section, I discuss the results, and provide some explanations for these deviations and for the observed treatment effect. I also derive implications for experimental design and the implementation of the experimental testing of leniency programs.

In the first stage, for both treatments, I observe higher transfer rates in low- than in high-endowment periods, which contradicts the theoretical prediction.[41] In the BH treatment the fraction of out-of-equilibrium choices is even higher than in the B treatment. A similar result is found for the context treatment in Krajčová and Ortmann (2008).

I note that the theoretical prediction is computed under the assumption of risk neutrality, which, as also suggested by the data from the risk-aversion questionnaire, is not likely to hold in my sample. My subjects appear to be modestly risk-averse, in accordance with the typical finding in the experimental literature (e.g. Holt and Laury 2002, Harrison, Johnson, McInnes, and Rutström 2005). When I computed the theoretical prediction for a (modestly) risk averse subject, I found that under some (reasonable) assumptions, my chosen parameterization can lead to a no-corruption equilibrium also for the high-endowment periods.[42] That is, for risk-averse subjects, it might in

[41] Recall that in the equilibrium Participant X always transfers with a high endowment and never transfers with a low. Or, in other words, given the leniency program currently in force, theoretically, with a low endowment (thus, a low bribe) an occasional illegal transaction is not implementable.

[42] I assume a standard CRRA utility function $u(x) = x^{(1-r)}$. The average risk-aversion

fact be optimal not to transfer a high endowment.

In addition, my subjects might exhibit the "preference for inclusion" reported by Cooper and Van Huyck (2003). The authors find that subjects presented with an extensive form game are significantly more likely to make choices that allow their co-player to make a choice – and thereby to affect final payoffs – rather than choosing a terminal node. In an extensive form game this "(non)inclusion" is more salient. In my game, "inclusion" introduces a risk of significant loss. Together with risk- or loss-avoidance, it might have resulted in subjects with a "preference for inclusion" being willing to transfer and continue playing the game, but only being ready to risk the low endowment and preferring to keep the high endowment for sure.

Furthermore, note that the difference in expected payoffs to Participant Y from choosing *Denounce* or *Action a* is relatively small[43] in both treatments (assuming that Participant X will react rationally), whereas the difference in payoffs to Participant X is substantial. Therefore, an altruistic Participant Y might prefer choosing *Action a* even in low-endowment periods, when this action is not maximizing the expected payoff. Or, alternatively, choosing *Action a* might be an act of positive reciprocity. In low-endowment periods, a rational Participant X might expect a rational Participant Y to choose *Denounce* and therefore he would not transfer. A Participant X who is trusting might expect Participant Y to choose *Action a* in the second stage and therefore he might want to transfer.

In the BH treatment, the difference in expected payoffs to Participant Y is about the same, but the possible gain to Participant X (after *Action a* has

coefficient in my sample is about 0.03; the maximal is about 0.1. As the bribery game involves nodes with negative payoffs, some assumptions need to be made about the utility function in the negative domain. Prospect theory suggests that in the negative domain, the steepness of the utility function might be about twice as much as in the positive domain. For illustration, I computed the theoretical prediction for a risk-neutral subject in the B treatment assuming two different levels of (dis)utility from paying a 300 CZK penalty after detection: $u(-300) = -u(450)$ and $u(-300) = -u(600)$. For low endowment, the theoretical prediction is the same as for a risk-neutral subject. For high endowment it changes. For an extremely risk-averse participant ($r = 0.1$), the disutility of 450 still predicts a corruption equilibrium, however, the disutility of 600 predicts a no-corruption equilibrium. For an average risk-aversion coefficient ($r = 0.03$), the disutility of 450 is sufficient to change the theoretical prediction. I obtained analogical results for the BH treatment (because of a different parameterization, the relevant disutilities for the BH treatment are $u(-200) = -u(300)$; and $u(-200) = -u(400)$).

[43]Note that this results from the nature of the game (see Figure 2.1).

been chosen) is considerably larger than in the B treatment. That is why, if the above arguments hold, the new paramererization might shift the choices even further away from the equilibrium. This is indeed what I observe in the data.

In the second stage, for both treatments, I observe about an equal split between choosing *Denounce* and *Action a*, for both low- and high-endowment periods. *Nothing* is almost never chosen.

Payoffs for Participant Y resulting from *Nothing* and *Action a* are the same, but taking into account the likely decisions of Participant X in the following stage, he is more likely to collect a higher payoff after he chooses *Action a*. This seems to be correctly recognized by the vast majority of my subjects. As regards the relative indecisiveness of subjects between choosing *Denounce* or *Action a*, I repeat the arguments mentioned above – the difference in expected payoffs is relatively small, which together with different preferences for altruism and reciprocity might have produced these results.

In the E_1 node, a new parameterization shifts the results closer to the prediction. Intuitively, if subjects exhibit negative reciprocity, this becomes the more salient the more there is at stake.

In the E_2 node, the majority of subjects plays equilibrium in both treatments. In the BH treatment the fraction of subjects who play equilibrium is slightly smaller. It is still the majority, though.

Altogether, my data to some extent confirm the main result of Buccirossi and Spagnolo (2006) – an occasional illegal transaction is implementable with a leniency policy in place. This becomes especially visible in the high-incentive treatment with high endowment. I observe a sensitive reaction to a parametric change that does not affect the theoretical prediction. My finding suggests that calibration, i.e. parameterization that reflects "real-life" situations reasonably well, might even be more important than in other scenarios. My data also suggest that other factors might be important as well. Trust and preferences towards others might play an important role. Further experimental testing of leniency policies might have to take these findings into account.

References

Abbink, K., Hennig-Schmidt, H., (2006). Neutral versus Loaded Instructions in a Bribery Experiment, *Experimental Economics 9(2)*, 103-121.

Abbink, K., Irlenbusch, B., Renner, E., (2002). An Experimental Bribery Game, *Journal of Law, Economics, and Organization 18(2)*, 428-454.

Apesteguia, J., Dufwenberg, M., Selten, R., (2007). Blowing the Whistle. *Economic Theory 31*, 143–166.

Bigoni, M., Fridolfsson, S.-O., Le Coq, C., Spagnolo, G., (2008a). Fines, Leniency, Rewards and Organized Crime: Evidence from Antitrust Experiments. *Working Paper Series in Economics and Finance, No. 698, Stockholm School of Economics.*

Bigoni, M., Fridolfsson, S.-O., Le Coq, C., Spagnolo, G., (2008b). Risk Aversion, Prospect Theory, and Strategic Risk in Law Enforcement: Evidence From an Antitrust Experiment. *Working Paper Series in Economics and Finance, No. 696, Stockholm School of Economics.*

Buccirossi, P., Spagnolo, G., (2006). Leniency Policies and Illegal Transactions, *Journal of Public Economics 90*, 1281–1297.

Cohen, J., (1988). Statistical Power for the Behavioral Sciences, 2^{nd} edition. *Lawrence Erlbaum Associates Inc, Hillsdale.*

Cooper, D., J., Van Huyck, J.,B., (2003). Evidence on the Equivalence of the Strategic and Extensive Form Representation of Games, *Journal of Economic Theory 110*, 290-308.

Dušek, L., Ortmann, A., Lízal, L., (2005). Understanding Corruption and Corruptibility through Experiments. *Prague Economic Papers 14*, 147-162.

Fischbacher, U., (2007). Z-tree: Zurich Toolbox for Ready-made Economic Experiments - Experimenter's Manual. *Experimental Economics 10(2)*, 171-178(8).

Gigerenzer, G., Hoffrage, U., (1995). How to Improve Bayesian Reasoning without Instruction: Frequency Formats. *Psychological Review 102*, 684–704.

Goeree, J., K., Holt, C., A., (2001). Ten Little Treasures of Game Theory, and Ten Intuitive Contradictions. *American Economic Review 91*, 1402-1422.

Gupta, S., Davoodi, H., Alonso-Terme, R., (2002). Does Corruption Affect Income Inequality and Poverty? *Economics of Governance 3(1)*, 23-45.

Harrison, G., W., Johnson, E., McInnes, M., Rutström, E., (2005). Risk Aversion and Incentive Effects: Comment. *American Economic Review 95 (3)*, 897-901.

Hertwig, R., Ortmann, A., (2004). The Cognitive Illusions Controversy: A Methodological Debate in Disguise that Matters to Economists. *in Zwick, R., Rapoport, A. (eds.), Experimental Business Research, Kluwer Academic Publishers, Boston, MA*, 361- 378.

Holt, C., A., Laury, S., K., (2002). Risk Aversion and Incentive Effects, *American Economic Review 92 (5)*, 1644-1655.

Hwang, J., (2002). A Note on the Relationship Between Corruption and Government Revenue. *Journal of Economic Development 27(2)*, 161-178.

Kamecke, U., (1997). Rotations: Matching Schemes that Efficiently Preserve the Best Reply Structure of a One Shot Game, *International Journal of Game Theory 26 (3)*, 409-417.

Krajčová, J., Ortmann, A., (2008). Testing Leniency Programs Experimentally: The Impact of Natural Framing, *CERGE-EI Working Paper No. 372*.

Mauro, P., (1995). Corruption and Growth. *Quarterly Journal of Economics, 110*, 681-712.

Ortmann, A., Lízal, L., (2003). Designing and Testing Incentive-compatible and Effective Anti-corruption Measures, grant proposal successfully submitted to the Grant Agency of the Czech Republic. Grant No. 402/04/0167.

Richmanová, J., (2006). In Search of Microeconomic Models of Anti-Corruption Measures – A Review, *CERGE-EI Discussion Paper No. 2006-157.*

Richmanová, J., Ortmann, A., (2008). A Generalization of the Buccirossi and Spagnolo (2006) Model. *CERGE-EI Discussion Paper No. 2008-194.*

Roth, A., (2002). The Economist as Engineer: Game Theory, Experimentation, and Computation as Tools for Design Economics, *Econometrica 70(4)*, 1341-1378.

Rydval, O., (2007). The Impact of Financial Incentives on Task Performance: The Role of Cognitive Abilities and Intrinsic Motivation, *PhD dissertation, CERGE-EI.*

Spagnolo, G., (2004). Divide et Impera: Optimal Leniency Programs, *C.E.P.R. Discussion Paper No. 4840.*

Tanzi, V., (1998). Corruption Around the World: Causes, Consequences, Scope and Cures. *IMF Working Paper 98/63.*

Williamson, O., (1983). Credible Commitments: Using Hostages to Support Exchange. *American Economic Review 73*, 519-540.

APPENDIX

Comparing the Data from Periods Before and After the Switch of Roles.

In Figure 2.4, I present the data from the before- and after-the-switch-of-roles periods (before-switch data in the upper rows and after-switch data below) from the low- and high-endowment periods of the B treatment.

In both cases, I observe a somewhat higher transfer rate in the second six periods. The transfer rate is higher in the low-endowment periods than in the high, before and after the switch of roles. In the B_0 node, more subjects chose the safe option (with no possibility of loss) after the switch of roles. In the E_2 node, results from before- and after-switch data are very similar, which is not the case of the E_1 node, where the relative percentages shifted closer to the equilibrium prediction.

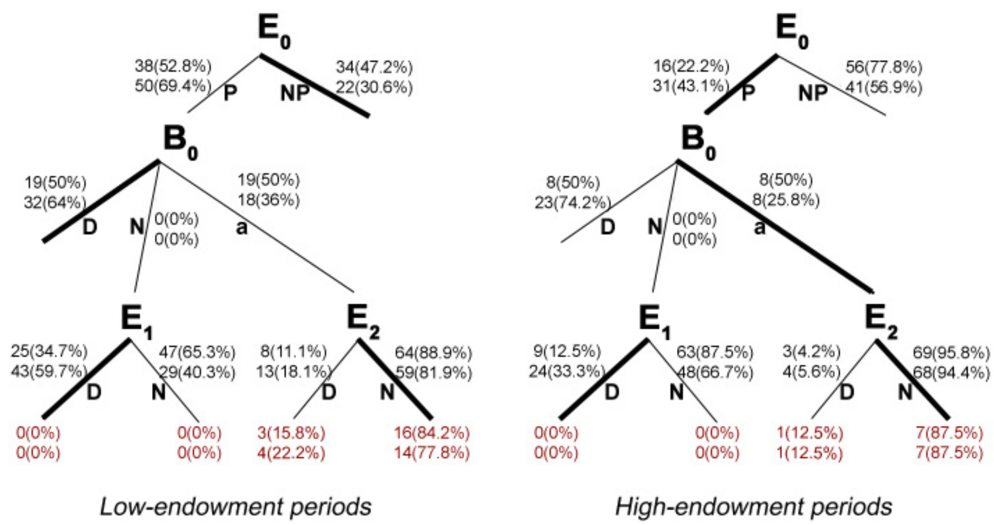

Figure 2.4: Before- vs. after-the-switch-of-roles data in the B treatment. Before-switch data are in the upper rows and after-switch data are below.

In Figure 2.5, I present the data from the before- and after-switch-of-roles periods (before-switch data in the upper row and after-switch data below) from the low- and high-endowment periods of the BH treatment.

In both cases, I observe no differences in the transfer rates after the switch

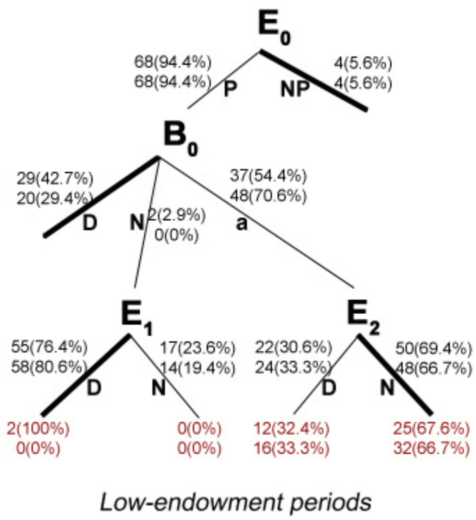

Figure 2.5: Before- vs. after-the-switch-of-roles data in the BH treatment. Before-switch data are in the upper rows and after-switch data are below.

of roles. Similarly as in the first part, the transfer rate is higher in the low-endowment periods than in the high after roles are switched. In the B_0 node, less subjects chose the safe option (with no possibility of loss) after the switch of roles in both low- and high-endowment periods. In the E_1 and E_2 nodes, results from before- and after-switch data are very similar. This is somewhat different evidence than that from the B treatment.

Robustness Checks

I performed two types of robustness check of my estimation results. The first regards the way I treated individual observations over rounds when running regressions – this is discussed in the subsection *Handling of the Data*. The second regards the experimental design – I run several sessions of alternative treatments in which I introduce only minor changes that do not appear to significantly affect behavior of subjects – this is discussed in the subsection *Pooling the Sessions*.

A. Handling of the Data

Throughout the analysis I have defined three alternative dependent variables, each of which captures slightly different information about the first-stage data. *Translog* is a 0/1 dummy variable identifying a transfer being made (value of 1) or not (value of 0) in a particular period. *Atranslog* is the average value of *translog* for one individual over periods $1, 3,$ and 5 (low-endowment periods) or $2, 4,$ and 6 (high-endowment periods). *Ltranslog* defines the dominant choice of a subject in periods $1, 3,$ and 5 or $2, 4,$ and 6. For a subject who has chosen *Pay* two or three times out of a total of three periods of interest, the dominant choice is 1; for a subject who has chosen *Not Pay* two or three times out of a total of three periods of interest, the dominant choice is 0.

Then, using one of the three types of dependent variable, I conducted four different types of regression analysis.

Clustered Regressions – as discussed in the main text, I run a clustered (robust)[44] linear probability model estimation with the binary variable *translog* as a dependent variable.

Regressions on Averaged Data – in this case, I run an ordinary least squares estimation of *atranslog*. I analyze only averaged data, where higher values of *atranslog* correspond to more transfers being made and thus to a stronger preference for this choice.[45]

Regressions on the 1^{st} *or* 2^{nd} *Period Data* – I estimate LPM only on the 1^{st} and 2^{nd} period *translog* (for low- and high-endowment periods, respectively). In this approach I omit part of the information, however I only use the part of the data that is not affected by the experience from previous rounds.[46]

Regressions on Dominant Choice – I estimate LPM using *ltranslog* as a dependent variable. Thus in this case, I am only looking at the dominant choice of each subject.

First I look at effect size measures, whether they give robust results for all four approaches to the data. The results are summarized in Table 2.6.

In all cases, the effects are large (recall that Cohen 1998 defines effect sizes of $d = 0.8$ as *large*). The directions of the effects are the same in all cases – transferring is higher in the high-incentive (BH) treatment.

Tables 2.7 and 2.8 summarize the main results from the estimation for low- and high-endowment periods, respectively.

Under all four approaches, the treatment effect is robust – I find the

[44]Standard errors are corrected for heteroskedasticity and for within-subject correlation.

[45]I also run poisson regressions on a count variable (counting the number of transfers made by an individual in the relevant three periods). The qualitative results are the same as with OLS and *atranslog*.

[46]I realize that for 2^{nd} period data this may not be completely true if subjects fail to realize that it is a different game they are playing in the high-endowment periods.

		B		BH		effect size	
	Data	mean	std.dev.	mean	std.dev.	odds ratio	Cohen's d
1,3,5	1^{st}-period	.583	.5036	.917	.2823	1.730	.8179
	average	.528	.4495	.944	.2123	1.788	1.1827
	dominant	.5	.5108	.958	.2041	1.916	1.1772
	all periods	.528	.5027	.944	.2306	1.788	1.0629
2,4,6	2^{nd}-period	.292	.4643	.708	.4643	2.425	1.1118
	average	.222	.3764	.681	.3330	3.068	1.2924
	dominant	.25	.4423	.708	.4643	2.832	1.0107
	all periods	.222	.4187	.681	.4695	3.068	1.0309

Table 2.6: Effect-size indices.

	Periods 1,3,5			
	clustered	averaged	1^{st}-period	dominant
age	**-.0761** (.001)	**-.0761** (.002)	**-.0751** (.023)	**-.0720** (.006)
male	**.0335** (.785)	**.0335** (.791)	**.1024** (.465)	**.0526** (706)
econ	**-.1990** (.056)	**-.1990** (.064)	**-.1349** (.298)	**-.2221** (.079)
BHtreat	**.3019** (.007)	**.3019** (.009)	**.2453** (.072)	**.3334** (.014)
const	**2.3445** (.000)	**2.3445** (.000)	**2.2788** (.004)	**2.2284** (.001)
mean $\widehat{p}$(y=1)	**.7361**	**.7361**	**.75**	**.7292**
# of obs.	144	48	48	48
joint p-value	.000	.000	.006	.000

Table 2.7: Results from *clustered* regressions vs. regressions on *averaged*, 1^{st}*-period*, and *dominant-choice* data from low-endowment periods.

	Periods 2,4,6			
	clustered	averaged	2^{nd}-period	dominant
age	**-.0452** (.053)	**-.0452** (.061)	**-.0454** (.166)	**-.0449** (.158)
male	**.3008** (.007)	**.3008** (.009)	**.1888** (.243)	**.3011** (.044)
econ	**-.0773** (.498)	**-.0773** (.511)	**-.2425** (.138)	**-.0448** (.785)
BHtreat	**.3972** (.001)	**.3972** (.002)	**.2844** (.083)	**.4122** (.015)
const	**1.0597** (.046)	**1.0597** (.046)	**1.3414** (.067)	**1.0561** (.139)
mean $\widehat{p}$(y=1)	**.4514**	**.4514**	**.5**	**.4792**
# of obs.	144	48	48	48
joint p-value	.000	.000	.003	.000

Table 2.8: Results from *clustered* regressions vs. regressions on *averaged*, 1^{st}*-period*, and *dominant-choice* data from high-endowment periods.

treatment dummy *BHtreat* significant. Except for the 1^{st}- and 2^{nd}-period data, it is significant at the 5% level. The sign on the coefficient is positive in all cases and for both low- and high-endowment periods, meaning that the transfer rate is higher in the high-incentive treatment.

As regards other explanatory variables, the results for low-endowment periods are very similar to results from clustered regressions – *age* is significant at 5%, *male* is not significant, and *econ* is significant only at the 10% level and not significant for the 1^{st}-period data.

For high-endowment periods, the the results are not as robust. *Age* is significant at the 10% level for clustered and averaged data, and not significant for the 2^{nd}-period and dominant-choice data. *Male* is significant at the 5% level with the exception of the 2^{nd}-period data where it is not significant. *Econ* is never significant. The results suggest that there is larger variation in high-endowment data, which is more difficult to explain by the explanatory variables. I believe that a larger sample size would lead to more robust results.

As regards the sizes and signs of coefficients, the results are very robust, especially for clustered, averaged and dominant choice data.

B. Pooling the Sessions

In addition to the benchmark treatment B and to the high-incentive treatment BH, I conducted two plus two sessions of "automatic" treatments A and AI. Under both treatments, A and AI, I used the same game and same parameterization as in the B treatment. The only difference was that in automatic treatments, each subject played against a computer program, with six subjects in the role of Participant X and six subjects in the role of Participant Y. The computer program always played a (subgame perfect) optimal strategy. Subjects were acquainted with these facts in the instructions.

The only difference between the A and AI treatments was that in the AI subjects received, as a separate part of the instructions, a so-called Backwards Induction Tutorial, intended to explain the basic principles of using backwards induction.

Before pooling the data from different treatments I performed basic statistical tests in order to discover significant differences in the distributions of choices – Fisher's Exact test and Wilcoxon rank-sum test. I find no evidence of significant differences in the distributions of choices between the A, AI and B treatments.

Afterwards, I performed two types of pooled analysis: 1) pooling the data from the A and B treatments vs. the BH treatment and 2) pooling the data from the A, AI and B treatments vs. the BH treatment. My main result, the significance of treatment dummy at 5%, remains unaffected.

	Periods 1,3,5			Periods 2,4,6		
	B vs. BH	B,A vs. BH	B,A,AI vs. BH	B vs. BH	B,A vs. BH	B,A,AI vs. BH
age	**-.0761**	**-.0652**	**-.0376**	**-.0452**	**-.0229**	**-.0186**
	(.001)	(.007)	(.160)	(.053)	(.338)	(.418)
male	**.0335**	**-.0097**	**-.0178**	**.3008**	**.1726**	**.1079**
	(.785)	(.932)	(.863)	(.007)	(.128)	(.317)
econ	**-.1990**	**-.2152**	**-.1572**	**-.0773**	**-.1638**	**-.1365**
	(.056)	(.026)	(.117)	(.498)	(.127)	(.164)
BHtreat	**.3019**	**.2898**	**.3550**	**.3972**	**.2949**	**.2831**
	(.007)	(.001)	(.000)	(.001)	(.007)	(.006)
const	**2.3445**	**2.1577**	**1.4756**	**1.0597**	**.7994**	**.7613**
	(.000)	(.000)	(.016)	(.046)	(.133)	(.130)
mean $\widehat{p}$(y=1)	**.7361**	**.7111**	**.6806**	**.4514**	**.4556**	**.4491**
# of obs.	144	180	216	144	180	216
joint p-value	.000	.000	.000	.000	.000	.002

Table 2.9: Results from the estimation of basic vs. extended data sets.

See Table 2.9 for the regression results for the low- and high-endowment periods. Clearly, pooling slightly different treatments leads to noisier results, which is not very surprising. Importantly, the treatment dummy remains significant at the 5% level.

Chapter 3

Testing Leniency Programs Experimentally: The Impact of “Natural” Framing

with Andreas Ortmann

Abstract

We study the effects of loaded instructions in a bribery experiment. We find a strong gender effect: men and women react differently to real-world framing. The treatment effect becomes significant once we allow for gender-specific coefficients. This chapter contributes to the (small) literature on experimental tests of (anti-)corruption measures and adds evidence to the (mixed) results on gender effects and the on-going discussion on the need for socio-demographic controls.

Keywords: corruption, anti-corruption mechanisms, optimal contract, monitoring
JEL classification: C91, D02, D73, K42

3.1 Introduction

The severe consequences of corruption have been widely documented in the empirical literature. For example, Mauro (1995) and Tanzi (1998) have shown a negative effect of corruption on economic growth; Hwang (2002) has demonstrated that corruption, through tax evasion, reduces government revenues; and Gupta, Davoodi and Alonso-Terme (2002) conclude that corruption increases income inequality and poverty. The design and implementation of effective anti-corruption measures therefore remains an important concern.

One promising anti-corruption measure is the leniency policy. Leniency policies award fine reductions of varying intensities to wrongdoers who "spontaneously" report an illegal agreement and thereby help to convict their accomplice(s). They serve as an enforcement mechanism as much as a means of deterrence in that, if appropriately designed and implemented, they have the potential to undermine the trust between wrongdoers. Leniency policies have been analyzed in the literature mostly as an anti-cartel mechanism.

Spagnolo (2004), for example, theoretically examines the effects of leniency policies of various degrees – from moderate (which reduce or cancel the penalty for a criminal who reports) to full (which, in addition, pay a reward). He shows that reward-paying leniency programs provide a (socially) costless[1] and very efficient measure for cartel deterrence. Drawing on earlier versions of Spagnolo (2004), Apesteguia, Dufwenberg and Selten (2004) conducted an experiment that confirms the promising cartel-deterring properties of leniency policies.

Bigoni, Fridolfsson, Le Coq and Spagnolo (2008 a,b) conducted related experiments. They find that without leniency past convictions reduce the number of cartels but increase collusive prices. Their results also suggest that past experience might have more important consequences for the perception of risk than an exogenous probability of detection, and that strategic risk plays a key role in the effectiveness of a leniency policy. In general, the deterrence is higher with leniency in place and rewards lead to almost complete deterrence.

The work of Bigoni et al. (2008 a,b) contributes to a better understanding of the cartel-deterring properties of leniency policies and highlights the importance of proper policy design.

[1]This is the case if the rewards are fully financed from fines imposed on other convicted members of the gang.

Leniency policies to deter cartels are not directly applicable as anti-corruption measures, since cartel deterrence is a simultaneous game while strategies, payoffs, and the move structure of anti-corruption measures are asymmetric.[2] A proper theoretical and experimental analysis is therefore called for.

To our knowledge the first theoretical work analyzing the various effects of leniency policies on corruption is Buccirossi and Spagnolo (2006). The authors show that poorly designed moderate policies may have a serious counter-productive effect: they might allow a briber to punish at relatively low cost a partner who does not respect an illegal agreement. In other words, some leniency policies might actually provide an enforcement mechanism for occasional illegal transactions.[3] Thus they can, contrary to the intention, increase corruption.

Buccirossi and Spagnolo's result together with the theoretical and experimental evidence from the literature on cartel deterrence suggests that the potential of leniency policies to undermine trust between wrongdoers hinges upon proper design and implementation.

Experimental methods have been widely used, albeit rarely, to study corruption (Dušek, Ortmann and Lízal 2005). They become especially useful when counterfactual institutional arrangements such as leniency programs need to be explored: they provide, for example, relatively cheap ways to examine the effects of such arrangements in controlled environments (see Dušek et al. 2005, Apesteguia et al. 2004, Buccirossi and Spagnolo 2006, Bigoni et al. 2008 a,b, Richmanová and Ortmann 2008 and also Roth 2002).

In Richmanová and Ortmann (2008) we proposed a generalization of the Buccirossi and Spagnolo (2006) model in which we introduce the probabilistic discovery of evidence by auditing inspectors.[4] Our generalization makes the model more realistic and more readily applicable for experimental testing without changing the qualitative results of Buccirossi and Spagnolo.

We use the generalized Buccirossi and Spagnolo model for the experimen-

[2]For a more detailed discussion see Richmanová (2006).

[3]Occasional illegal transactions are essentially one-shot transactions.

[4]In the original model, Buccirossi and Spagnolo assume that a briber and a bribee agree to produce hard evidence, which serves as a hostage. Without hard evidence being produced, the occasional illegal transaction is not enforceable. An audit, if it takes place, discovers the evidence with a probability of one. In Richmanová and Ortmann (2008), we argue that instead some evidence is created unintentionally and this can be discovered by the audit with some probability that is less than one.

tal testing of leniency policies as an anti-corruption measure. This chapter, and the closely related work reported in Krajčová (2008) (in Chapter 2), provide a new testbed for anti-corruption programs and address important methodological issues with which (anti-)corruption experiments are afflicted. Specifically, we are interested in to what extent home-grown priors that are related to corruption might translate into moral scruples and, for example, might induce subjects to make different decisions when loaded instructions are used that make it unambiguously clear what the context of the experiment is.[5]

Abbink and Hennig-Schmidt (2006) is the study most closely related to this chapter. In a bribery setting, these authors used a between-subject design with one treatment providing instructions framed in neutral terms and the other treatment "loading" the instructions with real-world framing. The authors used a finitely repeated reciprocity game as the bribery setting. With a very low probability (0.3%) bribery would be detected leading to both parties involved in the transaction being excluded from the experiment and forfeiting their accumulated earnings. In addition, a completed transaction imposed a negative externality on the public, in Abbink and Hennig-Schmidt (2006) represented by the other subjects in the session.

The authors find a treatment effect which is, however, not statistically significant;[6] they argue that this result may be caused by the nature of the

[5]Altogether, we designed three treatments: a benchmark, which is common to both studies (this chapter and Krajčová 2008) and in which all instructions are presented in neutral language; a context treatment, in which we use the same parameterization as in the benchmark but instructions are presented in real-world framing (this chapter); and, finally, a high-incentive treatment, which implements a new parameterization within neutral framing (Krajčová 2008). Krajčová (2008) examines the effect of a change in parameterization. It has been documented in the literature that a change in parameterization that does not affect the theoretical prediction might indeed have consequences for the behavior of subjects in the lab (e.g. Goeree and Holt 2001). In the generalized Buccirossi and Spagnolo game, the action bringing the highest possible payoff is also associated with a risk of a considerable loss. Therefore, it is possible that subjects in the lab will not behave in accordance with the theoretical predictions, especially when the predictions are made under the assumption of risk neutrality. Krajčová (2008) indeed finds a significant effect of the parametric change.

[6]This result is questionable. When looking at the evolutions of bribe offers and of permission frequencies, especially in the 20 central rounds, a clear difference between the two treatments is visible: for each round the average bribe offer is higher with neutral than with loaded instructions. The first rounds might likely be affected by a learning effect of the first kind (e.g., subjects becoming familiar with the lab setting rather than

game: it is very simple, and thus, even with neutral wording, subjects may have deciphered that the experiment was about corruption and corruptibility. The generalized Buccirossi and Spagnolo game is more complex (e.g., it involves the realization of two random outcomes) and therefore is likely to be less susceptible to inferences about the true nature of the strategic interactions.

Our results confirm that loading the instructions with real-world framing might affect subjects' behavior in a significant manner. Moreover, we find a strong gender effect - male and female participants react to a bribery context differently. This treatment effect becomes significant once we allow for gender-specific effects.

The remainder of this chapter is organized as follows. In the next section, we discuss the generalized Buccirossi and Spagnolo model in detail, and we also describe and compare our two experimental treatments. In Section 3 we explain how we implemented our two treatments and in Section 4 we discuss our results. Section 5 concludes.

3.2 Experimental Design

The experiment implements the bribery game in Richmanová and Ortmann (2008). An entrepreneur has an investment possibility of net present value v, if a bureaucrat is willing to perform an illegal action, *Action a*. For doing so, the bureaucrat may require compensation in the form of a bribe, b.

The timing of the game is as follows. First, the entrepreneur decides whether to *Pay* or *Not Pay* a bribe. If she does not pay a bribe, the game ends. If she does, the bureaucrat chooses one of three possible actions: *Denounce*, do *Nothing*,[7] or perform *Action a*.[8]

reacting to incentives, see e.g. Hertwig and Ortmann 2001) and the last five rounds by a possible termination effect, which is acknowledged by the authors. This treatment effect is confirmed by both the Wilcoxon rank-sum test and effect sizes for the 20 central rounds. It would likely become significant, even without excluding the possibly problematic first and last five periods, with a larger number of observations.

[7] *Nothing* denotes a passive action choice. For the bureaucrat, it means that he neither denounces nor respects (by providing the favor) the illegal agreement. For the entrepreneur, it means that she does not denounce in response to the bureaucrat's action.

[8] *Action a* means that the bureaucrat respects the illegal agreement and thus provides a (illegal) favor to the entrepreneur. That is, strictly speaking, not a corrupt action because it does not impose a negative externality on the public. According to Abbink, Irlenbusch

If the bureaucrat chooses *Denounce*, an audit is carried out. The audit may (with probability β, $\beta \in (0,1)$), or may not (with probability $1-\beta$), discover some evidence of bribery. If the bribery attempt is detected, the leniency policy guarantees that the bureaucrat will have to pay only a reduced fine whereas the entrepreneur will have to pay the full fine. In addition, bribe b is confiscated.[9] If the bribery is not detected, the bureaucrat will enjoy bribe b.

If the bureaucrat chooses *Nothing* or *Action a*, the entrepreneur has another move. In both cases, she may choose between *Denounce* and do *Nothing*.

If the entrepreneur chooses *Denounce* and the ensuing audit discovers evidence (which, again, happens with probability β), then she will have to pay a reduced fine whereas the bureaucrat will have to pay the full fine and, in addition, their illegal gains will be confiscated. If no evidence is discovered, both the bureaucrat and the entrepreneur will keep their illegal gains.

If the entrepreneur chooses *Nothing*, then an audit may still occur with some nonzero probability α. If the audit detects bribery (which happens with probability β), both parties are subject to a sanction, which consists of the confiscation of the illegal gains plus the full fine. The illegal gains include bribe b in any case and value v only in the case when the bureaucrat has chosen to perform *Action a*.

Figure 3.1 summarizes the extensive form of the game and the expected payoffs.

The contribution of the generalized model lies in the introduction of probability β. In Buccirossi and Spagnolo (2006) it is assumed that, before the illegal transaction takes place, the bureaucrat and the entrepreneur agree on the production of hard evidence. Without hard evidence being voluntarily produced by both of them the illegal transaction is not enforceable. In essence it is assumed that both involved are holding a hostage that commits each other to the desired outcome. It is furthermore assumed that, if an audit takes place, corruption is discovered and both culprits are convicted with a probability of one. Richmanová and Ortmann (2008) assume instead that some hard evidence is created unintentionally along the way and that this evidence may be discovered by an audit with probability $\beta \in (0,1)$. The basic

and Renner (2002) it is not such a problem since people do not care much about the costs they impose on others.

[9] Note that in this case the illegal transaction has been detected without *Action a* being performed and therefore there is no gain to the entrepreneur to be confiscated.

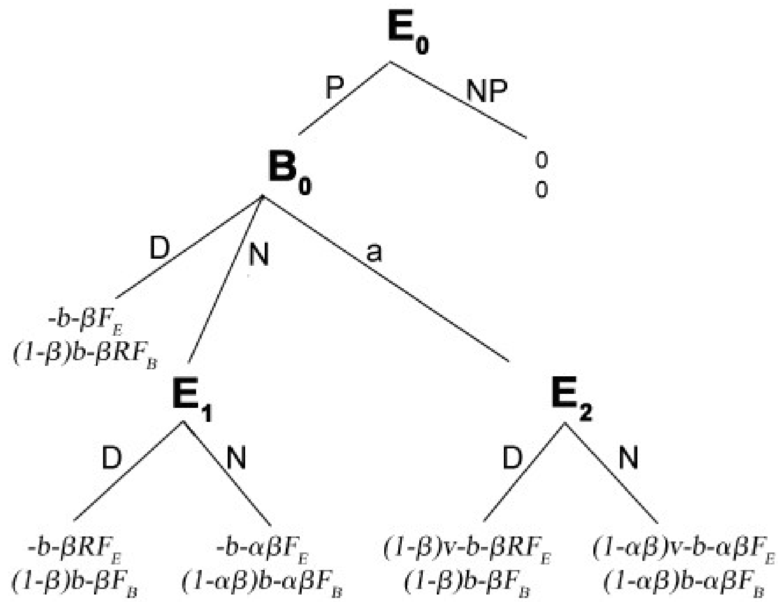

Figure 3.1: Extensive form of the corruption game in the generalized model. P stands for *Pay*, NP for *Not Pay*, D for *Denounce*, N for doing *Nothing*, a for performing *Action a*, b for a bribe, v for the value of the project to the entrepreneur, α for the exogenous probability of an audit, β for the probability of conviction, F_E and F_B for full fines and RF_E and RF_B for reduced fines to the entrepreneur and the bureaucrat, respectively.

structure of both the original and the modified game is the same except that in the original version the probability β is set to 1. The generalization makes the model more suitable for experimental testing, as no additional stage is needed in which subjects would have to agree on producing a hostage. In addition, the generalized model arguably resembles real-world situations more closely.[10]

Buccirossi and Spagnolo (2006) show that in the absence of a leniency program, occasional illegal transactions are not implementable.[11] The result

[10] We realize that in such a game beliefs about the probability of detection might play an important role. However, we believe that the introduction of beliefs would make the game more complex than necessary for experimental testing. Instead, we view probability β as an empirical success rate, or effectiveness, of a detection technology that is known to subjects.

[11] Facing the full fine even after reporting, the entrepreneur cannot credibly threaten to report the bureaucrat in the case when he would not deliver. Therefore, the bureaucrat would keep the bribe and not perform *Action a*, knowing that it is not profitable for the

carries over into the generalized model. After the introduction of a modest leniency program,[12] occasional illegal transactions are enforceable if the following three conditions are satisfied simultaneously. First, the *no-reporting condition* for the bureaucrat: the reduced fine must be such that the bureaucrat prefers performing *Action a* to *Denouncing* once the bribe has been paid. Second, the *credible-threat condition* for the entrepreneur: the reduced fine and the full fine must be set such that the entrepreneur can credibly threaten to report if the bureaucrat does not deliver. Third, the *credible-promise condition*: the entrepreneur must be able to credibly promise not to report if the bureaucrat respects the illegal agreement.

These three conditions, given the value of the project together with the full and reduced fines, define a bribe range for which the occasional illegal transaction is implementable. Even though these conditions are modified in the generalized model, the qualitative result remains unaffected.

We used the generalized version of the game for experimental testing of the theoretical prediction under two different scenarios: when the occasional illegal transaction is implementable in equilibrium, and when it is not. Implementability is a function of the per-round endowment for the entrepreneur. The per-round endowment exogenously defines the value of the bribe if the entrepreneur decides to pay it.[13] For each treatment we use two possible values of the per-round endowment: a low endowment that theoretically leads to a no-corruption equilibrium, and a high endowment that theoretically leads to a corruption equilibrium.

Following Abbink and Hennig-Schmidt (2006), the focus of this chapter is the question whether loaded instructions in a bribery experiment affect the behavior of subjects in a lab. For that purpose, we designed two treatments: a Benchmark (B) and a Context (C) treatment.[14]

---

entrepreneur to punish him. Consequently, the entrepreneur would not enter the illegal agreement in the first place.

[12] Similarly to Spagnolo (2004), "modest" means that a leniency program does not reward for reporting, at best it cancels the fine.

[13] This way we reduce the cognitive demand on subjects: the only decision they have to make is whether they want to transfer their per-round endowment or not.

[14] We also conducted two exploratory sessions of a partial context treatment (C- treatment), where we only provided context on the types of roles. In this treatment Participant X was called "Entrepreneur" and Participant Y "Bureaucrat". All actions and realizations of random outcomes were denoted by neutral letters, as in the B treatment. We do not report these data in the main text as it is not possible to control for subjects' interpretation of the game in this case and therefore it is hard to recognize all the possible effects

Table 3.1 summarizes the parameterization chosen for both treatments.

Treatment	α	β	**v**	$\mathbf{RF}_E$	$\mathbf{RF}_B$	$\mathbf{F}_E$	$\mathbf{F}_B$	$\mathbf{E}_L$	$\mathbf{E}_H$	**show-up**
B and **C**	0.1	0.2	100	0	0	300	300	20	40	300

Table 3.1: Experimental parameterization. α and β denote the probability of an audit and of discovering evidence of bribery, respectively; v denotes the value of the project to the entrepreneur; RF_E and RF_B denote reduced fines and F_E and F_B full fines to the entrepreneur and to the bureaucrat, respectively; E_L and E_H are low and high per-round endowment, respectively; and show-up stands for the show-up fee.

The probabilities α and β were chosen such that they approximately correspond to the real world exogenous probabilities of an audit and to real-world conviction rates; at the same time they are intuitively comprehensible for subjects. The value of the project v was chosen together with the full fines F_E and F_B such that the subject faces a considerable gain from the investment but also severe punishment in the case of detection. We set reduced fines RF_E and RF_B equal to zero to analyze the case of full leniency programs that, according to Apesteguia et al. (2004), have shown promising anti-cartel properties. Endowment determines the value of the bribe to be (not) paid. The "low endowment" of 20 leads (theoretically) to a no-corruption equilibrium, whereas the "high endowment" of 40 leads to a corruption equilibrium. Finally, the show-up fee was set such that we eliminate the possibility of earning a negative total from the experiment. The parameterization does not differ between B and C treatments as we are interested purely in the effects of "neutral" and "loaded" instructions.

Extended game forms and expected payoffs resulting from our parameterization are illustrated in Figure 3.2 for both low- and for high-endowment periods. The branches identifying the equilibrium choices of risk-neutral agents are in bold.

The instructions for the B treatment were presented in a completely context-free fashion. Subjects were called Participant X and Participant Y, actions were denoted by neutral letters and the realization of "detection" or "no detection" as "outcome A" or "outcome B", respectively.

In the C treatment, the roles that subjects were assigned were called *"Entrepreneur"* and *"Bureaucrat"*; actions were called *"Pay bribe"*, *"Not Pay bribe"*, *"Denounce"*, *"do Nothing"* and *"Provide the favor a"*; and the realizations of random outcomes were called *"corruption has been detected"* and

in this treatment. Some results from this treatment are discussed in the appendix.

"corruption has not been detected". Figure 3.3 provides a comparison of the wording for the two treatments, with the neutral wording always in the upper row.

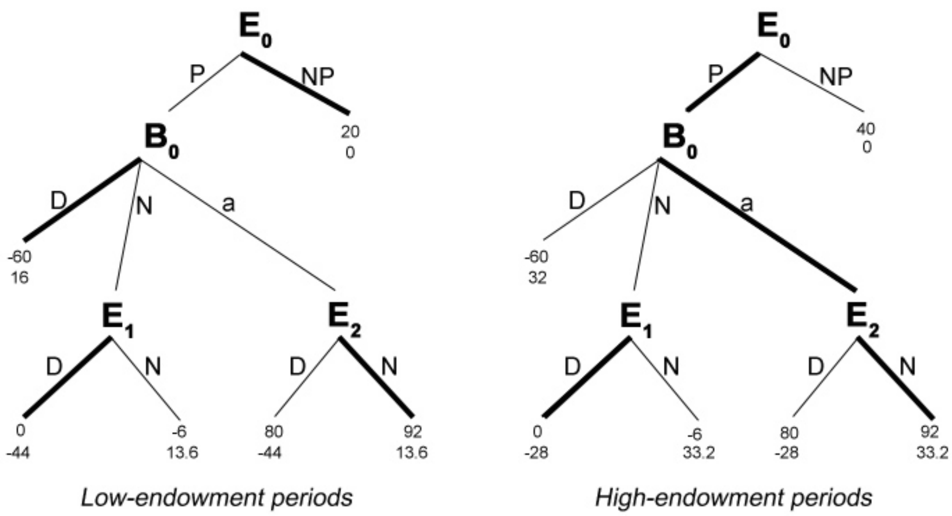

Figure 3.2: Expected payoffs from the corruption game.

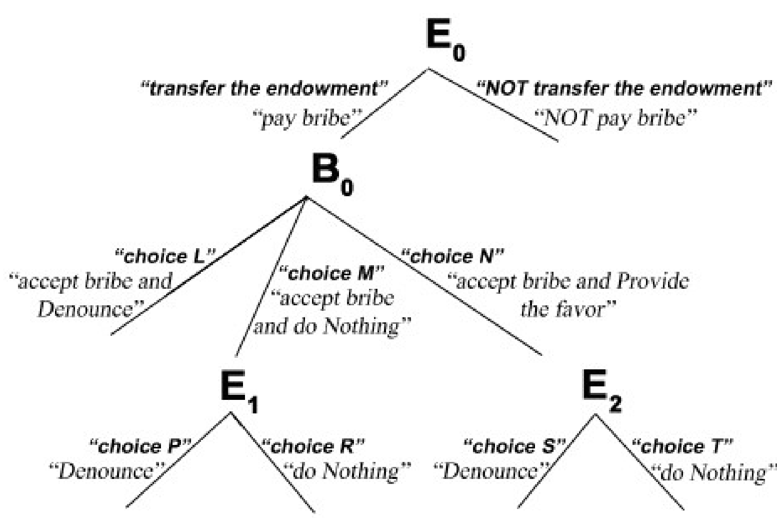

Figure 3.3: Neutral vs. loaded instructions wording. For each branch, the upper line provides the neutral labels of the B treatment (bolded); below are the loaded labels of the C treatment.

3.3 Implementation

The experiment was conducted in November and December 2006 at CERGE-EI in Prague, using a mobile experimental laboratory.[15]

Participants were recruited from the Faculty of Social Sciences of Charles University in Prague, from various faculties of the Czech Technical University in Prague and of the University of Economics in Prague. Students were approached via posters distributed on campus and via e-mail.[16]

Treatment	Subject Source[17]	M/F ratio[18]	mean (age)	mean (RA score)	mean (final pay)	Irreg[19]
B	FSS	8/4	20.9	29.7	320	1
B	FSS	10/2	21.75	28.8	330	0
B	CTU	11/1	22.9	34.7	330	0
B	FSS	9/3	22.3	26.4	323.3	0
C	CTU	9/3	21.9	33.7	340	0
C	UE	7/5	22.9	28	318.3	1
C	CTU	10/2	23	31.4	318.3	0
C	UE	7/5	21.7	28.1	315	0

Table 3.2: Summary of the demographic characteristics of subjects for all eight sessions.

We conducted four sessions of each treatment. Twelve participants, six in the role of Participant X – the entrepreneur – and six in the role of Participant Y – the bureaucrat – interacted in each session. In each session, all subjects participated in six rounds during which they kept the role that was assigned to them at the beginning of the first round.[20] Participants were anonymously

[15]http://home.cerge-ei.cz/ortmann/BA-PEL.htm

[16]By e-mail, we also directly invited students from these schools who participated earlier in unrelated experiments conducted at CERGE-EI.

[17]For each session, subjects were recruited from one source. FSS stands for the Faculty of Social Sciences in Prague, CTU for the Czech Technical University in Prague, UE for the University of Economics in Prague. We control for imbalance of the subject pool by including the econ and gender dummies in the econometric analysis.

[18]Male/Female ratio in the session.

[19]Irreg stands for a dummy variable for session irregularities. It identifies any unusual activities by subjects or any irregularities on the experimenter's side. In the first B treatment session an experimenter effect is possible; in the second C treatment session, one of the subjects reports "building engineering" as a field of study in a demographics questionnaire, which may mean that a CTU student participated in a UE session. We do not believe that these irregularities matter but wanted to control nonetheless. After running the preliminary regressions we concluded that they indeed did not matter.

[20]After each Participant X interacted exactly once with each Participant Y, the roles were switched for another six rounds. Subjects were not informed about the switch of

re-matched so that no subject was matched twice with the same co-player. This was common knowledge. The incentive compatibility of this matching scheme is discussed in Kamecke (1997).

Table 3.2 summarizes the demographic characteristics of subjects participating in the experiment. The majority of our subjects were male, reflecting the composition of the subject pools that we drew on. Mean age ranged between 20.9 and 23, over all sessions the minimum is 18 and the maximum is 29. We also measured subjects' risk aversion using a questionnaire based on Holt and Laury (2002). Mean RA score in the sample ranged between 26.4 and 34.7, over all sessions the minimum is 13 and maximum 51.[21] Average final payoffs for the B treatment ranged from 320 to 330, with the minimum being 300 and the maximum 400; for the C treatment it ranged between 315 and 340, with the minimum being 300 and the maximum 400.[22]

Each session began with general instructions. Afterwards, subjects were asked to fill in Risk-aversion and Demographics questionnaires, for which they earned their show-up fee. Then the instructions to the computerized part of the experiment were distributed. Understanding of the instructions was tested by a brief questionnaire. The computerized part of the experiment started only after every participant answered all testing questions correctly.[23] Each session concluded with a questionnaire asking for the subject's feedback on the experiment.[24]

---

roles in advance in order to avoid a possible impact on their behavior in the first six rounds. Before the beginning of the seventh round the announcement about the switch of roles appeared on their screens. The decisions in the last six rounds are likely affected by subjects' experience from the first six rounds and therefore we do not report them in the main text. A comparison of the before-switch and after-switch data is provided in the appendix. For both treatments, we observe more transferring in the after-switch data. In the B treatment, we also observe more denouncing in both the second and the third stage. In the C treatment, the differences for the second- and the third-stage data are very small.

[21]The higher the score, the more risk averse the subject is. The maximum possible RA score is 60 which, using the standard CRRA utility function $x^{(1-r)}$, approximately corresponds to a relative risk aversion coefficient of .17. The minimum possible RA score is 0, which approximately corresponds to a relative risk aversion coefficient of $-.13$. An RA score of 23 corresponds to risk-neutrality.

[22]At that point 400 CZK corresponded to about 16 USD, in purchasing power up to twice as much. Subjects were informed during recruitment that their final payoff from the experiment might be zero, but could not be negative. The non-negativity of the final payoff was ensured by the show-up fee.

[23]This was common knowledge.

[24]For filling this last questionnaire, subjects were paid an additional 50-200 CZK (cor-

All instructions were read aloud by the experimenter. As a part of the instructions, subjects received a pictorial representation of the game with a minimum use of game-theoretic terminology. Probabilistic outcomes were presented in both probabilistic terms and frequency representation (see e.g. Gigerenzer and Hoffrage 1995, or Hertwig and Ortmann 2004).[25]

The experiment was computerized using Z-tree software (Fischbacher 2007). At the beginning of each round, each participant was notified of her/his role. Participants X also learned their current per-round endowment. Afterwards, each pair interacted sequentially.[26] Between the second and the third stage, Participants X were asked about their choices at each node of the third stage if they were to reach it. After they made their conditional choices, they learned the actual decision of their co-player and they were asked to confirm, or to change, their previous choice. This mechanism allowed us to collect some additional data in rounds when the third stage was not reached.

At the end of each round subjects were given feedback about their action(s), the action(s) of the player they were paired with, the realization of the random outcome (detection vs. no detection, or outcome A vs. outcome B) and their resulting payoff. At the end, one round was randomly chosen to determine the final payoff from the computerized part of the experiment. This mechanism was chosen in order to ensure that the decision in every round is made as if in a one-shot game. This payment procedure was common knowledge ex-ante.

Participants were paid anonymously in cash right after each session. We used the Czech crown as the currency unit throughout the whole experiment.

3.4 Results

3.4.1 Summary Data

In Figure 3.4, the results from low- and high-endowment periods are presented. Each figure integrates the results from both treatments – the B

responded to about 2-9 USD) - the amount varied between sessions. This mechanism was used to adjust average earnings per session to the levels promised during recruitment.

[25]Originals (in Czech) of all materials that subjects received during the experiment are available at http://home.cerge-ei.cz/richmanova/WorkInProgress.html.

[26]Choices were made by clicking the respective buttons on the screen. Subjects were notified that once they made their choice it would not be possible to take it back.

treatment data in the upper rows and C treatment data below. The equilibrium choices for each case are in bold face.

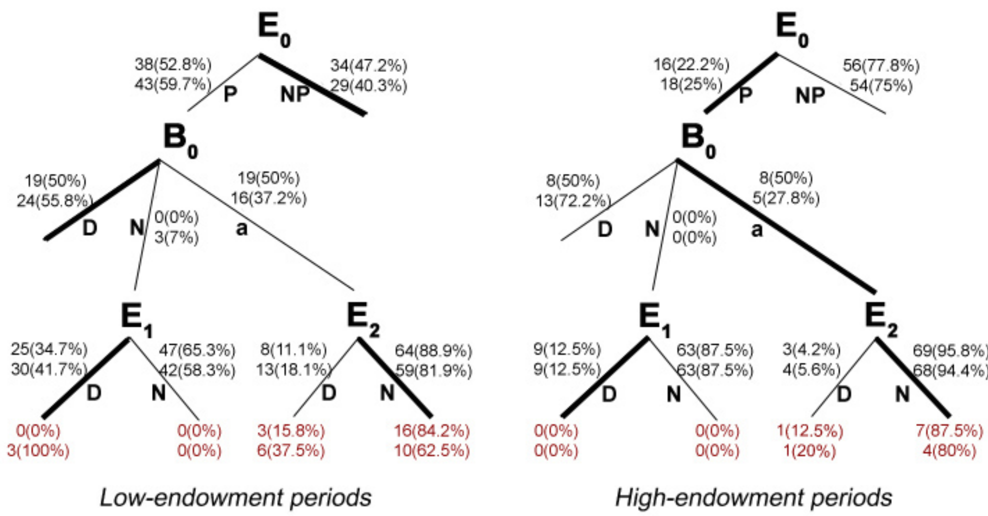

Figure 3.4: Experimental results. For each branch of the extensive form of the game, the upper row always displays the frequency of the action in the B treatment, while the lower row displays the frequency of the action in the C treatment (with the corresponding percentage in parentheses). For stages E_1 and E_2, above the branches, we present the conditional choices subjects were asked to report before they made their actual choice. The frequencies of real choices, which depend on the preceding decision of Participant Y, are presented at the bottom part of each figure.

For the aggregate first-stage data we observe surprisingly small differences between the two treatments. In addition, in both treatments, the frequencies of choosing *Pay* are higher in the low-endowment periods than in the high-endowment periods, which contradicts the theoretical prediction. Intuitively, subjects seem to be willing to transfer their (low) endowment to have a chance of receiving a high payoff, but their willingness to risk their endowment is limited. Instead of risking the high endowment, they seem to prefer the sure outcome.

As to the second-stage data, it is only relative percentages that can be compared across treatments, as the numbers of subjects that entered this stage of the game varied. In the B treatment, subjects split their choices evenly between playing *Denounce* or *Action a* for both low- and high-endowment periods. The difference in expected payoffs resulting from *Denounce* and *Action a* is, however, very small and might explain why we do not observe a stronger inclination to either choice. Also note that in both treatments *Denounce* is the only action through which the bureaucrat can avoid a negative expected round-payoff with certainty.[27]

In the C treatment, choices are shifted in favor of *Denounce*. Arguably, in the high-endowment periods, this result contradicts the theoretical prediction, but it is in line with our conjecture – knowing the context of their action choice, reporting corruption might be more attractive for subjects.

In line with the theoretical prediction and also our intuition, *Nothing*[28] was almost never chosen.

As to the third-stage data, conditional choices provide mixed evidence. In both treatments, subjects seem to prefer playing *Nothing* in either case. For the E_1 node that contradicts the theoretical prediction, while it is in line with the theoretical prediction for the E_2 node. When we look at the sequential choices, the results seem in line with the theoretical prediction for both treatments, inferring from relatively few observations.[29] We observe essentially no framing effect for high-endowment periods. For low-endowment periods, we observe a small shift in favor of *Denounce*, which is in line with our expectations.

Note that for the second- and the third-stage data we have too few independent observations (especially so for the high-endowment periods)[30] to perform a reliable formal analysis. Therefore, we only perform statistical and regression analyses of the first-stage data.

[27]See Figure 3.2 and Table 3.1 for more details. Even though the subject could possibly earn a negative round payoff, each subject also received a show-up fee which ensured a non-negative final payoff.

[28]Payoffs for Participant Y resulting from *Nothing* and *Action a* are the same, but taking into account the likely decisions of Participant X in the following stage, he is more likely to collect a higher payoff after choosing *Action a*.

[29]When we asked the subjects to make their real choices in the B treatment, only one of them changed her/his decision in the E_2 node from *Denounce* to *Nothing* (after observing what Participant Y has chosen) in the low-endowment period. In the C treatment, three subjects changed her/his decision in the E_2 node – two switched from *Nothing* to *Denounce* after Participant Y played *Action a* and one from *Denounce* to *Nothing* after Participant Y played *Action a* – and one subject changed her/his decision in the E_1 node from *Nothing* to *Denounce* after Participant Y played *Nothing*. All four cases occurred in low-endowment periods.

[30]Recall that Figure 3.4 presents the aggregated data from all relevant periods, therefore it contains repeated observations for individual subjects.

3.4.2 Analysis of the First-Stage Data

We performed standard non-parametric tests with the null hypothesis of no differences in the distributions of choices under the two treatments. We also computed the effect sizes to measure the magnitude of the treatment effect. Finally, we report the results from the estimation of a linear probability model in which we control for some demographic characteristics of subjects.

Due to the panel nature of the data, we considered four different approaches to formal regression analysis: 1) clustered data analysis – data from periods $1, 3$, and 5 (low-endowment) and from periods $2, 4$, and 6 (high-endowment) are clustered by subject to correct standard errors for likely within-subject correlation; 2) first-period data analysis – only first-period data (for the low-endowment case) and only second-period data (for the high-endowment case) are analyzed; 3) averaged data analysis – averaged data for periods $1, 3$, and 5 and for periods $2, 4$, and 6 are analyzed; and 4) dominant-choice data analysis – for each endowment value (low or high) each subject makes choices in three periods, and the dominant choice is the one that is played more often.

Clustered data have the advantage of using all the available information, while the other three approaches use only a part of the information we have. Therefore, in the main text we discuss the results for clustered data. The analysis of averaged, first-period, and dominant-choice data can be found in the appendix, as a robustness check of the main results. By and large, there are no major findings in these robustness tests.

In addition to the robustness checks based on different "data handling", we also run a few exploratory sessions of treatments in which the experimental conditions are only slightly modified compared to the benchmark and the context treatment. The results from the analysis on the extended data set is provided in the appendix, as an additional robustness check of the main results. By and large, there are no major findings in these robustness checks. Pooling slightly different treatments leads to noisier results, which is not very surprising.

Statistical Analysis

In Table 3.3 we report the results of three standard non-parametric tests in order to identify the differences in the distributions of choices under the two treatments. Specifically, we test the null hypothesis of no differences between the two treatments using the averages of the binary transfer variable[31] over periods $1, 3$, and 5 and $2, 4$, and 6.

periods	Ranksum[32]	Ksmirnov[33]	Fisher[34]
1,3,5	-.526 (.599)	.083 (.846)	(.947)
2,4,6	-.715 (.475)	.167 (.513)	(.218)

Table 3.3: Non-parametric tests.

According to Wilcoxon rank-sum, Kolmogorov-Smirnov, and Fisher's exact tests, we cannot reject the hypothesis of no differences in the distributions of choices under the two treatments at the 5% significance level.

To assess the magnitude of the effect for practical purposes, we in addition compute two standardized measures of effect size: Cohen's d and odds ratio, again, using the averages of the binary transfer variable over periods $1, 3$, and 5 and $2, 4$, and 6. We also want to look separately at male and female data to discover possible gender effects. The results for the full sample and for the male and female subsamples are reported in Table 3.4.

		B			C			effect size	
Periods	Sample	N	mean	std.dev.	N	mean	std.dev.	odds ratio	Cohen's d
1,3,5	full	24	.528	.4495	24	.597	.4282	1.131	.1571
	male	18	.519	.4461	17	.667	.4082	1.285	.346
	female	6	.556	.5018	7	.429	.4600	.772	-.264
2,4,6	full	24	.222	.3764	24	.25	.3147	1.126	.0807
	male	18	.296	.4105	17	.275	.3170	.929	-.057
	female	6	0	0	7	.190	.3253	NA[35]	.826

Table 3.4: Effect-size indices.

Cohen (1998) defines effect sizes of $d = 0.2$ as *small*, $d = 0.5$ as *medium*, and $d = 0.8$ as *large*. For the full sample, the results suggest only a small effect. However, when we look at the male and female subsamples separately, the effect size appears larger than in the full sample. It is also noticeable

[31]Transfer has a value of one if Participant X chooses *Pay* and a value of zero if s/he chooses *Not Pay* in the respective period.

[32]Ranksum stands for the two-sample Wilcoxon rank-sum (or Mann-Whitney) test. We report the normalized z statistic and corresponding p-value below.

[33]Ksmirnov stands for the Kolmogorov-Smirnov test. We report the statistic and below the corresponding p-value from testing the hypothesis that average transfer is lower in the B treatment.

[34]Fisher stands for Fisher's exact test. We report the resulting p-value.

[35]A division-by-zero problem occurs, due to no variation in this subsample.

that the effects for the male and for the female subsamples have opposite directions, which naturally results in a very small total effect. We observe very similar results when looking at the *odds ratio* – the effect is smaller in the full sample than in the two subsamples. These results suggest a non-negligible gender effect.

Altogether, both statistical tests and effect-size measures suggest that there are only minor differences between the first-stage choices in the C and B treatments. Effect-size measures for the male and female subsamples suggest that this result might be caused by counteracting gender effects. Therefore, further analysis which would control for gender and for other subjects' characteristics is called for.

Gender Differences

Before the estimation, we want to look more closely into the gender-specific data. Figures 3.5 and 3.6 provide the summary data separately for men and women.

For low-endowment periods, in the first stage of the B treatment the difference in the behavior of men and of women does not appear substantial – slightly more than half of each makes the transfer. However, in the C treatment, the transferring decisions of males and of females shift in opposite directions – two thirds of men, whereas less than a half of women, decide to make the transfer. This suggests that the corruption framing affects men and women differently.

Similarly in the second stage we can clearly see from the descriptive data that facing a full context, women become much more likely to report. Men's decisions seem to remain unaffected.

The results from the last stage are not so clearly distributed. In the E_1 node we observe the opposite effect of context on men than on women. In the E_2 node, the direction of the effect does not vary with gender. In general, both men and women prefer doing *Nothing* to *Denouncing*.

In the first stage of the high-endowment periods, the results are somewhat different. We still observe considerably more women refraining from making transfers but the framing effect seems to increase the transfer rate. Recall, however, that the results from the regression analysis suggest that these four observations might be just random realization. We observe almost no framing effect in the male subsample. In general, both men and women prefer not making the transfer.

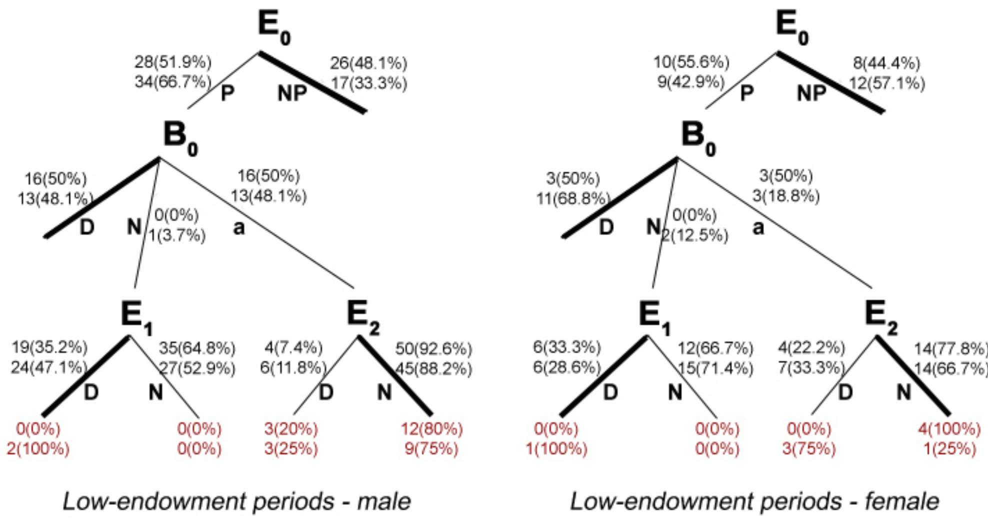

Figure 3.5: Experimental results for male and for female subjects in low-endowment periods. For each branch of the extensive form of the game, the upper row always displays the frequency of the action in the B treatment, while the lower row displays the frequency of the action in the C treatment (with the corresponding percentage in parentheses). For the nodes E_1 and E_2, above the branches, we present the conditional choices subjects were asked to report before they made their actual choice. The frequencies of real choices, which depend on the preceding decision of Participant Y, are presented at the bottom part of each figure.

In the second stage, the female subsample in the role of Participant Y is very small. In both treatments, all the women choose *Denounce*. For the male subsample, we observe some (possible) treatment effect, which shifts the choices more in favor of playing *Denounce* in the C treatment.

In the third stage the percentage of men choosing *Denounce* slightly decreases with framing, while for females it goes slightly up. In both subsamples, the prevailing choice is doing *Nothing*, though.

Econometric Analysis

In this section we report the results from econometric analysis controlling for some of the subjects' characteristics and for the treatment effect. We are also trying to confirm gender-specific effects.

During the experiment we distributed questionnaires in order to collect basic demographic data. Specifically, we have information about subjects'

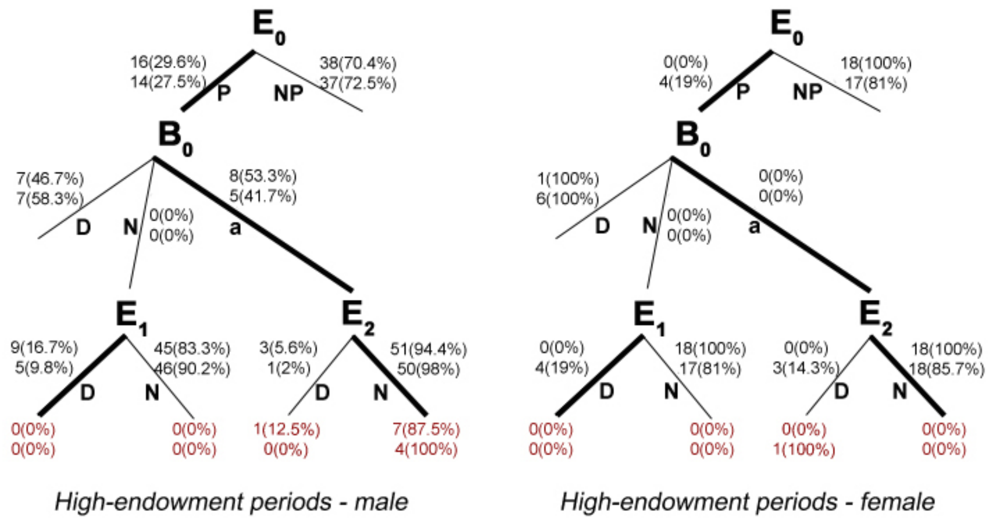

Figure 3.6: Experimental results for male and for female subjects in high-endowment periods. For each branch of the extensive form of the game, the upper row always displays the frequency of the action in the B treatment, while the lower row displays the frequency of the action in the C treatment (with the corresponding percentage in parentheses). For the nodes E_1 and E_2, above the branches, we present the conditional choices subjects were asked to report before they made their actual choice. Frequencies of real choices, which depend on the preceding decision of Participant Y, are presented at the bottom part of each figure.

age, gender, university and field of study.[36] We also measure each subject's risk aversion.

The dependent variable was defined as a 0/1 dummy variable *translog* identifying *Pay* being chosen (value of 1) or not (value of 0) in a particular period. We estimate a clustered linear probability model. We prefer a linear

[36]In addition, we collected data on: size of subject's household, number of cars in the household, and whether the subject himself has his own car and what is its approximate value, all of which serve as proxies for income. We also asked the subjects whether they considered themselves as technical types compared to their peers. We recorded the occurrence of any inconsistencies in the after-instructions questionnaire, which served as a simple test of understanding of the basic structure of the game, and in the risk-aversion questionnaire. At the end of the session we asked our subjects whether they understood the experiment. Finally, we recorded some general information about each session – the time of day it started and any session irregularities if they occurred. After running some preliminary regressions we, however, conclude that none of these variables is significant for explaining subjects' decisions. The demographic and the risk-aversion questionnaires are based on Rydval (2007).

probability model to other non-linear alternatives, as it does not rely on very specific distributional assumptions, the violation of which leads to inconsistent estimates if non-linear models are employed. Another advantage of the linear probability model lies in the straightforward interpretation of the estimated coefficients. We ran clustered robust estimation to correct standard errors for likely within-subject correlation and for heteroskedasticity.

In the appendix, we provide a discussion of the robustness checks we conducted in addition to the clustered regressions analysis. As the theoretical prediction differs for low- and for high-endowment periods,[37] these two groups were analyzed separately.

We start with a basic minimal model:[38]

$$P(translog = 1|x) = \beta_0 + \beta_1 \cdot age + \beta_2 \cdot male + \beta_3 \cdot econ + \beta_4 \cdot Ctreat,$$

where *age* corresponds to subject's reported age, *male* is a dummy variable defined based on subject's reported gender, and *econ* is a dummy variable identifying a subject having (value of 1) or not having (value of 0) an economic background, which is defined based on the subject's reported field of study. As we are mainly interested in the treatment effect, we also include a C-treatment dummy *Ctreat* in the model.

The results from the estimation are summarized in Table 3.5, denoted as Model 1. This model is, however, not significant. In the next step, we extend the basic minimal model by interaction terms with *male* to allow for gender-specific effects. This leads to Model 2:

$$P(translog = 1|x) = \beta_0 + \beta_1 \cdot age + \beta_2 \cdot male + \beta_3 \cdot econ + \beta_4 \cdot Ctreat +$$

$$+\beta_5 \cdot male * age + \beta_6 \cdot male * econ + \beta_7 \cdot male * Ctreat.$$

[37] Recall that in periods $1, 3$, and 5 the endowment was low, while in periods $2, 4$, and 6 the endowment was high.

[38] The second approach we used was $P(translog = 1|x) = \beta_0 + \beta_1 \cdot ra\_score$, where ra_score is a risk aversion score computed based on data from the risk-aversion questionnaire. Preliminary analysis suggests that age, male and econ predict ra_score well (all three are jointly significant at the 5% level, age and male with a negative sign on the coefficient, age with a positive; our proxy for income appeared insignificant, which is reasonable given our population sample). It was natural to consider these two sets of independent variables - one including ra_score only, and the other including male, age and income - as candidates for minimal models for our analysis. However, in $P(translog = 1|x) = \beta_0 + \beta_1 \cdot ra\_score$ ra_score never appeared significant and only rarely we observed the joint significance of the estimated models. Therefore, we omit a discussion of these results.

The results from the estimation of Model 2 are also summarized in Table 3.5.

	periods 1,3,5		**periods 2,4,6**	
Model	1	2	1	2
age	**-.0287** (.302)	**.1280** (.007)	**.0220** (.381)	**.0913** (.000)
male	**.0686** (.646)	**3.3442** (.010)	**.1706** (.055)	**2.5462** (.014)
econ	**-.1601** (.212)	**-.6307** (.000)	**-.0731** (.503)	**.2210** (.001)
Ctreat	**.0559** (.657)	**-.7156** (.004)	**.0230** (.809)	**-.0375** (.644)
age$*$ male	-	**-.1852** (.002)	-	**-.0941** (.032)
econ$*$ male	-	**.5354** (.002)	-	**-.3395** (.019)
Ctreat$*$ male	-	**.7983** (.006)	-	**.0036** (.983)
const	**1.2342** (.068)	**-1.4593** (.139)	**-.3400** (.553)	**-2.1070** (.000)
mean $\widehat{p}$(y=1)	**.5625**	**.5625**	**.2361**	**.2361**
# of obs.	144	144	144	144
joint p-value	(.488)	(.000)	(.078)	(.000)

Table 3.5: Results from estimation of the linear probability model(s). The first row of each cell reports estimated coefficients. The second row reports the corresponding p-value. *Mean* $\widehat{p}$*(y=1)* denotes the mean predicted probability of a transfer being made.

Model 2 is strongly significant and this confirms a strong gender effect. Therefore, in the discussion that follows, we will concentrate on the results from Model 2.

For both low- and high-endowment-period data, the joint p-value of the model is .000. All demographic characteristics – *age*, *male*, and *econ* – and their interaction terms are significant at the 5% level. Interestingly, the treatment dummy together with its interaction term is only significant for the low-endowment periods. This suggests that only for the low-endowment periods the presentation of the game matters.

The mean predicted probability of transfer in the low-endowment periods is .56; in the high-endowment periods it is only .24, which is considerably lower. This result contradicts the theoretical prediction.[39]

For the low-endowment periods, *age* has a positive sign on the coefficient for female, but a negative sign for male. *Econ* has a negative sign on the

[39]Recall that in the equilibrium Participant X always transfers high endowment and never transfers low.

coefficient for both male and female. The intercept is negative for women and positive for men. This means that with the same characteristics, women are less likely to make the transfer than men.

Treatment dummy *Ctreat* has a negative sign for female but positive for male subjects. This suggests a negative impact of a corruption context on the transferring decision for women but a positive impact for men, which is an intriguing result.

For the high-endowment periods, both *age* and *econ* have a positive sign on the coefficient for female, but negative for male. Similarly to low-endowment periods, the intercept is negative for women and positive for men. Thus, also when the endowment is high, having the same characteristics, women are less likely to make the transfer than men.

The treatment dummy *Ctreat* has a negative sign for both female and male subjects, which is yet another difference from low-endowment periods. This suggests a negative impact of a corruption context on the transferring decision – in high-endowment periods subjects are less likely to transfer when they are fully aware of the context. Note, however, that even though the sign reflects the expected impact of context, the coefficient is not significant.

3.5 Discussion

Some of the results confirm our expectations whereas some do not.

In the aggregate data, we find only a small and statistically insignificant treatment effect, which is in line with Abbink and Hennig-Schmidt (2006) but not with our expectations. Once we look at the male and female subsamples separately, we discover (significant) gender effects that cancel each other out and are responsible for the reduced overall effect of non-neutral framing.

For the aggregate second-stage data, the treatment effect shows in an increased denouncing rate, which is in line with our expectations. For male and female subsamples, as much as we can tell given the low number of observations, denouncing rates are lower or the same[40] in the B treatment. Also for the aggregate third-stage data the treatment effect goes in the predicted direction.

Different attitudes of men and of women towards corruption have previously been reported by, for example, Alatas et al. (2006). These authors

[40]This refers to all possible cases, when we are looking separately at male and female subsamples for high- and for low-endowment periods.

find significant differences in the behavior of men and women in a corruption experiment. Their results, however, appear to be culture-specific.[41]

The observed negative impact of non-neutral framing on the transferring decisions of women, together with the positive impact on the denouncing decisions of women, are in line with earlier findings of women being less likely to engage in, as well as less tolerant of (thus more likely to act against), corruption than men (e.g., Swamy, Knack, Lee and Azfar, 2001, or Dollar, Fisman and Gatti, 2001).

For low-endowment periods, we find a positive impact of a bribery context on the transfer rates of men. This suggests the opposite treatment effect to what we expected, but only for the male subsample. Women react to the context by reduced transferring. The (significant) result for the male subsample is very surprising and difficult to understand.

For high-endowment periods, the treatment effect appears insignificant. We find a (slightly) reduced transfer rate for male and a (slightly) increased transfer rate for the female subsample. The result for the female subsample is counterintuitive, however, the results of the t-test suggest that it might be due to random realization.

Another interesting result is that for both low- and high-endowment periods more than 50% of subjects do not play the equilibrium. Recall that theoretically, the optimal strategies are to transfer when the endowment is high and not to transfer when the endowment is low. For both treatments we observe just the opposite – relatively high transfer rates for low-endowment periods and relatively low transfer rates for high-endowment periods.

This phenomenon, in addition, appears robust. In Krajčová (2008) the same result is found for the high-incentive treatment.[42]

There are several possible explanations. One of them is the "preference for inclusion" conjectured by Cooper and Van Huyck (2003). They find that subjects are significantly more likely to make "non-rational" choices that allow their co-player to make a choice – and thereby to affect the final payoffs – when given extensive form versions of a game. In our game, "inclusion" introduces the risk of a significant loss. Our results suggest that subjects with a "preference for inclusion" were willing to transfer and to continue playing the game as long as the stakes were not too high (low endowment).

[41]The authors run the experiment in Melbourne (Australia), Delhi (India), Jakarta (Indonesia), and Singapore. Only the Australian data confirm a significant gender effect.

[42]Recall that the benchmark treatment is the same for both Chapters 2 and 3.

We also note that the theoretical prediction is computed under the assumption of risk neutrality, which, as also suggested by the data from the risk-aversion questionnaire, is not likely to hold in our sample. When we computed the theoretical prediction for a (modestly) risk averse subject, we found that under some (reasonable) assumptions, our chosen parameterization can lead to a no-corruption equilibrium also for the high-endowment periods.[43] That is, for risk-averse subjects, it might be in fact optimal not to transfer a high endowment.

Finally, the nature of the game implies that the endowment has explicit payoff consequences in the second and the third stage of the game only for Participant Y, not for Participant X. Therefore, the strategic importance of the endowment level might have been less obvious to Participants X than we thought.

Our results suggest that context indeed plays an important role for a subject's behavior in a bribery game. Importantly, the effect on male participants might be different than the effect on female participants. Some of our results are not significant, but this might be caused by the relatively small sample and the gender-unbalanced subject pool. With more subjects, possibly observed over more periods, and with a better gender-balanced sample, our results might become more conclusive.[44] Altogether, we conclude that subjects seem to engage in all sorts of social considerations in a bribery game, including moral scruples, which should not be dismissed by experimenters looking for relevant policy implications.

---

[43] We assume the standard CRRA utility function $u(x) = x^{(1-r)}$. The average risk-aversion coefficient in our sample is about 0.03, and the maximal is about 0.1. As the bribery game involves nodes with negative payoffs, some assumptions need to be made about the utility function in the negative domain. The prospect theory suggests that in the negative domain, the steepness of the utility function might be about twice as much as in the positive domain. For illustration, we computed the theoretical prediction for a risk-neutral subject assuming two different levels of (dis)utility from paying a 300 CZK penalty after detection: $u(-300) = -u(450)$ and $u(-300) = -u(600)$. For low endowment, the theoretical prediction is the same as for a risk-neutral subject. For high endowment it changes. For an extremely risk-averse participant ($r = 0.1$), the disutility of 450 still predicts a corruption equilibrium, however, the disutility of 600 predicts a no-corruption equilibrium. For an average risk-aversion coefficient ($r = 0.03$), the disutility of 450 is sufficient to change the theoretical prediction.

[44] Ortmann and Tichy (1999) also report some evidence of differences in the (cooperative) behavior of men and women. Also the gender composition of the subject pool in the experimental session matters. When controlling for past experience, gender differences, however, disappear.

References

Abbink, K., Hennig-Schmidt, H., (2006). Neutral versus Loaded Instructions in a Bribery Experiment, *Experimental Economics 9(2)*, 103-121.

Abbink, K., Irlenbusch, B., Renner, E., (2002). An Experimental Bribery Game, *Journal of Law, Economics, and Organization 18(2)*, 428-454.

Alatas, V., Cameron, L., Chaudhuri, A., Erkal, N., Gangadharan, L., (2006). Gender and Corruption: Insights from an Experimental Analysis. *The University of Melbourne Research Paper No. 974.*

Apesteguia, J., Dufwenberg, M., Selten, R., (2007). Blowing the Whistle. *Economic Theory 31*, 143–166.

Bigoni, M., Fridolfsson, S.-O., Le Coq, C., Spagnolo, G., (2008a). Fines, Leniency, Rewards and Organized Crime: Evidence from Antitrust Experiments. *Working Paper Series in Economics and Finance, No. 698, Stockholm School of Economics.*

Bigoni, M., Fridolfsson, S.-O., Le Coq, C., Spagnolo, G., (2008b). Risk Aversion, Prospect Theory, and Strategic Risk in Law Enforcement: Evidence From an Antitrust Experiment. *Working Paper Series in Economics and Finance, No. 696, Stockholm School of Economics.*

Buccirossi, P., Spagnolo, G., (2006). Leniency Policies and Illegal Transactions, *Journal of Public Economics 90*, 1281-1297.

Cohen, J., (1988). Statistical Power for the Behavioral Sciences, 2^{nd} edition. *Lawrence Erlbaum Associates Inc, Hillsdale.*

Cooper, D., J., Van Huyck, J.,B., (2003). Evidence on the Equivalence of the Strategic and Extensive Form Representation of Games, *Journal of Economic Theory 110*, 290-308.

Dollar, D., Fisman, R., Gatti, R., (2001). Are Women Really the “Fairer” Sex? Corruption and Women in Government. *Journal of Economic Behavior and Organization 46(4)*, 423-29.

Dušek, L., Ortmann, A., Lízal, L., (2005). Understanding Corruption and Corruptibility through Experiments. *Prague Economic Papers 14*, 147-162.

Fischbacher, U., (2007). Z-tree: Zurich Toolbox for Ready-made Economic Experiments - Experimenter's Manual. *Experimental Economics 10(2)*, 171-178(8).

Gigerenzer, G., Hoffrage, U., (1995). How to Improve Bayesian Reasoning without Instruction: Frequency Formats. *Psychological Review 102*, 684–704.

Goeree, J., K., Holt, C., A., (2001). Ten Little Treasures of Game Theory, and Ten Intuitive Contradictions. *American Economic Review 91*, 1402-1422.

Gupta, S., Davoodi, H., Alonso-Terme, R., (2002). Does Corruption Affect Income Inequality and Poverty? *Economics of Governance 3(1)*, 23-45.

Hertwig, R., Ortmann, A., (2001). Experimental Practices in Economics: A Methodological Challenge for Psychologists? *Behavioral and Brain Sciences 24(3)*, 383-451.

Hertwig, R., Ortmann, A., (2004). The Cognitive Illusions Controversy: A Methodological Debate in Disguise that Matters to Economists. *in Zwick, R., Rapoport, A. (eds.), Experimental Business Research, Kluwer Academic Publishers, Boston, MA*, 361- 378.

Holt, C., A., Laury, S., K., (2002). Risk Aversion and Incentive Effects, *American Economic Review 92(5)*, 1644-1655.

Hwang, J., (2002). A Note on the Relationship Between Corruption and Government Revenue. *Journal of Economic Development 27(2)*, 161-178.

Kamecke, U., (1997). Rotations: Matching Schemes that Efficiently Preserve the Best Reply Structure of a One Shot Game, *International Journal of Game Theory 26(3)*, 409-417.

Krajčová, J., (2008). Testing Leniency Programs Experimentally: The Impact of Change in Parameterization, *CERGE-EI Working Paper No. 370.*

Mauro, P., (1995). Corruption and Growth. *Quarterly Journal of Economics 110*, 681-712.

Ortmann, A., Lízal, L., (2003). Designing and Testing Incentive-compatible and Effective Anti-corruption Measures, grant proposal accepted by the Grant Agency of the Czech Republic. Grant No. 402/04/0167.

Ortmann, A., Tichy, L., K., (1999). Gender Differences in the Laboratory: Evidence from Prisoner's Dilemma Games, *Journal of Economic Behavior & Organization 39*, 327-339.

Richmanová, J., (2006). In Search of Microeconomic Models of Anti-Corruption Measures – A Review, *CERGE-EI Discussion Paper No. 2006-157.*

Richmanová, J., Ortmann, A., (2008). A Generalization of the Buccirossi and Spagnolo (2006) Model, *CERGE-EI Discussion Paper No. 2008-194*.

Roth, A., (2002). The Economist as Engineer: Game Theory, Experimentation, and Computation as Tools for Design Economics, *Econometrica 70(4)*, 1341-1378.

Rydval, O., (2007). The Impact of Financial Incentives on Task Performance: The Role of Cognitive Abilities and Intrinsic Motivation, *PhD dissertation, CERGE-EI.*

Spagnolo, G., (2004). Divide et Impera: Optimal Leniency Programs, *C.E.P.R. Discussion Paper No. 4840.*

Swamy, A., Knack, S., Lee, Y., Azfar, O., (2001). Gender and Corruption. *Journal of Development Economics 64(1)*, 25-55.

Tanzi, V., (1998). Corruption around the World: Causes, Consequences, Scope and Cures. *IMF Working Paper 98/63.*

Williamson, O., (1983). Credible Commitments: Using Hostages to Support Exchange. *American Economic Review 73*, 519-540.

APPENDIX

Comparing Data from Periods Before and After the Switching of Roles.

In Figure 3.7, we present the data from the before- and after-the-switch-of-roles periods for low and high endowments of the B treatment.

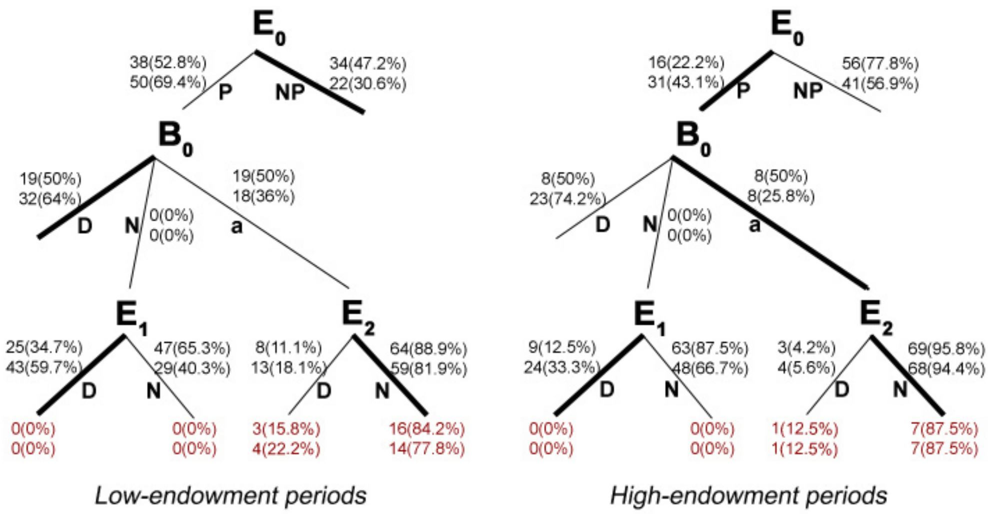

Figure 3.7: Before- vs. after-the-switch-of-roles data in the B treatment. Before-switch data are in the upper rows and after-switch data are below.

In both cases, we observe a somewhat higher transfer rate in the second six periods. Similarly as in the first part of the experiment, the transfer rate is higher in periods when the endowment is low than when it is high. In the B_0 node, more subjects chose the safe option (with no possibility of loss) after the switch of roles. This means for low-endowment periods a shift towards, but for high-endowment periods a shift further away from, the theoretical prediction. In the E_2 node, results from before- and after-switch data are very similar and for both low and high endowment, and they are in line with the theoretical prediction. In the E_1 node, we observe a shift towards the equilibrium after the switch of roles.

In Figure 3.8, we present the data from before- and after-the-switch-of-roles periods from the low- and high-endowment periods of the C treatment.

In the C treatment, the transfer rate drops after the switch of roles, more so in periods when the endowment is high. This is just the opposite effect as in the B treatment. The transfer rate is higher when the endowment is low in both cases, before and after the switch of roles, which contradicts the theoretical prediction. In the B_0 node, a higher fraction of subjects chose the safe option (with no possibility of loss) after the switch of roles. This is a similar result as in the B treatment – for low-endowment periods it means a shift towards, but for high-endowment periods a shift further away from, the theoretical prediction. In the E_1 and E_2 nodes, the results from before- and after-switch data are similar for low-endowment periods (more so in the E_1 than in the E_2 node). In high endowment periods we observe no difference at all.

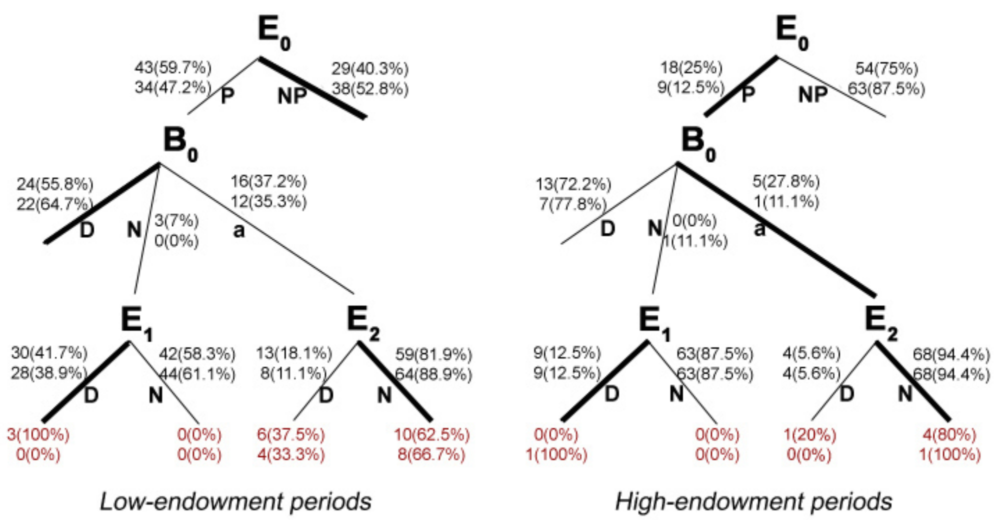

Figure 3.8: Before- vs. after-the-switch-of-roles data in the C treatment. Before-switch data are in the upper rows and after-switch data are below.

Robustness Checks

We performed two types of robustness check of our estimation results. The first regards the way we treated individual observations over rounds when running regressions – this is discussed in the subsection *Handling of the Data*. The second regards the experimental design – we run several sessions of alternative treatments in which we introduce only minor changes that do not appear to significantly affect behavior of subjects – this is discussed in the subsection *Pooling the Sessions*.

A. Handling of the Data

Throughout the analysis we have defined three alternative dependent variables, each of which captures slightly different information about the first-stage data. *Translog* is a 0/1 dummy variable identifying transfer being made (value of 1) or not (value of 0) in a particular period. *Atranslog* is the average value of *translog* for one individual over periods $1, 3$, and 5 (low-endowment periods) or $2, 4$, and 6 (high-endowment periods). *Ltranslog* defines the dominant choice of a subject in periods $1, 3$, and 5 or $2, 4$, and 6. For a subject who has chosen *Pay* two or three times out of a total three periods of interest, the dominant choice is 1; for a subject who has chosen *Not Pay* two or three times out of total three periods of interest, the dominant choice is 0.

Then, using one of the three types of dependent variable, we conducted four different types of regression analysis.

Clustered Regressions – as discussed in the main text, we run a clustered (robust)[45] linear probability model estimation with the binary variable *translog* as a dependent variable.

Regressions on Averaged Data – in this case, we run an ordinary least squares estimation of *atranslog*. We analyze only averaged data, where higher values of *atranslog* correspond to more transfers being made and thus to a stronger preference for this choice.[46]

Regressions on the 1^{st} *or* 2^{nd} *Period Data* – we estimate LPM only on the 1^{st} and 2^{nd} period *translog* (for low- and high-endowment periods, respectively). In this approach we are omitting part of the information, however we only use the part of the data that is not affected by the experience from previous rounds.[47]

Regressions on Dominant Choice – we estimate LPM using *ltranslog* as a dependent variable. Thus in this case, we are only looking at the dominant choice of each subject.

First we look at effect size measures, whether they give robust results for all four approaches to the data. The results are summarized in Table 3.6.

In all cases, the effects are small (recall that Cohen 1998 defines effect sizes of $d = 0.2$ as *small*), for high-endowment dominant choice data the

[45] Standard errors are corrected for heteroskedasticity and for within-subject correlation.

[46] We also run poisson regressions on a count variable (counting the number of transfers made by an individual in the relevant three periods). The qualitative results are the same as with OLS and *atranslog*.

[47] We realize that for 2^{nd} period data this may not be completely true if subjects fail to realize that it is a different game they are playing in the high-endowment periods.

		B		C		effect size	
	Data	mean	std.dev.	mean	std.dev.	odds ratio	Cohen's d
1,3,5	1^{st} period	.583	.5036	.625	.4945	1.072	.0841
	average	.528	.4495	.597	.4282	1.131	.1571
	dominant	.5	.5108	.583	.5036	1.166	.1635
	all periods	.528	.5027	.597	.4939	1.131	.1385
2,4,6	2^{nd} period	.292	.4643	.25	.4423	0.856	-.0926
	average	.222	.3764	.25	.3147	1.126	.0807
	dominant	.25	.4423	.25	.4423	1	0
	all periods	.222	.4187	.25	.4361	1.126	.0655

Table 3.6: Effect-size indices.

effect is zero (but we need to keep in mind that only part of the available information is used). Except for 2^{nd} period data, also the direction of effect is the same in all cases. This suggests that initially, the transfer rate was lower for high-endowment periods in the context treatment but in later periods it increased. When we look at the male and female subsamples, the results are also robust for all four approaches – suggesting a counteracting gender effect (we omit reporting all numbers here as they are very similar to the results for averaged data reported in Table 3.4 in the main text).

Tables 3.7 and 3.8 summarize the main results from the estimation for low- and high-endowment periods. For all four approaches, the models that do not allow for gender-specific effects are not significant. Therefore in the discussion that follows we will concentrate only on models containing interaction terms.

For the low-endowment periods, the results from the *averaged*, 1^{st} *period* and *dominant-choice* data analysis confirm the results from the *clustered* regressions. We find the directions of all the effects the same, the explanatory variables are significant in most cases and there are no dramatic differences in coefficient sizes. Only *econ* and *econ*male* are not significant in the 1^{st} *period* data case. They both become significant once we include the information from later rounds – for *clustered*, *averaged* and *dominant-choice* data.

For the high-endowment periods, only the results from *averaged* and *dominant-choice* data analysis confirm the results from *clustered* regressions – the treatment dummy is not significant, neither is its interaction term, the directions of all the effects are the same, and the sizes of the coefficients are comparable. For the 2^{nd} *period* data the estimated model is not significant. This suggests that the behavior in the second period is different, and more difficult to explain by demographic characteristics. To be able to say whether in later rounds the behavior really stabilizes, we would need to observe more high-endowment periods.

	Periods 1,3,5							
	clustered		averaged		1^{st} period		dominant	
age	**-.0287** (.302)	**.1280** (.007)	**-.0287** (.317)	**.1280** (.011)	**-.0069** (.822)	**.1255** (.034)	**-.0470** (.160)	**.1141** (.058)
male	**.0686** (.646)	**3.3442** (.010)	**.0686** (.656)	**3.3442** (.015)	**.2748** (.117)	**2.8556** (.067)	**.0758** (.664)	**3.4441** (.042)
econ	**-.1601** (.212)	**-.6307** (.000)	**-.1601** (.226)	**-.6307** (.000)	**-.1269** (.416)	**-.3627** (.313)	**-.1529** (.342)	**-.8570** (.000)
Ctreat	**.0559** (.657)	**-.7156** (.004)	**.0559** (.666)	**-.7156** (.006)	**.0381** (.794)	**-.7183** (.007)	**.0733** (.621)	**-.6802** (.010)
age*male	-	**-.1852** (.002)	-	**-.1852** (.003)	-	**-.1446** (.035)	-	**-.1983** (.009)
econ*male	-	**.5354** (.002)	-	**.5354** (.004)	-	**.2467** (.534)	-	**.8200** (.000)
Ctreat*male	-	**.7983** (.006)	-	**.7983** (.010)	-	**.8433** (.009)	-	**.7390** (.025)
const	**1.2342** (.068)	**-1.4593** (.139)	**1.2342** (.077)	**-1.4593** (.162)	**.6265** (.382)	**-1.7304** (.194)	**1.6046** (.047)	**-1.0004** (.444)
mean $\widehat{p}$(y=1)	**.5625**	**.5625**	**.5625**	**.5625**	**.6042**	**.6042**	**.5417**	**.5417**
# of obs.	144	144	48	48	48	48	48	48
joint p-value	.488	.000	.519	.000	.370	.001	.370	.000

Table 3.7: Results from *clustered* regressions vs. regressions on *averaged*, 1^{st} *period*, and *dominant-choice* data from low-endowment periods.

B. Pooling the Sessions

In addition to the benchmark treatment B we conducted two plus two sessions of "automatic" treatments A and AI. Under both treatments, A and AI, we used the same game and same parameterization as in the B treatment. The only difference was that in automatic treatments, each subject played against a computer program, with six subjects in the role of Participant X and six subjects in the role of Participant Y. The computer program always played a (subgame perfect) optimal strategy. Subjects were acquainted with these facts in the instructions.

The only difference between A and AI treatments was that in AI subjects received, as a separate part of the instructions, a so-called Backwards Induction Tutorial, intended to explain the basic principles of using backwards induction.

In addition to the full-context C treatment, we conducted two sessions with partial context – the C- treatment. In the C- treatment, the subjects receive only limited information about the context – Participant X is called "Entrepreneur" and Participant Y is called "Bureaucrat". Actions are, however, denoted by neutral letters – the same as in the B treatment.

Before pooling the data from different treatments we performed basic

	Periods 2,4,6							
	clustered		averaged		2^{nd} period		dominant	
age	**.0220** (.381)	**.0913** (.000)	**.0220** (.396)	**.0913** (.000)	**.0052** (.873)	**.0456** (.397)	**.0486** (.108)	**.1369** (.000)
male	**.1706** (.055)	**2.5462** (.014)	**.1706** (.063)	**2.5462** (.021)	**.2312** (.051)	**2.0233** (.241)	**.1133** (.327)	**3.2569** (.010)
econ	**-.0731** (.503)	**.2210** (.001)	**-.0731** (.516)	**.2210** (.002)	**-.1495** (.334)	**.1772** (.274)	**-.1039** (.483)	**.3316** (.002)
Ctreat	**.0230** (.809)	**-.0375** (.644)	**.0230** (.815)	**-.0375** (.663)	**-.0514** (.699)	**.0480** (.627)	**-.0143** (.908)	**-.0563** (.663)
age*male	-	**-.0941** (.032)	-	**-.0941** (.043)	-	**-.0625** (.397)	-	**-.1215** (.023)
econ*male	-	**-.3395** (.019)	-	**-.3395** (.026)	-	**-.3710** (.133)	-	**-.5021** (.016)
Ctreat*male	-	**.0036** (.983)	-	**.0036** (.984)	-	**-.1750** (.424)	-	**-.0395** (.863)
const	**-.3400** (.553)	**-2.1070** (.000)	**-.3400** (.565)	**-2.1070** (.000)	**.1145** (.875)	**-1.1202** (.370)	**-.8378** (.214)	**-3.1605** (.000)
mean $\widehat{p}$(y=1)	**.2361**	**.2361**	**.2361**	**.2361**	**.2708**	**.2708**	**.25**	**.25**
# of obs.	144	144	48	48	48	48	48	48
joint p-value	.078	.000	.095	.000	.175	.027	.183	.000

Table 3.8: Results from *clustered* regressions vs. regressions on *averaged*, 1^{st} *period*, and *dominant-choice* data from high-endowment periods.

statistical tests in order to discover significant differences in the distributions of choices – Fisher's exact test and the Wilcoxon rank-sum test. We find no evidence of significant differences in the distributions of choices between A, AI and B treatments, nor between C- and C treatments.

Afterwards, we performed two types of pooled analysis: 1) pooling the data from A and B treatments vs. pooling the data from C- and C treatments; and 2) pooling the data from A, AI and B treatments vs. pooling the data from C- and C treatments. Note that in 1) both pools contain the same number of subjects, which is not the case after we extend the benchmark-type pool by data from AI.

See Tables 3.9 and 3.10 for the regression results for low- and high-endowment periods, respectively. Clearly, pooling slightly different treatments leads to noisier results, which is not very surprising.

For both low- and high-endowment periods, the significance of *econ* (and its interaction term) disappears.

As regards the treatment dummy, on the one hand, the significance for low-endowment periods disappears, but on the other hand, the treatment dummy becomes significant for high-endowment-period data.

	Periods 1,3,5					
	B vs. C		B,A vs. C,C-		B,A,AI vs. C,C-	
age	**-.0287** (.302)	**.1280** (.007)	**-.0191** (.380)	**.0854** (.073)	**-.0093** (.641)	**.0784** (.040)
male	**.0686** (.646)	**3.3442** (.010)	**.0162** (.890)	**2.4523** (.039)	**.0076** (.940)	**2.2189** (.030)
econ	**-.1601** (.212)	**-.6307** (.000)	**-.1754** (.089)	**-.3061** (.025)	**-.1343** (.164)	**-.0944** (.532)
CCtreat	**.0559** (.657)	**-.7156** (.004)	**.0609** (.550)	**-.2708** (.218)	**.0736** (.449)	**-.2738** (.178)
age*male	-	**-.1852** (.002)	-	**-.1235** .(023)	-	**-.1080** (.018)
econ*male	-	**.5354** (.002)	-	**.1801** (.318)	-	**.0005** (.998)
CCtreat*male	-	**.7983** (.006)	-	**.3586** (.154)	-	**.3877** (.101)
const	**1.2342** (.068)	**-1.4593** (.139)	**1.0848** (.035)	**-.9740** (.339)	**.8308** (.071)	**-.9774** (.243)
mean p(y=1)	**.5625**	**.5625**	**.5787**	**.5787**	**.5714**	**.5714**
# of obs.	144	144	216	216	252	252
joint p-value	.488	.000	.439	.066	.675	.2194

Table 3.9: Results from estimation on basic vs. extended data sets for low-endowment periods. CCtreat is a dummy identifying context-type treatment(s) – C, or C and C- treatments.

	Periods 2,4,6					
	B vs. C		B,A vs. C,C-		B,A,AI vs. C,C-	
age	**.0220** (.381)	**.0913** (.000)	**.0310** (.100)	**.1134** (.000)	**.0253** (.160)	**.0758** (.007)
male	**.1706** (.055)	**2.5462** (.014)	**.0620** (.461)	**2.3301** (.001)	**.0144** (.867)	**1.4000** (.081)
econ	**-.0731** (.503)	**.2210** (.001)	**-.1751** (.070)	**-.1029** (.616)	**-.1424** (.113)	**.0227** (.876)
CCtreat	**.0230** (.809)	**-.0375** (.644)	**-.0780** (.331)	**-.2268** (.072)	**-.1172** (.133)	**-.2955** (.026)
age*male	-	**-.0941** (.032)	-	**-.1019** (.001)	-	**-.0601** (.099)
econ*male	-	**-.3395** (.019)	-	**-.0768** (.740)	-	**-.2101** (.248)
CCtreat*male	-	**.0036** (.983)	-	**.1282** (.436)	-	**.1958** (.243)
const	**-.3900** (.553)	**-2.1070** (.000)	**-.3103** (.469)	**-2.1335** (.000)	**-.1354** (.737)	**-1.2879** (.029)
mean p(y=1)	**.2361**	**.2361**	**.2593**	**.2593**	**.2817**	**.2817**
# of obs.	144	144	216	216	252	252
joint p-value	.078	.000	.045	.000	.075	.040

Table 3.10: Results from estimation on basic vs. extended data sets for high-endowment periods. CCtreat is a dummy identifying context-type treatment(s) – C, or C and C- treatments.