Understanding Corruption and Corruptibility Through Experiments: A Primer

Libor Dusek, Andreas Ortmann, Lubomir Lizal[1]

December 20, 2004

Abstract

Corruption and corruptibility - due to their illegal and therefore secretive nature – are difficult to assess either with traditional tools such as hard data on criminal convictions or soft data elicited through opinion polls, questionnaires, or case(s) studies. While there seems to be agreement nowadays that corruption does have a negative impact on (foreign) private investment and growth, government revenue and infrastructure, and social equality, and while there seems to be evidence that low economic development, federal structure and short histories of experience with democracy and free trade all favor corruption on the macro-level, it is poorly understood what exactly, on the micro-level, the determinants of corruptibility are and what institutional arrangements could be used to fight (the causes of) corruption. In this article we review a third, complementary mode of investigation of corruption and corruptibility: experiments. We assess their strengths and weaknesses, and identify areas where they could be particularly useful in guiding policy choices – namely in designing incentive-compatible and effective anti-corruption measures in public procurement.

1. Introduction

Corruption remains an important policy concern in virtually all countries, the Czech Republic being no exception. Due to its secretive nature, the extent and pervasiveness of corruption has been difficult to assess although examples of creditable assessment tools such as the well-established Corruption Perception Index of Transparency International (www.transparency.org; Treisman 2000), or the new V4 City Corruption Propensity Index of Transparency International CR

(www.transparency.cz; Ortmann 2004), provide approximations that suggest that the available hard data (e.g., criminal convictions) are but the tip of the iceberg. There seems to be agreement now that corruption does affect negatively (foreign) private investment and growth (e.g., Mauro 1995, 1997), government revenue (Hwang 2002) and infrastructure (Mauro 1998), as well as social equality (Gupta, Davoodi, & Alonso-Terme 2002). In the following we therefore take as a well-proven fact that corruption is welfare-reducing. There also seems to be evidence that low economic development, federal structure and short histories of experience with democracy and free trade all favor corruption on the macro-level (Treisman 2000). However, it is poorly understood what exactly, on the micro-level, the determinants of corruptibility are and what institutional arrangements could be used to fight (the causes of) corruption.

How important, for example, are detection probabilities for bribe-giving and bribe-taking and how do they interact with the severity of penalties? Are detection probabilities correctly perceived? Can the perception of detection probabilities systematically be manipulated (e.g., by going after high-visibility violators rather than routine violators)? Is corruptibility also a function of people's perception of the pervasiveness of corruption in society? Is the distant past of a country indeed as important as current policy, as Treisman (2000) claims? What do laws and regulations have to look like if they are to stand a chance to effectively undermine "the tenacity of the past" (Treisman 2000, p. 438)?

Because of the secrecy in which corruption typically takes place, these are tough questions to answer. It is therefore, maybe, not surprising that one typically finds a trial-and-error approach to laws and regulations that try to curb corruption. This approach is manifested in frequent legal and regulatory revisions through which authorities try to react to deficiencies of laws and regulations that have become too obvious to ignore. Such an evolutionary approach to finding optimal solutions works sometimes but often it does not as, for example, the frequent revisions of public procurement law in the Czech Republic demonstrate.

Laboratory experiments have been used increasingly, and successfully, as the method of choice by economists to understand a plethora of design and implementation problems ranging from the analysis of matching mechanisms in a variety of labor markets (Roth 2002) to auction mechanisms (Milgrom 2002, Klemperer 2004). The increased and successful use of experiments has three drivers. First, laboratory experiments allow us to control the behavior of subjects in ways that are typically not possible in the field. Second, laboratory experiments allow one to systematically manipulate the environment and the resulting behavior changes and hence to address the issue of causality in ways not possible in field contexts. Third, it is often less expensive to test alternative institutional arrangements (e.g., subtle differences in auction procedures for public procurement projects) in the experimental laboratory rather than in the laboratory of real life.

For these reasons, laboratory experiments on corruption, corruptibility, and measures to fight them, seem self-suggesting. It is therefore interesting to note that experiments on corruption and corruptibility are few and far between, with the earliest arguable being Frank & Schulze (2000) and Abbink, Irlenbusch, & Renner (2002). Experiments that address the incentive-compatibility and effectiveness of anti-

corruption measures in public procurement are even fewer in number, with Apesteguia, Dufwenberg, & Selten (2003) being the prominent example to date.

Indeed, in this article we review what is, to the best of our knowledge, the universe of laboratory experiments that speak to the issue of corruption and corruptibility as of December 2004[2]. We assess the strengths and weaknesses of this method of investigation, and identify areas where laboratory experiments could be particularly useful (e.g., the design of incentive-compatible and effective anti-corruption measures). The remainder of the article is organized as follows:

In section 2, we review experiments on the determinants of corruptibility and corruption and an experiment on the efficacy of whistle-blower provisions, and summarize what can be learned from this initial set of experiments. In section 3 we address important methodological issues such as representative sampling and representative stimuli. In section 4 we ponder the question where exactly experiments could be of use.

2.   A review of experiments

2.a.   Experiments on the determinants of corruption – bilateral settings

Game theoretically speaking, corruption is a three-player game involving a briber (the principal), a bribee (the agent, typically assumed to be some public official), and a third party (possibly, society), that is damaged by the bribe.[3] Of course, corruption also exists in private interactions: the bribee may be some manager and the third party damaged by the bribe may be various other stakeholders. Some of the players may be collective actors that either collaborate or compete. Stripped to its essence, corruption is a principal-agent (P-A) game with an important twist: the externality or welfare reduction imposed on the third party. This twist complicates the analysis of the basic principal-agent game which has been extensively studied and is reasonably well understood analytically, albeit typically in the context of legal activities such as employer-employee relationships.[4] It is also at the heart of a recent flurry of publications on corruption and competition.[5]

The P-A game has also been studied experimentally, mostly in the context of gift-exchange games (which are P-A games under a different name). Prominent papers are Fehr, Kirchsteiger, & Riedl (1993) and Fehr, Gaechter, & Kirchsteiger (1997). These papers have generated an important industry of their own, analytically and experimentally, on trust and reciprocity.[6]

---

[2] More precisely, we should say: the universe of English-language articles on corruption experiments. After a first draft of this article existed, we came across the excellent survey article by Renner (2004) which, unfortunately, is in German. Renner reports that Gneuss (2002) contains experiments on leniency provisions. We have not yet been able to lay our hands on this book. Frank (2004), also in German, is an excellent summary of key results from the empirical research on corruption; we benefited greatly from it.

[3] The third party (possibly, society), that is damaged by the bribe, is typically a rather passive player, more like the recipient in a dictator "game" than the responder in an ultimatum game.

[4] E.g., Kreps 1990; Mas-Colell et al. 1995; Ortmann & Colander 1997; Martimort 2002).

[5] E.g., Burguet & Che 2004; Celentani & Ganuza 2002; Compte, Lambert-Mogiliansky, & Verdier 2004.

[6] E.g., Fehr & Schmidt 1999; Bolton & Ockenfels 2000; Charness & Rabin 2003; Engelmann & Strobel forthcoming; Berg, Dickhaut, McCabe 1995; Ortmann, Fitzgerald, Boeing 2000; Cox 2004;

Initial experimental results suggested that laboratory subjects were much fairer and more reciprocal than the predictions of deductive game theory suggested[7]. If these results would indeed tell us something about real-world corruption and corruptibility, they would obviously be bad news. Recent literature, however, has asked methodological questions regarding the robustness of the early results.[8] We will address these results in more detail below.

Abbink, Irlenbusch, & Renner [AIR](2000), draw on the gift-exchange literature, and extend it, by introducing the "moonlighting game" which models legally not enforceable types of contracts. A principal "hires" a moonlighter (the agent) to perform some task; he also provides the resources. The moonlighter can either steal the resources or perform the task, thus generating an economic surplus which the principal can either share with the moonlighter, or which he can pocket. In analogy to gift-exchange games, this situation has the potential to improve the lot of both players but efficiency gains of that kind require a (non-binding) agreement to generate an economic surplus, and hence trust (on the part of the principal) and reciprocity (on the part of the agent), with the moonlighter facing (non-rational) retribution if he does not reciprocate the trust. Even though moonlighting is a form of corruption that typically imposes externalities on society, these costs are not modeled here.

AIR implemented their moonlighting scenario experimentally by letting both players start with an identical endowment. They ran two treatments: in one explicit, but non-binding contracting, was possible; in the other (which constitutes a so-called baseline), it was not. Apart from this explicit contracting stage, the two treatments were the same: the Principal was given the possibility to offer up to one half of her initial endowment. The Agent could either keep what was given to her (steal the resources) or pass it back to the Principal (perform the task). Whatever the Agent passed to the Principal was tripled by the experimenter, thus generating an economic surplus for the principal and modeling the efficiency gains from trust. The Principal then could pass up to 100 % of this amount to the Agent, or spent up to half of her initial endowment to punish the Agent for keeping the initial investment. Such punishment reduced the Agent's holdings three times more than Principal spent on punishment. The first two stages are more or less identical to the trust game popularized by Berg, Dickhaut, & McCabe (1995), the retribution stage is AIR's innovation

The prediction of deductive game theory is that the Principal will not spend anything on punishment (if the Agent kept what was given to her), and in any case will not share whatever economic surplus might get generated (if the Agent passed back the initial investment of the Principal to have it tripled by the experimenter). Hence the

---

Dufwenberg & Kirchsteiger 2004; see also the excellent discussion of trust and reciprocity experiments and theories in Camerer 2003.

[7] In the early nineties game theory bifurcated once again. While traditional deductive (or, eductive) game theory continued to be built under the maintained heroic assumptions of full rationality and common knowledge, inductive (or, evolutive) game theory emerged as a serious competitor that made do with significantly reduced rationality and knowledge assumptions. Deductive game theory is best exemplified by Mas-Colell, Whinston. & Green (1995); inductive game theory is well exemplified by Vega-Redondo (2003).

[8] E.g., Engelmann & Ortmann 2004; Charness, Frechette, & Kagel 2004; Healy 2004; Steiner 2004.

Agent will never pass back the initial investment, and hence the Principal will never make an offer in the first place. Or so goes the prediction of traditional game theory for one-shot or finitely repeated games. Not surprisingly, in light of previous results on trust and gift-exchange games, the experimental results showed that this prediction is systematically falsified in both treatments.

In both experimental treatments, retribution was found to be quite common -- the Principal often punished the Agent for stealing the resources, although this action was costly and reduced his payoff as well.[9] Reciprocity was found to be less common – the Principal often did not "return" a fair amount to the Agent. Therefore, summarizing, hostile actions are consistently punished while the friendly ones are less consistently rewarded. The possibility to verbally negotiate contracts increased the probability of the Agent passing (increased trust) but did not change the reciprocity attitude of the Principal (i.e., no decrease in exploitation of the trust of the Agent).

In a follow-up article (AIR 2002), the authors model explicitly a bribery game. In the baseline or "pure reciprocity" treatment, the briber proposes to the bribee a deal. The bribee can decide whether to accept or reject the deal. (Independent of that decision the briber gets stuck with a small initiation fee.) If the bribee rejects the proposed deal, it will not materialize. If the bribee accepts the proposed deal, it brings about (through the experimenter) a tripling of the briber's initial investment. Next the bribee has to choose one of two decisions, with the first decision benefiting the briber significantly more than the bribee, and the second decision benefiting the bribee somewhat more than the briber. No retribution is modeled although the move structure otherwise is that of the moonlighting game. As in the moonlighting game there are also no externalities modeled in this baseline or "pure reciprocity" treatment.

In order to study how externalities and detection probabilities and punishment affect the briber's and bribee's behavior, the experimenters conducted two other treatments. In the "negative externality" treatment, accepting a bribe inflicted damage on a third party, while in the "sudden death" treatment subjects faced a probability of detection if they accepted a bribe. All three treatments were conducted as 30 round "partners" treatments, meaning a subject was matched with one other person throughout.

The game-theoretic prediction is the same in all three treatments.

The results suggest that the behavior of bribers and bribees is unaffected by the damage inflicted on a third party (although here the third party was all the other participants in the experimental session.) The threat of a drastic penalty (although extremely small in the experimental parameterization), decreases attempted bribes. Subjects tend to underestimate the probability of disqualification, though.

Abbink (2002) builds on the bribery game in AIR (2002) but, rather than the third party being represented by all the other participants in the experimental session, the

---

[9] It is an interesting fact that the two treatments are rather similar. The result prompts questions about the baseline set-up (whose results deviate from the game-theoretic prediction.)

third party was now represented by additional subjects that performed a task (evaluating video clips) for which they were paid. Importantly, the wages these workers were given was either high or low relative to that of the public official, making the bribee (the public official) either better off or worse off (if he was not hit by sudden death). Again, there was no treatment effect: whether the bribee received a relatively high or low wage did not affect significantly her or his reciprocity (to the giving of the briber). This observation is important in light of arguments, and empirical evidence (e.g.. Van Rijckeghem & Weder 2001), which suggest that higher wages for public officials would make them more resistant to bribe offers.

Abbink (2004) also builds on the sudden-death treatment of AIR (2002) to study experimentally the corruption-reducing effects of staff rotation, a practice introduced by the German federal government. Staff rotation is implemented by re-matching the participants in the experiment in each round ("strangers" treatment) rather than letting fixed pairs play all thirty interactions ("partners" treatment). The results are in line with our intuition (but arguably contradicts earlier findings on partners/strangers treatments) in that the number of offered transfers, i.e. bribery attempts, and their volume is cut by about half in the strangers treatment. [10]

The papers by Abbink and his collaborators use as a basic template the bi-lateral P-A game template that also underlies gift-exchange games and explore variants of that basic game (namely, reciprocity to bribe offers, and the "trusting" behavior on the part of those that offer the bribes will reciprocate, the effects of small detection possibilities, negative externalities, and relative wages, as well as matching patterns that attempt to explore the consequences of repeat encounters). To those that know the experimental gift-exchange literature, the reported results don't come as surprise. Of course, the same methodological questions that apply to that literature have to be asked here, too. In addition, there is the important question of the framing of the experiments. Following a long-standing tradition in the experimental literature in economics, Abbink and his colleagues stripped the instructions from all references to the reality of what they were investigating (but see Abbink & Hennig-Schmidt 2002 about which more below.) Specifically, nowhere was it mentioned that the problem that the subject faced was a moonlighting or a bribery game. Such abstraction from reality has long been considered best practice in experimental economics because it was widely believed that "home-grown priors" could be kept out of the laboratory this way. As we will see presently, there is increasing doubts about the validity of this claim.

2.b.   Experiments on the determinants of corruption – unilateral settings

Another class of experiments models corruption as a unilateral decision by a public official. Unlike in the experiments by Abbink and his collaborators, there is no reciprocal relationship between a bribee (the public official) and a briber who might choose to offer a bribe in order to induce the bribee to make a more favorable decision. Instead, the public official decides unilaterally how much money to "divert"

---

[10] The result is also interesting from a methodological perspective. The partners/strangers treatment has been used previously in public good experiments where the evidence, however, is rather mixed (see Andreoni & Croson 2004). It is an important question for future research to understand why Abbink (2004) finds these very strong strangers effect in (asymmetric) principal-agent games that do not exist in (symmetric) public good experiments.

from public funds, subject to the risk of being discovered and punished. Such experiments are in principle useful to study the determinants of corruption in public procurement, or other misuses of public funds.

Specifically, Frank & Schulze (2000) and Schulze & Frank (2003) conducted two series of experiments with members of a university student film club in Germany. Before watching a movie, each member was placed in the (fictitious) position of the manager of the club and presented with the following situation: The club needed to obtain some service from a private firm, and the manager had to choose a firm that would perform that service. Each subject was presented with offers from several firms, which differed in (a) the price that the film club would have to pay and (b) the side payment (bribe) from the firm to the manager that the manager would keep for himself. The higher the price, the higher was the bribe. For example, an honest manager would receive no bribe and the club would pay only DM20[11]. If, however, the manager took the maximum bribe offered, he would receive DM144 while the club would have to pay DM200. The subject indicated their secret choice on a form, the forms were collected, one of the forms was drawn, and the payoffs were made according to the subject's choice on the form that was drawn. The club would receive DM200 minus the price chosen by the subject whose form was drawn, and that subject would receive the bribe that he opted for.[12] Subjects were paid in private so that others could not observe the identity of the subject who may have harmed the club by his corrupt choice.

Note that these experiments were unusual in that the public entity that would suffer from corruption was real: The experimenters could study corruption "as is" in the real world, while having other important variables under control.

The two key determinants of corruption investigated in these experiments were the officials' wages and the probability of detection. One half of the subjects were given a fixed payoff in addition to the possible bribe. Comparing these two groups allowed testing the hypothesis that higher wages of public officials lead to less corruption by inducing loyalty. One half of the subjects (both with the positive "wage" and zero wage) were also told that if their sheet would be drawn and they accepted bribes, their payoff from the experiment would be nullified with a known probability. Comparing subjects who faced the risk of punishment with those who did not, allowed the authors to test whether "deterrence works" in this laboratory corruption settings. Within the group that faced the risk of punishment, comparing subjects with positive wage and zero wage allowed the authors to test the hypothesis that higher wages of public officials lead to less corruption by increasing the opportunity costs of corruption because the official with higher wages has more to lose if detected (Becker and Stigler (1974)).

Frank and Schulze find that the relationship between officials' wages and corruption is non-trivial: When there was no risk of punishment, giving subjects an additional fixed payoff did not significantly reduce their proclivity to behave corruptly. Thus there

---

[11] About 10 Euros.

[12] The subjects therefore knew that their choice would materialize into payoffs to themselves and to their club only with some probability. One can speculate whether the subjects' choices were the same if they mattered with certainty; however, there is some evidence (e.g., Bolle 1990) that at least in some settings, deterministic and probabilistic payoffs give similar results.

was no evidence of the loyalty effect of wages. However, when punishment was possible, subjects receiving also the fixed payoff did choose lower bribes, which is consistent with the deterrence effect of higher wages. [13]

The risk of punishment produced the expected result on the other end of the distribution: the share of subjects choosing the maximum possible bribe fell from 28.8% to 12.6% when risk of punishment was introduced. A surprising result of this experiment was that the risk of punishment actually increased corruption: 9.4% of the population was honest when there was no risk of punishment while only 0.9% were honest when punishment was possible. Frank and Schulze hypothesized that the introduction of monetary incentives reduced the intrinsic incentives to behave honestly, which has been observed in different experimental contexts (e.g., Gneezy & Rusticchini 2000).

The absence of a loyalty effect of higher wages and the intrinsic motivation effect of the risk of punishment may have only limited external validity. In real-life situations, it is the principal himself who may induce loyalty or intrinsic motivations by offering higher wages or promising not to audit the agents. In the experiment, higher wage and risk of punishment were controlled by the experimenters, who were not connected with the principal (the film club) in a way that would meaningfully induce some loyalty. As regards the loyalty effect of higher wages, this explanation finds some support in Frank (1998). But the argument that subjects do not care for the welfare of experimenters seems to invalidate the intrinsic motivation effect of the risk of punishment.

Azfar & Nelson (2003) studied not only the impact of wages, but also two other possible determinants of corruption: transparency and separation of powers. Similar to Frank and Schulze, the experimental design models situations when an official is in a position to divert some public funds for himself. Limited transparency makes this possible: the size of the public budget is a realization of a random variable and the public is unable to observe the realization. In the experiment, limited transparency was implemented by having a group of eight subjects, one of them being the official. The president rolled a dice which determined how many valuable tiles he received. This information was not revealed to other subjects. The valuable tiles were topped up with some worthless tiles, which together comprised the group's "budget". The president then made a secret choice of how many of the valuable or worthless tiles to keep for himself and how many to pass on the regular members of the group. From the tiles passed on the group, each regular member of the group drew a single tile. Valuable tiles gained by a subject translated into a monetary payoff at the end of the experiment.

The lack of transparency in the experiment resulted from the fact that drawing a worthless tile did, in no way, prove that the president was corrupt: by keeping

---

[13] Another interesting finding from these experiments is that students of economics were more corrupt when there was no risk of punishment – they were choosing the individually rational maximum bribe more often than students from other fields. The experiment also revealed that this effect is due to self-selection of students into economics rather than due to economists being more able to determine the self-interested optimum after mastering the concepts of economics. This result is in line with previous results of a well-known debate in the *Journal of Economic Perspectives* (See also Rubinstein 2004) to which, however, Yezer, Goldfarb, & Poppen (1996), provided contradictory evidence.

valuable tiles for himself, the president would only increase the probability of drawing a worthless tile, but subjects were more likely to draw worthless tiles if the president drew few valuable tiles in the first place or if a higher number of worthless tiles was mixed with the valuable ones. Azfar and Nelson control the degree of transparency by varying the number of worthless tiles. More worthless tiles made it harder for the regular members to infer whether the president behaved corruptly and therefore encouraged more corruption.

The mechanism that constrained the president from keeping all tiles for himself was election. In subsequent rounds of the experiment, the current president competed for reelection with two other members of the group. In a departure from standard protocol in experimental economics, candidates actually gave speeches to induce members to vote for them. To provide members with additional (albeit, again, only indicative) information about the president's corruption, one member also played the role of "attorney general". The attorney had the option to randomly draw up to four tiles that the president kept for himself and show those to other members of the group. (The first two draws were free for the attorney, two other reduced his payoff). Exposing a valuable tile would unambiguously prove corruption, while exposing worthless tiles would not necessarily prove honesty. The experimenters varied the institutional set-up for choosing the attorney general: in some groups, he was elected together with the president, while in other groups he was chosen by the president. This was intended to investigate possible consequences of "separation of powers". Like in Frank & Schulze (2003), the experimenters also investigated the effect of officials' wages on corruption by varying the fixed payoff that the president would receive regardless of the number of tiles kept.

Deductive game theory predicts that the president will keep all valuable tiles for himself, the attorney general will not reveal any tiles held by the president, and the members will vote in any way that minimizes personal effort. In fact, the actual behavior of subjects was different. At the election stage, members were very unlikely to re-elect a president found corrupt. While the overall probability of re-election was 32%, the corrupt president was re-elected only once out of the 14 cases when corruption was exposed. The attorneys were quite active in revealing the tiles held by the president: in 92% of cases they took at least the two free draws, and in 21% of cases they undertook all four draws, even though that reduced their individual payoffs. The elected attorney generals were more active in exposing president's choices than the appointed ones. Also, the appointed attorneys were more passive in their "job" in the sense that out of 12 cases when attorneys did not undertake the two free draws, 10 of them were appointed.

Overall, the presidents chose to give all the valuable tiles to the group 74% of the time. A higher wage significantly reduced corruption; the structure of the experiment does not, however, allow us to assess to what extent this was due to the loyalty or deterrence effect. The deterrence incentive was indisputably present since a president with a higher wage who was found corrupt had more to lose by not being re-elected. The degree of transparency, controlled by the number of worthless tiles, had the predicted effect on corruption, although the effect was barely statistically significant in some estimation procedures. Whether the attorney general was appointed or elected had no effect on the presidents' corruption which is rather

surprising given that the elected attorneys were more active in exposing the presidents' behavior.

In an experiment that was explicitly inspired by Azfar & Nelson (2003), Barr, Lindelow, & Serneels (2003) studied similarly the effects of embezzlement by public servants, controlling for detection probabilities and the severity of punishment. Their experimental set-up was essentially a relabeling of the Azfar and Nelson experimental set-up tailored to the temptations that healthcare workers in developing countries face (in their experiment, Ethiopian nursing students). The results are quite similar to those in Azfar & Nelson (2003), notwithstanding that the framing of this experiment was "natural" in the sense that the laboratory setting was not abstract. One interesting quantitative result, which Van Rijckeghem & Weder (2001) also find, is that the effect of wage increases is relatively small: a 200% increase in wages leads to only a 30% reduction in resource expropriation.

Falk & Fischbacher (2002) study an experimental setting where subjects are offered the possibility to steal from other participants. These authors are, in the spirit of the emerging literature on conditional cooperation, particularly interested in the question to what extent it matters that other players also engage in criminal activities such as tax evasion. There is evidence from the related survey literature on tax evasion that indeed such social "peer" effects matter (e.g., Torgler 2003; see also the clever littering experiment by Cialdini et al. 1991) Clearly, these results are likely to transfer to the question to what extent corruptibility is a function of the extent of corruption that people encounter in their environment.

In their experiment, Falk and Fischbacher allocated subjects randomly and anonymously to groups of four. Each subject was given the opportunity to take away from the other three between 0 and 20 points. The value of a stolen point to the "thief" was either low (1/2) or high (1). In addition to welfare losses, Falk and Fischbacher controlled for conditional and unconditional stealing decisions. The former was implemented by the strategy method (i.e., by asking a subject how much they would steal given that others stole on average such and such). Another interesting feature of this experiment was its "asset legitimacy": participants had to first earn the money with which the game was played (an experimental implementation feature that in other contexts has been shown to have significant influence on outcomes (e.g., Cherry, Frykblom, & List 2002).

The key result of Falk & Fischbacher is that that norm violations are conditional and that they are remarkably strong. But here, too, it is important to note that the instructions were framed "neutrally": the participants in the experiment did not engage in criminal activities; for all they knew, they simply engaged in some optimization problem. Yes, they had to take from others but this "norm violation" was not labeled a crime explicitly. The authors also did not control for inequality aversion. Babicky, Ortmann, Semerak, Soukenikova, Vrazda (2004) use a set-up inspired by Falk & Fischbacher (2002) to tease apart the effects of reciprocity and inequality aversion.

2.c.   Experiments on the design of anti-corruption measures

There are other incentive mechanisms that may potentially be effective in fighting corruption. For example, in most countries it is illegal to accept as well as offer bribes. It has been suggested that decriminalizing the offering of bribes, or granting legal immunity to bribers who voluntarily report their corrupt acts, would reduce corruption since public officials would face the additional risk of being reported and punished. The effect of possible whistle-blowing on illegal behavior has already been analyzed in the context of anti-trust policy, both theoretically (e.g., Berentsen, Bruegger & Loertscher 2003; Spagnolo 2004) and experimentally (prominently, Apesteguia, Dufwenberg & Selten 2004).

The authorities wish to promote competition and discourage cartel deals. As cartel agreements are illegal, the members of a cartel have to rely on trust (and reciprocity) rather than on written agreements that might incriminate them. In the past, policymakers and regulators punished all firms violating the competition policy rules. Recently, they have become concerned that equal punishment of all firms could be counterproductive and they have tried to strengthen the incentives for members of the cartel to report agreements. These "leniency" provisions essentially guarantee immunity to whistleblowers even if they were involved in the cartel agreement, or the attempt to bribe. Unfortunately, these provisions are a two-edged sword in that they can be used as disciplining tool: A firm, for example, that has accepted a bribe but does not want to deliver on the implicit deal, can now – under certain conditions – be punished for not having paid. Whether indeed leniency provisions have these perverse incentive properties, and how they could be broken down is indeed difficult to assess in the field since undetected (and also unreported) cartels cannot be observed in the real world. The experimental approach seems a highly viable tool, and arguably the only one now available, to analyze the effects of new policies, as illustrated by Apesteguia, Dufwenberg, & Selten [ADS] (2004).

These authors compare experimentally three possible anti-trust policies with the ideal market outcome. Their goal is to assess new leniency policies that guarantee the whistle-blowing member of a cartel immunity from prosecution even if the firm has originally taken an active part in the cartel agreement. These policies grant whistleblowers immunity from various fines and were recently adopted both in the US and by the European Commission. A similar leniency mechanism is also a key provision in the context of recent Czech anti-bribery legislation (see Lizal & Ortmann 2003).

ADS (2004) use stylized market games of Bertand price to compare four scenarios: the Standard case in which all cartel members are punished, the Leniency case in which whistleblowers are granted partial or full immunity dependent on how many other firms also blow the whistle, the theoretically best approach in which the incentives to blow the whistle are increased through rewards that consist of a share of the fines non-reporting members of the cartel are made to pay (Bonus), and a competitive market scenario in which cartel formation is theoretically not possible and empirically unlikely (Ideal; see Ortmann 2003). While the Ideal treatment is a one-stage game, the Standard, Leniency, and Bonus treatments are multi-stage games with a communication stage in which potential cartelists could hammer out agreements.

Importantly, for the Standard and Leniency treatments deductive game theory predicts a multiplicity of equilibria (all symmetric price equilibria) while it predicts a unique and symmetric lowest-price equilibrium for Bonus and Ideal. Note that the Bonus treatment precludes cartels and induces competitive bidding, at least theoretically. Note also that the Leniency treatment does not necessarily do so, and in fact leads to the same set of – from the point of view of the policy maker or regulator – undesirable consequences as the Standard treatment (although the incentives for high-price equilibria are weaker in the former than in the latter).

These theoretical predictions were tested with twelve groups of three subjects ("firms") in each of these treatments. The key results were:

1. The market price in the Standard treatment was significantly higher than the market price in the treatment that induced competitive bidding;
2. The Leniency treatment, in contrast, resulted in the second-lowest price; there was no significant difference between this price and the one induced by competitive bidding;
3. The Leniency treatment led to significantly lower prices than those in the Standard treatment, and (although, insignificantly) fewer cartels and more cartel members reporting;
4. The theoretically best approach ("Bonus") did not live up to its billing: Market prices were significantly higher than in the Ideal or Leniency treatments and statistically not different from the results in the Standard treatment. Moreover, this environment led to the highest (although not significantly) number of cartel formations.

3.   Methodological issues

There are many legitimate questions that one can ask about the external validity of this very small initial set of corruption experiments, or for that matter of experiments more generally. These questions fall roughly into two distinct categories: that of representative samples and that of representative stimuli (see Hertwig & Ortmann 2004).

As regards representative samples, most experimental economists work with a convenience sample of subjects – traditional college students. This is also true of all corruption studies reviewed above, with the paper by Barr et al. being the sole exception. As regards representative stimuli, most studies follow the convention among experimental economists of using abstract laboratory environments. The bribery game in AIR (2002), for example, is not framed in terms of bribe-giving or bribe-taking, nor is the rotation experiment in Abbink (2004) framed as such, nor is the Falk & Fischbacher (2000) experiment framed as a stealing experiment. In fact, only the experiments by Barr et al. and Frank and Schulze feature aspects of real life in their set-ups. One could call this a (one of many) framing issue (s). Another issue is that all experimental games that we report here are designed and implemented as one-shot, or finitely repeated, games whose game-theoretic predictions differ dramatically from those for indefinitely repeated games arguably pervasive in the game of life. It is therefore an interesting question to what extent subjects can understand the strategic situation they have been put in by the experimenter – a

situation that is rather different from what they encounter in their daily lives and for which they have probably adapted, and finely calibrated, good copying mechanisms (Cosmides & Tooby 1996; Juslin, Winman, & Olsson 2000). Since indefinitely repeated play tends to produce more trust and reciprocity, experimental results underestimate systematically the degree of corruption relative to that benchmark. To the extent that typically one-shot and finitely repeated game situations are implemented, findings of trust and reciprocity or retribution, in contrast, suggest a higher degree of corruption than predicted by deductive game theory.

We now briefly discuss the issue of representative samples and representative stimuli. The purpose of this part of the paper is not to give a comprehensive review of these issues but to convey to the reader that there is a voluminous literature out there on these issues and to provide a guide for those among the readers eager to explore these issues. The reader might want to think of it as an annotated list of further readings.

3.a. Representative samples

Harrison & List (2004) discuss the relevance of inferences drawn from laboratory experiments with convenience samples such as students. A simple way to address this issue is to run control treatments with other, less convenient, but arguably more representative samples. Alternatively, the authors argue, one can identify what it is that is non-representative about a given subject pool and then statistically correct for it. They also suggest – somewhat in contrast to the key earlier contribution by Ball & Cech (1996) -- that subject pool effects are something to worry about.

3.b. Representative stimuli

The issue of representative stimuli is more complex than that of representative samples. Harrison & List (2004; see also Carpenter, Harrison,& List 2005) mention four aspects of representative stimuli: the nature of the commodity, the nature of the task or trading rules applied, the nature of the stakes, and the nature of the environment that the subject operates in (e.g., whether it is abstract or not). All four aspects address ultimately the issue of how the laboratory setting is framed.

3.b.1. Framing

Questions about the nature of the commodity/service being traded or the nature of the task or trading rules applied are discussed in Harrison & List (2004). The authors document that it can matter whether one abstracts from the commodity/service traded, or the nature of the task. It is intuitive to expect a similar result in the context of corruption. As mentioned, the bribery game in AIR (2002), for example, is not framed in terms of bribe-giving or bribe-taking, nor is the rotation experiment in Abbink (2004) framed as such, nor is the Falk & Fischbacher (2000) study framed as a stealing experiment. Interestingly, Abbink & Hennig-Schmidt (2002) have tested for such framing effects of "neutral" and "loaded" instructions and find no statistical difference. We doubt that their result will be the last word in the matter of corruption experiments.

Questions about the nature of the stakes and the nature of the environment that the subject operates in are discussed in Harrison & List (2004) and also in Hertwig & Ortmann (2001, 2003). By and large, the articles by Hertwig & Ortmann suggest that the effects of stakes are in line with an economic theory of cognition: the higher the stakes the closer behavior moves to the game-theoretic prediction and the smaller the variance becomes. That said, as shown in Rydval & Ortmann (2004), cognitive capital (maybe produced by experience) is often a good substitute for cognitive labor (effort) and is therefore an important confound.

3.b.2.   Calibration

An issue related to the discussion in the previous section is the question to what extent the model underlying the experimental test are small-scale replications of the real world, or the "field". (Experimentalists maintain that their laboratory worlds are very real worlds.) Recently, important questions have been raised (e.g., Engelmann & Ortmann 2003, Healy 2004), for example, about the efficiency gains and implicit punishment possibilities of the early gift exchange, and trust or moonlighting, experiments. The results of Barr et al. (2003) exemplify a related problem: Can the numbers – a 200 percent in wage increases affecting only 30% reduction in corruption – guide policy considerations in field contexts? Or, how do the clever but ultimately simplistic experiments of Frank and Schulze (2000) and Schulze and Frank (2003) translate into policy guidance, if at all?

4.   Conclusion: Where experiments could be of use

What can be learned from the experiments on corruption, corruptibility, and procurement?

While the experimental research on corruption is premature to allow definite conclusions, some patterns emerge from the experiments already conducted:

- Words or deeds? The welfare-reducing effects on third parties hardly affect corrupt behavior (e.g., AIR 2002; Abbink 2002). This suggests that clean-hands campaigns or attempts to change the public sense of propriety through advertising in metro and trams are not likely to be successful in curbing corruption.
- Does deterrence work? Yes, it does. Increasing the probability of detecting bribe-giving and bribe-taking and the size of the punishment does by and large restrain corrupt behavior (e.g., AIR 2002; Frank & Schulze 2000; Schulze & Frank 2003; see also ADS 2004).
- Are detection probabilities correctly perceived? The results in AIR 2002 suggest that subjects do not gauge the detection probabilities correctly. This could well be a function of the probabilistic way in which the information was given to students (e.g., Hoffrage, Gigerenzer, Krauss, Martignon 2002) and at this point we should therefore not generalize from one bribery experiment.
- Do higher wages reduce corruption? Higher wages of officials do reduce corruption, but only when officials face the risk of detection and punishment (e.g., Frank & Schulze 2000, Schulze & Frank 2003; Azfar & Nelson 2003; Barr et al. 2003; but see Abbink 2002). Higher wages reduce corruption through the risk of losing a well-paying job if detected, an argument well-known from the literature

on efficiency wages. Increasing wages without subjecting the officials to the risk of punishment does not appear to induce sufficient loyalty to reduce the amount of corruption. The results of Barr et al. (2003) suggest furthermore that any increase in wages has to be considerable to affect significant reductions in corruption.

- Is corruptibility a function of people's perception of the pervasiveness of corruption in society? The results by Falk & Fischbacher (2002) suggest strongly that the extent of corruption in a society is a major determinant of corruptibility. The related but preliminary results of Babicky et al. (2004) confirm this result and suggest that inequality aversion is an additional determinant.
- Can skillfully formulated laws and regulations overcome the distant past of a country? How exactly do laws and regulations have to be formulated that stand a chance to effectively undermine the tenacity of the past? The answer to the first question seems affirmative (e.g., ADS 2004 but also all the evidence that shows that there are systematic treatment effects). The second question requires significantly more realistic modeling than is provided by models of Bertrand price competition. Such modeling seems eminently doable.
- Other results: rotation policy, transparency, and separation of power. The results of Abbink (2002) are intriguing but can't be the last word in this matter because rotation policies do have cost in terms of learning effects and reduced degrees of specialization. The intuitive transparency effects that Azfar & Nelson (2003) found warrant further investigation, as do the counterintuitive, and troubling. separation-of-power effects that the same authors found.

In the previous section we have sketched a number of important questions about the ability of experiments to help us understand the "real world", or to give policy advice. This discussion should not be misunderstood: Yes, experiments are not the panacea that everyone doing research on the determinants of corruption and corruptibility is desperately looking for but they do allow us to address many questions in systematic and relatively low-cost ways. If for example, someone objects to the efficiency gains and implicit punishment possibilities used in a particular experiment, it is easy to test the robustness of the experimental results with another parameterization. If someone objects to the stakes being too small to be telling us something, the experimenter can always scale them up. Even then experiments may be a low-cost means of exploration of the effects of detection probabilities, penalty sizes, and their interactions, and the importance of people's perceptions of the pervasiveness of corruption in society.

Most relevant, however, we believe them to be in the design of anti-corruption laws and regulations. Present anti-corruption law, for example, has some clear-cut problems that invite corruption. A case in point is the leniency provision. Take as an example the "regret" provision which – while well-intentioned – under certain conditions allows someone who has offered or promised a bribe to report this to the police or a public prosecutor and get off scot-free. At the same time, the person who accepted the bribe, might see significant prison time. It should be obvious (but obviously wasn't to those who write the law) that this scheme has undesirable incentive properties in that it gives the person who initiates the bribe the possibility to punish the bribee if he or she does not deliver on the proposal (illegal or sweet) deal. This provision of Czech anti-corruption law is hence another classic example where some experimental testing would have quickly revealed the pervasive incentive

effects. Experimental testing would also allow to study incentive-compatible and effective anti-corruption measures.

References

Abbink, K. (2004), Staff rotation as an anti-corruption policy: an experimental study. European Journal of Political Economy, forthcoming.

Abbink, K. (2002), Fair Salaries and the Moral Costs of Corruption. CeDEx Working Paper 2002-5, University of Nottingham.

Abbink, K., Irlenbusch, B., & Renner, E. (2002), An Experimental Bribery Game. Journal of Law Economics & Organization, 18, 428 – 454.

Abbink, K., Irlenbusch, B., & Renner, E. (2000), The Moonlighting Game. Journal of Economic Behavior and Organization, 22, 265 – 277.

Abbink, K. & Hennig-Schmidt, H. (2002), Neutral versus Loaded Instructions in a Bribery Experiment. Manuscript.

Andreoni, J. & Croson, R. (2004), Partners vs. Strangers; Randon rematching in public goods experiments. In: C.R. Plott & V.L. Smith (Eds), Handbook of Experimental Results. North-Holland, Amsterdam.

Apesteguia, J., Dufwenberg, M. & Selten, R. (2003), Blowing the Whistle, Working Paper 2003-5, Economics, University of Stockholm.

Azfar, O. & W.R.J. Nelson (2003), Transparency, wages, and the separation of power. An experimental analysis of the causes of corruption. Working Paper, U of New York (Buffalo), featured in Transparency International 2004 yearbook.

Babicky, V., Ortmann, A., Semerak, V., Soukenikova, L., & Vrazda, B. (2004), Willingness to Contribute When Incomes Are Earned: The Relative Importance of Reciprocity and Inequality Aversion, Presentation at the GDN RRC IV, Prague, August 2004.

Ball, S.B. & Cech, P.A. (1996), Subject pool choice and treatment effects in economic laboratory research. In: R.M. Isaac (Ed), Research in Experimental Economics (Vol. 6, pp. 239 – 292).

Barr, A., Lindelow, M., & Serneels, P. (2003), To Serve the Community or Oneself: The Public Servant's Dilemma, CSAE WPS 2003-11.

Becker, G. & G.J. Stigler (1974), Law Enforcement, Malfeasance and Compensation of Enforcers, Journal of Legal Studies 3, 1-18.

Berentsen, A., Bruegger, E., & Loertscher, S. (2003), On Cheating and Whisteblowing. Working Paper, Econonomics Department, University of Basel.

Berg, J.E., Dickhaut, J.W., & McCabe, K.A. (1995), Trust, Reciprocity, and Social History, Games and Economic Behavior 10, 122-142.

Bolle, F. (1990), High reward experiments without high expenditure for the experimenter? Journal of Economic Psychology 11, 157-167.

Bolton, G. & Ockenfels, A. (2000), ERC: A Theory of Equity, Reciprocity and Competition, American Economic Review 90, 166-193.

Burguet, R. & Che, Y.-K. (2004), Competitive procurement with corruption. RAND Journal of Economics 35, 50–68.

Camerer, C.F. (2003), Behavioral Game Theory: Experiments in Strategic Interaction. Russell Sage Foundation, New York, New York. Princeton University Press, Princeton, New Jersey.

Carpenter, J., Harrison, G., & List, J. (2005), Field Experiments. . In: R.M. Isaac (Ed), Research in Experimental Economics.

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

Charness, G., Frechette, G.R., & Kagel, J.H. (2004), How Robust is Laboratory Gift Exchange. Experimental Economics 7, 189-205.

Charness, G. & Rabin, M. (2003), Understanding Social Preferences with Simple Tests, Quarterly Journal of Economics 117, 817-869.

Cherry, T.L., Frykblom, P., & Shogren, J.F. (2002). Hardnose the Dictator. American Economic Review 92, 1218-1221.

Cialdini, R.B., Kallgren, C.A., Reno, R.R. (1991), A focus theory of normative conduct: A theoretical refinement and re-evaluation of the role of norms in human behavior. In: Zanna, M.P. (Ed.) Advances in Social Psychology, Vol. 24, Academic Press, New York, pp. 201-234.

Compte, O., Lambert-Mogiliansky, A. & Verdier, T. (2004), Corruption and Competition in Public Market Auctions. RAND Journal of Economics (forthcoming)

Cosmides, L. & Tooby, J. (1996), Are humans good intuitive statisticians after all? Rethinking some of conclusions from the literature under uncertainty. Cognition 58, 1 – 73.

Cox, J.C. (2004), How to Identify Trust and Reciprocity, Games and Economic Behavior 46, 260-281.

Dufwenberg, M. & Kirchsteiger, G. (2004), A Theory of Sequential Reciprocity, Games and Economic Behavior 47(2), 268-298.

Engelmann, D. & Strobel, M. (2004),

Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments. American Economic Review forthcoming.

Engelmann, D. & Ortmann, A. (2003), The Robustness of Laboratory Gift-Exchange. CERGE-EI Discussion Paper.

Falk, A. & Fischbacher, U. (2002), "Crime" in the lab – detecting social interaction. European Economic Review 46, 859-869.

Fehr, E., Kirchsteiger, G., & Riedl, A. (1993), Does Fairness Prevent Market Clearing? An Experimental Investigation, Quarterly Journal of Economics 108, 437-459.

Fehr, E., Gaechter, S., & Kirchsteiger, G. (1997), Reciprocity as a Contract Enforcement Device, Econometrica 65, 833-860.

Fehr, E. & Schmidt, K. (1999), A theory of Fairness, Competition and Cooperation, Quarterly Journal of Economics 114, 817-868.

Frank, B. (2004), Zehn Jahre empirische Korruptionsforschung. Vierteljahreshefte zur Wirtschaftsforschung 2004 (2), 184 – 199.

Frank , B. (1998), Good news for experimenters: subjects do not care about your welfare. Economics Letters 61, 171-174.

Frank, B. & Schulze, G.G. (2000), Does economics make citizens corrupt? Journal of Economic Behavior and Organization 43, 101–103.

Gneezy, U. & Rustichini, A. (2000), Pay enough of don't pay at all. Quarterly Journal of Economics ?, 791-810.

Gneuss, S. (2002), Strafmildernde Selbstanzeige und Korruptionsbekaempfung – Eine rechtsoekonomische Analyse. Wiesbaden, Deutscher Universitaets-Verlag.

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

Harrison, G.W. & List, J. (2004), Field experiments, Journal of Economic Perspectives, forthcoming.

Healy, P.J. (2004), Group Reputations and Stereotypes As Contract Enforcement Devices. Cal Tech working paper.

Hertwig, R. & Ortmann, A. (2001), Experimental practices in economics: A methodological challenge for psychologists? Behavioral and Brain Sciences 24, 383–403.

Hertwig, R. & Ortmann, A. (2003), Economists' and psychologists' experimental practices: How they differ, why they differ, and how they could converge. In: I. Brocas

& J.D. Carrillo (Eds),The Psychology of Economic Decisions (pp. 253-272), Oxford University Press, Oxford.

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, Boston, MA: Kluwer, 361-378.

Hoffrage, U., Gigerenzer, G., Krauss, S., & Martignon, L. (2002), Representation facilitates reasoning: what natural frequencies are and what they are not. Cognition 84, 343 – 352.

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

Juslin, P., Winman, A., & Olson, H. (2000), Naïve empiricism and dogmatism in confidence research: A critical examination of the hard-easy effect. Psychological Review 107, 384-396.

Klemperer, P. (2004), Auctions: Theory and Practice. Princeton: Princeton University Press.

Kreps, D.M. (1990), Game Theory and Economic Modelling. Oxford: Oxford University Press.

Laffont, J.-J. & Martimort, D. (2002), The Theory of Incentives. The Principal-Agent Model. Princeton: Princeton University Press.

Lizal, L. & Ortmann, A. (2003), Designing and Testing Incentive-Compatible and Effictive Anticorruption Measures. Grant proposal to GA CR. See home.cerge-ei.cz/Ortmann/PAPERS/CERGEEIGrantAgCRApplicMarch2003.PDF

Mas-Colell, A., Whinston, M.D., & Green, J.R. (1995), Microeconomic Theory. Oxford: Oxford University Press.

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

Mauro, P. (1997), The Effects of Corruption on Growth, Investment, and Government Expenditure: A Cross-Country Analysis. In: K.A. Elliott (Ed), Corruption and the Global Economy. Washington, Institute for International Economics, 83-107.

Mauro, P. (1998), Corruption and the Composition of Government Expenditure. Journal of Public Economics 69, 263-279.

Milgrom, P. (2004), Putting Auction Theory to Work. Cambridge: University of Cambridge Press.

Ortmann, A. (2004), Corruption Prague. The Prague Post, August 5, A5.

Ortmann, A. (2003), Bertrand Price Undercutting: A Brief Classroom Demonstration. Journal of Economic Education 34, 21-26.

Ortmann, A. & Colander, D. (1997). A Simple Principal-Agent game for the class room. Economic Inquiry 35, 443-450.

Ortmann, A., Fitzgerald, J., & Boeing, C. (2000), Trust, Reciprocity, and Social History: A Reexamination, Experimental Economics 3, 81-100.

Renner, E. (2004), Wie laesst sich die Korruption wirksam bekaempfen? Vierteljahreshefte zur Wirtschaftsforschung 2004 (2), 292 – 300.

Roth, A. E. (2002), The Economist as an Engineer: Game theory, Experimentation, and Computation as tools for Design Economics. Econometrica, 70, 1341-1378.

Rubinstein, A. (2004), Instinctive and Cognitive Reasoning: Response Time Study. Working paper.

Rydval, O. & Ortmann, A. (2004), How financial incentives and cognitive abilities affect task performance in laboratory settings: an illustration. Economics Letters 83, 315-320.

Schulze, G.G. & Frank, B. (2003), Deterrence vs Intrinsic Motivation: Experimental evidence on the determinants of corruptibility. Economics of Governance, 4, 143-160.

Spagnolo, G. (2004), Divide et impera: Optimal Leniency Policies for Cartel Deterrence. University of Mannheim Working paper.

Steiner, J. (2004), A Trace of Anger Is Enough: On the Enforcement of Social Norms. CERGE-EI Working paper.

Torgler, B. (2003). Tax Morale in Central and Eastern European Countries, Conference oin "Tax Evasion, Trust, and State Capacities" in St. Gallen (Oct 17-19, 2003)

Treisman, D. (2000), The causes of corruption: a cross-national study. Journal of Public Economics 76, 399–457.

Van Rijckeghem, C. & Weder, B. (2001), Bureaucratic Corruption and the Rate of Temptation: Do Wages in the Civil Service Affect Corruption, and by How Much? Journal of Development Economics 65, 307-331.

Vega-Redondo, F. (2003), Economics and the Theory of Games, Cambridge: Cambridge University Press.

Yezer, A.M., R.Goldfarb, & Poppen, P.J. (1996), Does studying economics discourage cooperation? Watch What We Do, Not What We Say or How We Play, Journal of Economic Perspectives 10, 177-186.