On Monetary and Non-monetary Interventions to Combat Corruption

Ritwik Banerjee
Indian Institute of Management

Arnab Mitra
Portland State University, arnab@pdx.edu

Citation Details
Banerjee, Ritwik & Mitra, Arnab, 2018. "On monetary and non-monetary interventions to combat corruption," Journal of Economic Behavior & Organization, Elsevier, vol. 149(C), pages 332-355.

On Monetary and Non-monetary Interventions to Combat Corruption

Ritwik Banerjee
Indian Institute of Management Bangalore
Bengaluru
Karnataka, India 560076

Arnab Mitra
Department of Economics
Portland State University
Oregon, USA 97201

Abstract: We study the relative effectiveness of extrinsic monetary disincentives and intrinsic non-monetary disincentives to corruption, using a harassment bribery game. In doing so, we also test the Beckerian prediction that at the same level of expected payoff, a low probability of detection with high fine is a stronger deterrent to corruption than a high probability of detection with low fine. In Experiment 1, two treatments are designed to study the effect of a low probability of detection with high fine and a high probability of detection with low fine, on bribe taking behavior. In Experiment 2, subjects participate in the same baseline harassment bribery game either without or after having gone through a four-week ethics education program. Results show that: (a) a low probability of detection with high fine reduces both the amount and the likelihood of bribe demand, (b) a high probability of detection with low fine has no effect on bribe demand, (c) normative appeals of ethics education has a small effect on the likelihood but not on the amount of bribe demand, when measured immediately after the intervention, (d) the effect of ethics education vanishes when measured four weeks after the intervention.

JEL codes: C91 C92 D03 K42

Keywords: corruption, harassment bribes, fine/penalty, probability of audit, ethics education

☆ The authors have contributed to this article equally, and senior authorship is not assigned. The authors thank Marie Claire Villeval for her prescient comments on an earlier version of the paper. The authors are also thankful to seminar participants at Portland State University (Randall Bluffstone, John Gallup, Forrest Williams and James Woods), IIM Bangalore, ISI Delhi, Ashoka University, Shiv Nadar University for the general discussions and comments. The authors are indebted to Gautam Gupta and Abhishek Das for their cheerful patronage. Financial support from the Danish Council for Independent Research, Social Sciences, Grant no: DFF-1327-00037; and the Faculty Enhancement Award 2016–17, Portland State University is gratefully acknowledged.

⁎ Corresponding author.
Email addresses: ritwikbanerjee@iimb.ac.in, ritwikbanerjee@outlook.com (R. Banerjee); amitra@pdx.edu (A. Mitra)

"[…]it is efficient to hang offenders with probability zero."

\- S.C. Kolm, 1973

1. Introduction

Ashoka, the emperor who once ruled ancient India, propagated among his subjects the idea of *Dhamma* (loosely translated as righteous or moral life), through a series of edicts inscribed on stone pillars. After having laid out the demands of *Dhamma*, the great monarch surmised that though regulation and persuasion were two means available for its promotion, it was persuasion which was more effective. Despite having the entire state machinery at his disposal, Ashoka thought a change of heart through persuasion had a far greater effect than regulation.

In this paper, we evaluate the efficacy of regulation and persuasion when it comes to moral decision making (e.g. whether to engage in corruption or not), i.e., we ask if extrinsic (dis)incentive induced by a regulatory framework is more successful in dissuading people from making immoral choices, or if inducing a higher intrinsic moral cost is more effective. In other words, we investigate whether Ashoka was indeed right in his surmise; some 2400 years later.

The focus of economic theories that study different aspects of corruption has been on comparing the private returns to corruption versus honesty. From this "rational" vantage point, the overarching objective has been to design institutions which offer a credible disincentive to corruption by suitably altering the two principal anti-corruption policy levers, namely, probability of audit and penalty or fine (Becker and Stigler,1974; Rose-Ackerman, 1975). These two policy levers have remained the bulwark of anti-corruption mechanisms since then.[1] The idea that intrinsic motivation (i.e., the moral cost associated with the shame and guilt of making unethical choices) may serve as a prominent corruption restraining trait has emerged in the literature rather recently. Cultural anthropologists and sociologists have proposed that a change

[1] That is not to say that alternative policies have not been examined or tested. For instance, Basu (2011) proposed and Abbink et al. (2014) examined asymmetric liability for the suppliers and recipients of corrupt practices, Ryvkin and Serra (2012) analyzed the effects of staff rotation, Shleifer and Vishny (1993) proposed competition among public officials.

in the “value system” be made as opposed to the “incentive system”, to make a significant dent into corruption (Bardhan, 2006). According to this view, education and awareness campaigns alongside a strong social disapproval mechanism are effective measures that help change the value system. Notice, this is “persuasion” in the sense Ashoka used the term.

We investigate the fault line between “rational” extrinsic incentive driven approach and intrinsic moral cost driven approach, using a laboratory corruption experiment. In particular, we aim to identify the relative strengths and weaknesses of the two approaches with the help of a harassment bribery game. Within that overarching objective, we also investigate which of the two policy levers, probability of detection and fine, has a stronger bite in terms of reducing corruption. Thus, in addition to analyzing non-monetary and monetary disincentives to corruption, our study contrasts the alternative mechanisms within the monetary disincentive framework.

We take an experimental route to study the effects and the deeper mechanisms driving corrupt behavior. Our methodological choice of using laboratory based experimental tools for investigating these questions is guided in part by the complexity of the questions, nature of the outcome of interest (namely, corruption), and recent methodological advances made in the context of laboratory corruption games. First, identification of the effects of either monetary or non-monetary interventions using observational data is difficult because of endogeneity related issues. Any such attempt will likely give us at best correlational, not causal, results. Second, corruption is a clandestine activity and may not be measured without measurement errors. And third, studies show that results obtained through laboratory experiments, particularly laboratory corruption games, are externally valid and they do measure moral cost of engaging in corruption (Armantier and Boly, 2013; Banerjee, 2016).

We conduct two experiments – one each to study monetary and non-monetary disincentives to corruption. Both experiments share the same canonical harassment bribery game.[2] In the bribery game, a "Citizen" (C, hereafter) performs a real effort task. If successful, the C is entitled to a prize, subject to a procedural approval from a "Public Official" (PO, hereafter). However, the PO may demand a bribe from the C to approve of the prize.[3] The C can either (a) accept the bribe demand, in which case she earns the prize less the bribe, and the PO earns his salary plus the bribe, or (b) reject the bribe demand, in which case both individuals earn their basic payoffs. In this baseline treatment, the bribery game does not have any anti-corruption mechanism.

Experiment 1 introduces two treatments with different anti-corruption mechanisms. In the first treatment, the players engage in the same bribery game as before, however, there is a small probability that the PO is audited. If the audit finds him accepting bribe, he pays a heavy fine. On the other hand, in the second treatment, there is a high probability that the PO is audited. However, if the audit finds him accepting a bribe, he pays a small fine. The two treatments are so designed that the PO's expected payoff at each bribe demand level stays constant.

The two audit treatments offer us a novel setting where Becker's (1968) classic prediction – for risk averse individuals, quantum of fine serves more of a deterrent than an equivalent increase in chances of being caught – can be tested. There is a further policy-relevant, critically important reason to examine which of the two strategies is more effective. Clearly, if the two strategies are costless, a planner will choose the maximum value for both the policy levers. However, in practice, increasing the probability of detecting corruption is costly, e.g., the

[2] A harassment bribery game, as opposed to games based on collusive bribery, simulates a situation where a government official can potentially extort bribe from a citizen for a service the citizen is naturally entitled to, e.g., obtaining a driving license after having passed the driving test. Past studies such as Abbink et al. (2014)

and Banerjee (2016) have used different versions of laboratory based harassment bribery games.

[3] From now on we will treat the PO as male and the Citizen as female.

government needs a larger number of vigilance officers, increased capability of existing anti-corruption personnel, and a large, swift and effective judiciary. Increasing fine too is expensive: first, an increase in fine raises the cost of getting a conviction, and second, increasing fine in the statutes involves significant political costs. As a result, a planner faces a trade-off (in terms of cost) between probability of detection and fine. Consequently, the efficacy of these two alternative strategies becomes the decisive question.

Results from Experiment 1 suggest that low probability of detection and high fine significantly reduces both the amount and the likelihood of bribe demand, but high probability of detection and low fine has no effect on either. In other words, a high fine offers a stronger deterrence to bribe demanding behavior than a high probability of detection at the same expected payoff level.

We design a second experiment aimed at raising the moral cost or intrinsic disincentive for the unethical act of demanding bribe. Many business schools across the globe teach courses on ethics – the aim supposedly is to change one's perception of the moral (in)appropriateness of an action. We test the efficacy of one such ethics teaching module with bribe demanding behavior as an outcome variable of interest. In the baseline treatment of Experiment 2, the same harassment bribery game is implemented, without subjects having been through the ethics education. In Treatment 1, subjects participate in a four-week ethics module and in the fifth week they participate in the laboratory based harassment bribery game. In Treatment 2, subjects participate in the same four week ethics module, but play the bribery game only in the eighth week. Thus, the experimental treatments vary whether a subject goes through the ethics module or not and when the outcome of interest, namely unethical behavior as measured through the bribery game, is observed. To the best of our knowledge, this is the first carefully crafted experimental design that attempts to evaluate whether ethics education is effective, and it also helps us identify whether such effects, if any, hold longitudinally.

In an influential review paper, Abbink and Serra (2012) discuss the role of framing effects and externalities in reducing corruption, in the context of nonmonetary incentives and intrinsic motivations as anticorruption tools. Perhaps, their attempt can be construed as the first logical step in the formation of hither to thin experimental corruption literature on non-monetary interventions. The next pivotal question is: can behavioral change be induced by normative appeals, such as ethics courses? The relevance of evaluating the impact ethics courses may potentially have on curbing corruption cannot be overstated in today's world. Firstly, ethics is taught at the school level as moral science in many parts of the world. Secondly, in the recent past ethics has been introduced as a compulsory subject in a wide range of professional courses (such as Masters in Business Administration). The genesis of such introduction can be traced back to the gross unethical conduct in accounting scandals in Enron in 2001, financial crisis in 2008, and more recently, LIBOR rate fixing episode which involved a slew of international banks. As Beyond Grey Stripes reported in 2011, the number of MBA programs which had a compulsory ethics course increased from 34% to 79%, the aim of which presumably was moral transformation.[4] Whether or not such courses induce the desired goals is still an open question.

Results from Experiment 2 indicate that normative appeals of ethics education act in ways that are very different from extrinsic disincentive driven approaches. First, ethics education did not have any effect on the amount of bribe demand either in Treatment 1 or Treatment 2. Second, it did have an effect on the likelihood of bribe demand in Treatment 1, but not in Treatment 2. These results indicate that any effect ethics education may have had on the decision to engage in

[4] In a related but distinct literature, it has been shown that studying economics makes people more selfish (rational, if you will) (List et al., 2001; Frey and Meier, 2003; Haucap and Müller, 2014), and care less about fairness, more about efficiency (Fehr, Naef and Schmidt, 2006; Faravelli, 2007).

bribery (or not) is short-lived and does not survive beyond the fourth week. The results also suggest that the underlying mechanism through which normative appeals work is very different from the mechanism through which monetary disincentives function.

The two experiments are conducted with different subject pools. As such, a direct comparison of the treatment effects may not be too meaningful. However, Experiments 1 and 2 share the same baseline treatment – treatments are then introduced within each experiment. This allows us to measure the treatment effects relative to the respective baseline results within each experiment. Besides, the identical underlying structure in both the experiments allows us to pin down the behavioral channels through which monetary and non-monetary interventions work. Our treatment effects, when normalized relative to the baseline, indicate that not only does ethics education not have any statistically significant effect on bribe amount, the relative effect that it has on the likelihood of bribe demand is quantitatively smaller than that in low probability of audit with high fine. We also document that extrinsic incentive driven monetary interventions create perverse effects owing to strategic considerations unlike non-monetary interventions. Overall, in our setting, regulation has been more effective in reducing bribe demand than persuasion through normative appeals.[5]

Our contributions to the literature are several. First, in view of the fact that monetary and non-monetary interventions are hard to compare in observational settings, we provide a credible experimental structure using which the issue can be studied. Second, we study the relative effectiveness of the fine amount and the probability of audit in curbing unethical conduct – our novel approach keeps the expected payoff under the two parameter configurations the same and

[5] We acknowledge that since normative appeal (intrinsic motive) is not quantifiable, it is difficult to formulate an appropriate ethics education module that is exactly equivalent to a (quantifiable) monetary disincentive (extrinsic motive) based mechanism in terms of curbing corruption. As such, an insightful discussion on the relative comparison between the effectiveness of the ethics module and the effectiveness of the audit treatments is beyond the scope of this paper.

thus, allows us to test an important implication of the Beckerian paradigm of studying corruption that postulates that the effects of the two policy instruments are asymmetric. Third, we use beliefs about others' behavior to delineate the deeper psychological mechanisms at work when it comes to efficacy of different anti-corruption tools. Fourth, our study examines not only the very short run instantaneous effect of normative appeals on corruption, but also the intermediate effect after a significant passage of time.

The remaining sections are organized as follows: Section 2 offers an overview of the literature, Section 3 lays down a set of clearly defined hypotheses that may be tested using our experimental data, Section 4 lays out the experimental design and results from Experiment 1, Sections 5 present the experimental design and results from Experiment 2. Section 6 compares the results from the two experiments, offers policy recommendations, and Section 7 concludes.

2. Overview of the Literature

The effects of the two instruments, probability of detection and penalty (or fine), have been independently studied in quite a few contexts. In the tax compliance literature, studies show a small but significant deterring effect of an increase in detection probability (Slemrod et al., 2001), but no effect of an increase the associated fine (Pommerehne and Weck-Hannemann, 1996; Torgler 2005). A survey paper by Fischer et al. (1992) concludes that tax compliance increases with probability of audit or detection. The nature of the collected data in these studies is often controversial and potential endogeneity issues are many. Besides, only a few studies deal with changing both fine and probability of audit at the same time – while Alm et al. (1992) find tax compliance increases with audit rates more than it does with fine rates, Park and Hyun (2003)'s findings are opposite. However, even the last two studies do not change the probability and fine parameters in a manner such that the expected payoff is unchanged, thus making comparisons difficult.

In the context of criminal behavior, it is generally agreed that increases in probability of detection and fine reduce crime (Eide, 2000). This literature concludes that in terms of magnitude, certainty of punishment is a stronger deterrent to crime than severity of punishment. However, many of these studies rely on self-reported data collected from individuals with criminal history (see e.g. Grogger, 1991). Other contexts that address similar questions include Stafford (2002), who estimates the impact of an increase in fines on hazardous waste violations; and Shimshack and Ward (2005), who analyze the effect of past fines and frequency of inspection on water pollution.

Two papers attempt to change the probability and amount, for the same level of expected payoff and both take the experimental route. In the gain domain, Bruner (2009) studies the effect of changes in probability of winning rewards in a context free lottery and changes in reward amount, keeping the expected payoff the same. His finding that people prefer an increase in probability to an increase in reward for the same expected value is consistent with predictions from expected utility theory. In the loss domain, Friesen (2012) finds that severity of punishment is a stronger deterrent than certainty of punishment, when it comes to the issue of compliance. Though strategies to combat corruption has taken the center stage of discussions in many national and international fora, there is surprisingly little evidence regarding the efficacies of the two principal policy levers, namely, probability of detection and fine.

In the context of efficacy of teaching ethics, Konow (2014) evaluates the impact of philosophy class on fairness views and of studying business ethics on generosity and cooperation. However, it is hard to argue that an un-generous or non-cooperative behavior is also unethical, which perhaps leads to a disconnect between the topic of lecture and the outcome of interest. Waples et al. (2009) conduct a meta-analysis that shows there is, if any, a minimal impact of ethics courses on outcomes. In a related study, Antes et al. (2010) find that a course on responsible conduct of research has little impact on actual ethical decisions scientists take. The limitations of these studies lie in their outcome measures, which are attitudinal and not behavioral. In our study, we measure a behavioral outcome, which is precisely what the imparted ethics education aims to address, namely, unethical conduct. Once again, despite normative appeals being widely invoked in popular discourse to engender honest behavior among common people, civil servants or politicians, there has been no systematic study that analyzes the impact of such appeals on corruption.

Our main outcome of interest is corruption behavior, which is essentially clandestine in nature and therefore, difficult to measure. Luckily a large experimental literature has evolved which suggests that corruption behavior may be credibly measured through lab based experimental games (see Abbink and Serra (2012) for a review). Our ability to observe corruption in a controlled environment gives us a nice set up in which anti-corruption policies may be tested. Some extant studies have creatively used such games to study anti-corruption policies. For instance, Abbink, Irlenbusch and Renner (2002) introduce exogenously determined "sudden death" in their bribery game and find that a small threat to disqualification is a strong deterrent; a finding also observed in Berninghaus (2012). On the other hand, Schulze and Frank (2003) do not find any effect of introducing fine on corruption. The apparent contradiction between the two studies has been explained by differences in experimental protocol, payment rule etc. Interestingly, the two studies differ in one crucial aspect, which has been largely ignored by the literature - Abbink, Irlenbusch and Renner (2002) implement higher penalties and Schulze and Frank (2003) higher probability of detection. Our paper can potentially throw a light on this difference in parametrization as a reason for divergence in the results – we discuss this aspect in Section 7. Although some studies focus on social appropriateness norm violation (Banerjee, 2016) and externality as forms of non-monetary disincentive to corruption (Abbink et al., 2002; Barr and Serra, 2009) and others study the effectiveness of social observability on corruption (Salmon and Serra, 2016), none of them compare the relative effectiveness of extrinsic and intrinsic disincentives to corrupt behavior. One recent exception is Cason et al. (2016), who compare compliance behavior in the presence of formal regulatory framework with fine and informal mechanisms such as information observability. They find that social observability has no effect on misreporting behavior. Social observability is one important form of non-monetary channel. Our paper studies normative appeal as a form of non-monetary intervention. In doing so, we contribute precisely at the intersection of the monetary and non-monetary interventions, by analysing their relative effectiveness in the context of a standard harassment bribery game.

3. Hypotheses

According to Becker, an agent who decides whether to commit a crime (corruption, in our case), essentially compares the benefits of committing versus not committing the crime. In our framework, the PO decides whether to demand bribe or not. The expected utility of not demanding a bribe is given by

$$EU(honesty) = U(w),$$

where w is the wage he receives.[6] The expected utility of obtaining a bribe is given by

$$EU(corruption) = pU\big(w - F(b)\big) + (1-p)U(w+b) - m(b) \qquad (1),$$

[6] Since the PO's payoff from honesty does not involve any uncertainty, this is actually a payoff with certainty.

where p is the probability of audit, b is the bribe amount, $F(b)$ is the fine if audited ($F'(b) > 0$), and $m(b)$ is the intrinsic moral cost of demanding the bribe.[7] In Experiment 1, the moral cost component remains unchanged, while we manipulate p and F to change the expected utility from engaging in corruption. On the other hand, our main objective in Experiment 2 is to influence expected utility by changing $m(b)$ through ethics education.

Experiment 1 is based on two anticorruption policy levers, namely probability of audit (p), and quantum of fine (F). Introduction of p and F in itself reduces the expected utility of a (risk averse) PO and therefore acts as a deterrent. Such an environment will be characterized by the following two features: one, a rational C will expect less number of POs to demand bribe and two, if fine is monotonically related to bribe amount demanded, then the predictions for the proportion of the POs demanding bribe will also extend to the amount of bribe demanded. This leads us to the first set of empirically testable hypotheses.

Hypothesis 1A: *The proportion of corrupt POs and the amount of bribe demanded are higher in an environment where there is no chance of detection than one in which there is a chance of an audit and associated fine.*

Hypothesis 1B: *The proportion of Citizens who believe a bribe will be demanded and the expected bribe amount are higher in an environment where there is no chance of detection than one in which there is a chance of an audit and associated fine.*

Expected utility framework gives us some more interesting predictions: If the expected payoff remains unchanged, a (bribe demanding) risk averse PO prefers high probability of detection and low fine over low probability of detection and high fine. Appendix 1 lays out the theoretical framework for the prediction, and Figure A1.1 illustrates the same[8]. A corollary of

[7] Strictly speaking, the utility functions are indirect utility functions, which are often denoted by $V(\cdot)$.

[8] It is important to emphasize that the expected utility framework discussed above is based on just that: expected utility. A mathematical model for the framework is given in Appendix 1. The model shows that, for a given bribe demand and fixed expected payoff, a risk averse PO will derive higher expected utility under a policy of high probability of detection with low fine than under a policy of low probability of detection with high fine. However, the model does not accommodate the situation in which the PO considers the game as an expected utility

that prediction is that the two enforcement systems are equally effective for risk neutral individuals. Consequently, a rational Citizen will expect low probability of detection with high fine to have a higher bite in terms of reducing corruption than high probability of detection with low fine. A Citizen will also prefer to accept a higher bribe demand in the presence of audit and asymmetric liability i.e. if she is not liable to penalty but the PO is. This gives us a second set of testable hypotheses.

Hypothesis 2A: The proportion of POs who demand bribe and the bribe amount are higher in an environment with high probability of detection and low fine, than in an environment with low probability of detection and high fine, at the same level of expected payoff.

Hypothesis 2B: The proportion of Citizens who expect POs to demand bribe and the expected bribe amount are higher in an environment with high probability of detection and low fine, than in an environment with low probability of detection and high fine, at the same level of expected payoff.

Hypothesis 2C: The maximum bribe acceptable to Citizens will be higher in the audit treatments than that in the control treatment.

In Experiment 2, we aim to change the incidence of bribe demand by altering the moral cost of engaging in corruption, $m(b)$, by introducing ethics education. An ethics education is expected to increase the moral cost and thereby lower the expected utility from demanding bribe. We expect a lower bribe demand due to an increase in the intrinsic cost of demanding bribe. However, the effect of the ethics education may not remain unchanged with passage of time. In particular, we may expect that the mitigating effect of ethics education on corruption may taper (or even disappear) with time. Finally, so far as the relative effectiveness of intrinsic versus extrinsic cost of engaging in corruption is concerned, economic theory does not tell us anything

maximization exercise with respect to b. In such a situation, the policy maker may be interested in assessing how b^* will vary across the two policies. Such a situation can be captured by a more complicated model but it is beyond the scope of this paper. However, even in this model, it can be shown (under reasonable conditions) that a policy of low probability of detection with high fine will have a stronger bite on b^*, than a policy of high probability of detection with low fine.

about this. That, after all, is an empirical question. Thus, we have the next set of empirically testable hypotheses.

Hypothesis 3A: The proportion of POs who demand a bribe and the bribe demand are lower for the subjects who go through ethics education than for those who do not go through ethics education.

Hypothesis 3B: The immediate effect of ethics education on POs' bribe demand behavior is greater than the longer run effect.

Hypothesis 3C: The proportion of Citizens who accept bribe are lower for the subjects who go through ethics education than for those who do not go through ethics education. The immediate effect of ethics education on Citizens' acceptance behavior is greater than the long run effect.

Hypothesis 3D: The effect of ethics education or intrinsic motivation on POs bribe demand behavior, relative to the control, may be higher or lower or equal to the effect of extrinsic motivation oriented policies such as introducing probability of detection and fine.

We also expect that the Cs' beliefs will be consistent with the POs' actual decisions. The next section lays out the experimental design that allows us to test the above hypotheses.

4. Experiment 1

Each treatment in Experiment 1 comprises of multiple stages. We begin by describing Stage 1 that is common to all treatments. The subsections will then focus on the other stages that are specific to a given treatment.

In Stage 1 of each treatment, a C is randomly matched with a PO, and they play a one-shot game.[9] The Cs and the POs are located in two different rooms. Each C is paid a basic fee of 400 Mohor for performing a task, and each PO is paid a salary of 1000 Mohor (M, hereafter) for grading the task and (potentially) approving the prize.[10] The rules of the game, the set of strategies and the payoff function for each player are common knowledge to both players.

---

[9] We describe the experimental treatments from the perspective of a randomly matched representative pair that includes a C and a PO. The tasks of all other matched pairs are identical to the tasks of the representative pair.

[10] The experimental currency is termed as "Mohor". The conversion rate between Mohor and Indian Rupee (INR) is set at 1M = 0.4 INR. Therefore, the basic fee (salary) for a C (PO) is 160 INR (~ 9.41 PPP USD). In PPP terms, 1 US Dollar = 17 INR (source: http://data.worldbank.org/indicator/PA.NUS.PPP).

To start with, the C is asked to perform a simple task to qualify for a prize of 600M (over and above the basic fee of 400M). The experimenter provides each C with five sequences of letters on a piece of paper. The task requires each C to correctly indicate within seven minutes the number of times the letter "A" occurs in each sequence.[11] If a C correctly solves at least three of the sequences, then she is entitled to the prize. The experimenter provides the correct answers to the matched PO, who then determines whether the C has qualified for the prize by providing at least three (out of five) correct answers. In addition, the experimenter announces the correct answers in the Cs' room. Given the simplicity of the task, each C is expected to get at least three answers correct, and thereby form a sense of entitlement to the prize.[12]

4.1 Control Treatment

Figure 1 depicts the Control Treatment. After Stage 1, subjects proceed to Stage 2, which begins when the PO determines whether the C has won the prize. Following that determination, the PO chooses any one of four possible strategies: approve the prize, demand a bribe of 200M, or 400M, or 600M in return for approving the prize. While the PO chooses his strategy, the C simultaneously and independently decides whether to accept or reject each possible bribe demand the PO might opt for. Stated differently, the experiment employs the strategy method, wherein the C indicates any one of two possible strategies ('accept' or 'reject') for each possible bribe demand. Once the decisions are made, each player turns in the decision sheet to the experimenter.

[Figure 1 here]

[11] A typical sequence is: AEACBCBCCAEEAABBADEADDEADEADB.

[12] It turned out that each of 114 Citizens, who participated in the experiments, provided at least three correct answers. However, it was important that the citizens "earned" the prize rather than being endowed with it – this approach helped in engendering the sense of entitlement and subsequently a sense of harassment, if and when bribe is demanded.

The final payoffs from the game are determined by matching the two players' decisions. If the PO *approves*, both players earn 1000M (PO earns salary $(= w)$, and C earns prize plus basic fee). On the contrary, if the PO *demands a bribe* $b \in \{200\text{M}, 400\text{M}, 600\text{M}\}$, two possibilities emerge: the C accepts, in which case she earns $(1000\text{M} - b)$ and the PO earns $(1000\text{M} + b)$; or the C rejects, in which case (the transaction does not take place and) she earns 400M and the PO earns 1000M. Since the base earning of the PO is at least as large as that of the C, it ensures that bribe demand cannot be explained by other behavioral channels such as inequity aversion.

In addition to their decisions, the C and the PO respond to a set of questions related to their beliefs about what strategy their matched partner is likely to take. In particular, the C reports her expectation of whether the PO will demand bribe, and the PO reports his expectation of whether the C will accept a bribe demand. Additionally, each PO is asked (a second-order belief question) about the amount of bribe he believes the C thinks PO will demand; the choices being 0M (approve the prize), or 200M, or 400M, or 600M. Each subject is incentivized to report her belief accurately by an additional payment of 25M for each belief-related question, if the belief matches the actual response of the partner. The rich data on beliefs allow us to identify the deeper psychological mechanisms at play. Besides, a player's belief about the matched player's decisions may reflect how the former expects a representative individual of the larger society to act in a situation that potentially involves corruption.

The Control Treatment described above is devoid of any anticorruption mechanism. As such, the only cost of indulging in corruption is one of intrinsic moral cost. Each audit treatment incorporates a specific stage where an anticorruption mechanism is implemented.

4.2 Audit Treatments

Figure 2 depicts the Audit Treatments. The two Audit Treatments are almost similar to each other, except for a couple of parameters. We proceed by describing a generic Audit Treatment, followed by a discussion of the parameters that distinguish the treatments.

[Figure 2 here]

Notice that an Audit Treatment is comprised of three stages, where Stages 1 and 2 are identical to those of the Control Treatment. A new stage (Stage 3) is introduced at the end of Stage 2. The new stage involves an audit with probability p of all successful transactions (between C and PO) in which PO either approves the prize or demands a bribe which C accepts. Since no further consequence is needed if an audit finds that PO has approved the prize, for simplicity, Figure 2 does not depict the situation that there is an audit with probability p even when PO approves the prize. Instead, the figure shows that an audit is conducted with probability p for only those transactions in which a bribe is demanded *and* accepted.[13] If the audit finds a bribe demand, the C gets the bribe amount back and a fine is deducted from the PO's salary. The experimental design gives immunity to the C when she accepts a bribe demand, which conforms with the extant literature on harassment bribe (Basu, 2011; Abbink et al., 2014).[14]

Our main goal is to study the POs' decisions under two equivalent audit treatments. The equivalence originates from the fact that for a representative PO, the expected payoff from an audit remains unchanged across the two treatments, for each possible bribe demand: in one treatment the probability of an audit is low and the fine is high, whereas, in the other, the

[13] The audit is not directed to those situations where a bribe demand is rejected. In reality, if law-enforcement budget is limited, it might be difficult for authorities to determine (in an ex-post sense) whether a given potential transaction failed due to a rejection of bribe demand, or due to incorrectly filed documents that have been returned to the citizen. As such, we decided to restrict the scope of the audits to successful transactions.

[14] The design acknowledges the fact that in numerous instances, a citizen in India is literally forced to pay a harassment bribe for critical entitled services, when faced with the alternative option of the service not being provided. As such, for a citizen, acceptance is a strictly dominant strategy over rejection. It will be clear in the next section that, given our experimental design, a citizen's expected payoff from acceptance of a bribe demand is more than the payoff from rejection.

probability of an audit is high and the fine is low. Using the previous notations, a PO's expected payoff from demanding a bribe of b is given by

$$E(\pi) = p\big(1000 - F(b)\big) + (1 - p)(1000 + b) \qquad (2)$$

To test Hypothesis 2A, we design two treatments with identical $E(\pi)$; one with low p and high F, and the other with high p and low F. The experimental design is informed by plotting the iso-expected-payoff curves (IEPC) for each bribe demand. An IEPC is a locus of p and F such that $E(\pi)$ remains the same along a curve. The slope of an IEPC is obtained by totally differentiating equation (2), setting $d\big(E(\pi)\big) = 0$, and rearranging terms

$$\left.\frac{dF}{dp}\right|_{d(E(\pi))=0} = -\frac{F + b}{p} \qquad (3)$$

[Figure 3 here]

Figure 3 plots the IEPCs for three different bribe levels in the $p - F$ space and it is characterized by the following features: one, expected payoff from bribe demand is higher than that from being honest; two, higher the bribe demand, higher is the expected payoff; three, higher the bribe demand, higher is the imposed fine under an audit. We choose two (p, F) combinations for each expected payoff level. Each combination is realistic and has precedence of earlier use in the literature, in addition to being easy for subjects to make heuristics calculations with. It is worth noting here that parameterizing probability of detection is tricky. There is no reliable estimate of what proportion of all corruption is actually audited in reality, given the clandestine nature of the activity. As a result, we rely on parameters used in past studies as reference points. For instance, Friesen (2012) employ inspection probabilities of 10%, 20% 30% … 90% but does not find a differential deterrence effect between inspection probabilities of 10% and 20%. In a similar fashion, Bruner (2009) employs probability of reward in increments of 10 percentage points, from 10% to 100%. India National Crime Records Bureau Report (2010) mentions that around 40% of officials, who take bribe and are reported against, pay a fine.[15] Inspired by this evidence, Abbink et al. (2014) use 40% as detection probability. With these references in mind, we assign 40% probability of detection in the high probability with low fine treatment (HP, henceforth), and 20% probability of detection in the low probability with high fine treatment (LP, henceforth). Note that when moving from HP to LP, the detection probability is halved. Correspondingly, the fine is higher (lower) in LP (HP).[16]

[Table 1 here]

Table 1 describes the fine schedule for each bribe demand and the PO's expected payoff. As an illustration, assume that the PO demands a bribe of 400M and the C accepts. Note that if the PO is caught, the fine is 600M (100M) under LP (HP). As such, under LP, the PO earns (1000M – 600M) with probability 0.2 and (1000M + 400M) with probability 0.8; the expected payoff being 1200M. Similarly, under HP, the PO earns (1000M – 100M) with probability 0.4 and (1000M + 400M) with probability 0.6; the expected payoff being 1200M. Also, note that under any audit treatment and any bribe demand, the C is better off "accepting" than "rejecting" since the former may lead to the PO getting caught, in which case C's entitlement will be returned.[17,18] The above features of the experimental design captures the essence of harassment bribe in India.

---

[15] Being official data, this may actually be an overestimation of the actual rate of penalty.

[16] Table 1 makes it clear that it would have been logistically difficult for us to have designed a treatment with a detection probability of 10% (i.e., $p = 0.1$) or lower. Note that under $p = 0.2$, the fine for a $600M$ bribe demand is $900M$. As such, under $p = 0.1$, the fine for a $600M$ bribe demand would have exceeded $1000M$, thereby raising the possibility that a PO could end up with a negative payoff at the end of the experiment. Alternatively, it would have necessitated recalibration of the entire experiment.

[17] For instance, if $p = 0.2$, and $b = 400M$ (or, $600M$), a C is better off accepting the bribe demand for an expected payoff of 0.8*[1000M – 400M] + 0.2*1000M = 680M (or, 0.8*[1000M – 600M] + 0.2*1000M = 520M), than rejecting for a payoff of 400M. Thus, for a C, the strategy of "Accept" dominates "Reject".

[18] All treatments (Control, LP, and HP) share a common Nash equilibrium (NE): (Demand 600M, Accept), where the first (second) strategy in the parenthesis corresponds to the PO (Citizen). In addition, only for Control, (Demand 600M, Reject) is also a NE.

In our experiment, we do not directly measure risk attitude in an incentivized manner. Such incentivized risk preference measurement constitutes an experiment in itself and therefore, would have been logistically difficult to administer, given that our main experiment is quite long and significantly time consuming. Instead, we adopt a survey based measure of generalized risk from Dohmen et al. (2011), who experimentally validate the measure with incentivized choice. The survey question is the following: *"How do you see yourself: are you generally a person who is fully prepared to take risks or do you try to avoid taking risks? Please mention a whole number between 0 and 10, where the number 0 means 'not at all willing to take risks', and the number 10 means 'very willing to take risks'."* While response to this survey measure cannot be mapped to risk averse, risk neutral, and risk loving categories, a comparison of the underlying distribution of the survey based risk measure across the treatments can tell us if the observed treatment effects may be explained by the risk attitude. As an additional robustness check, we will examine if our results hold after eliminating potential risk loving and risk neutral agents, defined in terms of certain ranges of the risk attitude distribution, from our sample. More on this in section 4.4.

4.3 Experimental Procedure

Experiment 1 was conducted with undergraduate students at a large public university in India. A total of 304 subjects were recruited (38 pairs per treatment). Each subject participated in only one role of each treatment. Subjects were randomly assigned to treatments and roles. Two different rooms were used for the roles of the Cs and the POs. Since the students were well versed in English, the instructions were written (and read out) in English. Many examples, focused on possible earnings for a C and a PO, were discussed. Each experimental session lasted for approximately 1.5 hours. The participants' final earnings were converted from Mohors to

Indian Rupee, and were paid out in cash. The earnings ranged from Rs. 160 to Rs. 640 with an average of Rs. 373.07 (21.95 USD in PPP terms).

4.4 Results

[Figure 4]

Figure 4 provides a snapshot of the participants' decisions in Experiment 1. The average bribe demand is lower in any of the audit treatments, in comparison to Control. Similarly, the proportion of POs who demand bribe is lower in the audit treatments, in comparison to Control. For both measures, the difference between Control and LP is statistically significant. The lower panel of the figure shows that the percentage of the Cs who are willing to accept a bribe demand is a decreasing function of the amount demanded, irrespective of the treatment. In Control, if a PO demands 600M in return for approving the prize, the C will end up with the basic payoff of 400M, no matter whether she chooses to accept or reject the demand. To ensure that the PO does not earn any payoff from such bribe demand, it is likely that significantly more Cs will reject (than accept) in a tit for tat manner. This explains the rationale for why a very low percentage of the Cs are willing to accept a 600M bribe demand in Control.

Figure A1 in Appendix 2 depicts the distribution of bribe demand for each treatment. The distributions for Control and HP appear as closely similar. However, the distribution for LP seems different from the rest of the two distributions. LP witnesses a significantly higher percentage of POs approving the prize or demanding a bribe of 200M, in comparison to Control and HP.

The mean for variables of interest for each treatment and the treatment wise mean differences are reported in Tables 2A and 2B. Note that the Cs in LP and HP realize that if they accept a bribe demand, they may earn back their entire prize money because of the possibility of an audit.

Such strategic considerations are absent in the Control treatment. This explains the difference in acceptance rates between the Control and the audit treatments, particularly for higher bribe amounts.

For the POs, the possibilities of an audit with an associated fine are a deterrent to bribe demand. 15.79% POs choose to approve the prize in Control, as opposed to 42.11% and 23.68% in LP and HP, respectively. The difference between Control and LP is statistically significant (χ^2 test, p<0.01). On the intensive margin, the average bribe demand including the zeroes is 321M, 152M and 294M for the Control, LP and HP, respectively. The differences are statistically significant between Control and LP (t-test, p<0.01). Similar results hold if one considers average bribe demand excluding the zeroes. An interesting picture emerges from these analyzes which gives our first result.

Result 1A*:* In comparison to Control, LP has a significant effect on the likelihood of bribe demand and the amount demanded, however, HP has little and statistically insignificant effect on both fronts.

Result 1A suggests that Hypothesis 1A is only partly valid, i.e., possibilities of audit and fine in itself are not a sufficient deterrent, but possibility of audit characterized by high fine is.

The Cs expect a larger likelihood of honesty than is actually the case. However, the qualitative treatment effects that we observe among the POs are also valid for the Cs. They expect a higher proportion of the POs to not demand bribe in LP (55.26%) than either in Control (34.21%), or HP (36.84%). When it comes to bribe amount, once again, the Cs expect a smaller average bribe demand than is actually the case. However, qualitatively, the treatment effects in the Cs' belief are very similar to the treatment effects in actual bribe demand. The expected average bribe demand is smallest for LP (136M) and its difference from Control and HP are statistically significant (t-test, p=0.06 and t-test, p=0.03, respectively). The above discussion suggests that Hypothesis 1B is only partly valid and gives us our next result.

Result 1B: Possibility of audit and fine by itself is not sufficient to lower beliefs about how corrupt an environment is, but possibility of audit characterized by high fine is.

Do the POs behave in a manner consistent with the Beckerian framework in LP and HP? The proportion of the honest POs in LP and HP are 42.11% and 23.68%, respectively. The difference is statistically significant (χ^2 test, p=0.09). The average bribe demand including the zeroes (excluding the zeroes) are 152M and 294M (263M and 386M) in LP and HP, respectively; the difference between the treatments being statistically significant (t-test, p<0.01) in each case. This gives us one of the main results in this paper.

Result 2A: A comparison of the three different outcomes between LP and HP suggests that the LP treatment offers a more effective deterrent to corruption than the HP treatment.

Result 2A is consistent with the Beckerian framework, from which Hypothesis 2A is derived. While our results are consistent with the Beckerian framework, they may also be explained by alternative behavioral frameworks such as Prospect Theory, according to which individuals overweight small probabilities and underweight large probabilities. While our set up cannot rule out or identify alternative mechanisms at play, this is a promising line of future work.[19]

Our data further suggests that the Cs' belief about the proportion of honest POs in LP (55.26%) is greater than the same in HP (36.84%) (χ^2 test, p=0.11). Correspondingly, the Cs' believe that average bribe demand including the zeroes (excluding the zeroes) are 136M and 242M (305M and 383M) in LP and HP, respectively; the difference between the treatments is statistically significant when we include the zeroes (t-test, p=0.03), but not when we exclude the zeroes (t-test, p=0.13). This gives us our next result and it is consistent with Hypothesis 2B.

Result 2B: Citizens believe that POs are more likely to be honest in the LP treatment than in the HP treatment.

[19] We thank an anonymous referee for bringing this issue to our attention.

Our core results go through after controlling for demographic and other variables in a regression framework. Table A1 in Appendix 2 which reports the summary statistics of the explanatory variables also presents the definitions of these variables. We use these variables in the regression models presented in Table A2 in Appendix 2. The table presents marginal effects under multiple model specifications. An approval of the prize is treated as zero demand for bribe for each model. The OLS model (negative binomial model) treats bribe demand as a continuous variable (count variable: 0, 1, 2, 3). The logit model treats bribe demand as a binary variable (1, if demand > 0), and thereby focuses on the extensive margin. Across the model specifications, LP has a statistically significant deterrent effect on bribe demand, when compared to Control. At the intensive margin (OLS) [extensive margin (logit)], LP lowers bribe demand [probability of bribe demand] by approximately 168M [28% to 33%], in comparison to Control.

Among the other covariates, the variable "Risk" has a statistically significant positive impact on bribe demand. At the intensive (extensive) margin, a one-point increase in Risk leads to a 14.38M (3%) increase in bribe demand (probability of bribe demand). As pointed out earlier, Becker's theory predicts that low probability of detection with high fine is a greater deterrent than high probability with low fine when agents are risk averse. Given that our measure of Risk cannot identify a risk lover, risk neutral or risk averse agent, how do we know that our results are consistent with the prediction of Becker? To ensure that the results are consistent, we need to show, beyond any reasonable doubt, that: (a) our result is not driven by risk neutral or risk loving agents i.e. our result holds even if we exclude from the overall subject pool the potential risk loving and risk neutral POs, (b) the observed effect of LP on curbing bribe demand is not borne by a significant difference in the distribution of Risk across the treatments, and (c) overall, the regression models presented in Table A2 in Appendix 2 do not drastically change if we categorize Risk into fewer groups, instead of letting it enter the models on a (linear) 0 to 10 scale.[20] The next three paragraphs briefly address the above issues in the same order.

An emerging literature maps risk aversion in population in certain countries through representative samples. For instance, Dohmen et al. (2011) report that 9% of the Germans are risk loving. Harrison et al. (2007) note that "very few subjects are risk loving or risk neutral" and that "risk aversion is by far a better characterization of the risk preference of an average Dane." This is also consistent with Holt and Laury (2002, 2005). In the absence of a similar study in India, we do not have a strong prior on what proportion of our sample is risk averse. Instead, we conduct a set of robustness checks by restricting our sample to bottom 90, 80, 70, 60 and 50th percentiles of the Risk distribution, to make sure our results are not driven by risk loving or risk neutral agents. Table A3 in Appendix 2 presents the marginal effects from OLS regression results. Column (1) makes use of the full sample while subsequent columns use the restricted sample. Two main results emerge: First, the treatment effect of LP on bribe demand persists even as we restrict the sample. Second, as we exclude the top 20 percentile (column 4) and thereby potentially exclude the POs with risk neutral and risk loving preferences, the coefficient of Risk no longer turns out as statistically significant. Stated differently, this result is in accord with the extant literature discussed above: once we focus only on the risk averse POs, Risk does not influence bribe demand any longer, however, the treatment effect of LP remains significant.

Figure A2 in Appendix 2 presents a histogram of the POs' Risk distribution for each treatment. The mean of Risk considering all POs in all treatments in Experiment 1 is 5.96. Each histogram plots the relative frequency for four different ranges: 0–3 (low category), 4–6 (mild category), 7–8 (moderately high category), and 9–10 (high category). We also performed a Kolmogorov-Smirnov test for equality of the Risk distribution between pairs of treatments. The

[20] We thank an anonymous referee for bringing our attention to (b) and (c) above.

results are reported underneath the figure. We did not find any statistically significant difference between any two treatments.

Table A4 in Appendix 2 presents additional regression models in which Risk is categorized into fewer groups. Categorized Risk is defined as 1, 2, 3 and 4 for Risk range 0-3, 4-6, 7-8, and 9-10, respectively. Per expectation, for all models, the marginal effect of Categorized Risk emerges as statistically significant and positive. For instance, for the OLS model, a one unit upward movement in Categorized Risk induces a PO to increase bribe demand by 34.06M units. In addition, the coefficient of LP dummy continues to remain negative and statistically significant across all the models.

Finally, Table A5 in Appendix 2 presents regression models on the Cs' acceptance decision. The treatment features made it possible for us to make only one inference about a representative Cs' decision. Note that in the audit treatments, "Accept" is a strictly dominant strategy for any C over "Reject"; whereas in Control, "Accept" is a weakly dominant strategy over "Reject". As such, it is likely that the maximum bribe acceptable to the Citizens (MBAC, henceforth) will be higher on an average in the audit treatments than in Control. Indeed, this subtle difference between the treatments may serve as an explanation for the observed statistically significant and positive marginal effect for the audit treatment dummies, under the OLS and negative binomial models with full specifications. When we define the acceptance decision as a binary variable (which takes a value 0 if the citizen is unwilling to accept any bribe demand or 1 otherwise), we do not find a statistically significant treatment effect for any treatment. This gives us our next result.

Result 2C: The maximum bribe acceptable to the Citizens is higher in the audit treatments than that in the control treatment.

It is worthwhile to note here that the final objective of any policy is to reduce successful bribery. Now that may be achieved by lowering either bribe demand or bribe acceptance or both. So how do our treatments perform so far as this objective is concerned? Table 2A reports that the percentage of successful bribery decreased from 57.89% in Control to 44.74% in T1, which then increased to 52.63% in T2. However, none of the differences are statistically significant as per the *chi*-square tests.

5.1 Experiment 2: Experimental Design and Procedure

The aim of Experiment 2 was to examine if ethics education had an effect on unethical conduct, as measured by the control treatment of Experiment 1. In terms of equation (1), we expected to raise the intrinsic moral cost of demanding bribe ($m(b)$) through ethics education and thereby reduce the expected utility from demanding bribe. We conducted Experiment 2 with student subjects at a business school in India. Most MBA programs in India have a compulsory ethics course embedded in it. The aim of the course is to reinforce principled ethical conduct among future business professionals. Our experiment employed a between subject design and involved first year MBA students. Our choice of first year MBA students was based on two factors: first, they were a month into the program and having come from varied backgrounds, they had not yet imbibed the norms of the environment; and second, they had not yet been introduced to the compulsory course on ethics, which made treatment comparisons possible.

The ethics module was adapted from the regular Ethics course due to be taught later in the year. The module consisted of four lectures of two hours each, scheduled to be taught once each successive week. The topics of the four lectures were: Ethics, Morals and Values – An Overview (Lecture 1), Ethics and Business (Lecture 2), Ethical Conduct in Politics and Public Life (Lecture 3), and Ethics and Society (Lecture 4). The content was primarily drawn from the compulsory ethics course; the pedagogical details were decided by two regular instructors responsible for teaching ethics. Lectures 1 and 2 were delivered by one instructor and Lectures 3 and 4 were delivered by the other instructor across both the treatments. This approach ensured that there was no difference in the delivery of the material across treatments. Thus, the module was administered just as the regular course on ethics would be conducted. The lectures included a significant amount of discussion on corruption and its ills on developing countries like India.

Students did not receive credit for attending the module, but were actively encouraged to do so by the program administration. To minimize attrition and maximize participation, students were informed that they could participate in an event and earn prize money after the completion of the ethics module, but only if they had attended all the four lectures. We took special care to minimize experimenter demand effect – one, we did not tell the students upfront that the "event" was an experiment; two, experimenter was nowhere seen as the one either conducting or organizing the sessions on ethics module.

Per the admission policies of the university, students were admitted sequentially as per the merit list, however, the process of their assignment to any one of the sections is completely random. In fact, the administration takes special care in ensuring that students are randomized into sections, in order to avoid possibilities that one section goes academically ahead and another falls behind. Given that the students were randomized into sections only a month back and they were only a month into the program, we decided to randomize sections into treatments. This effectively meant that students were randomized into treatments and finally, we had three sections randomized into the three treatments. This was also logistically much more convenient for us to schedule the lectures and conduct the experimental sessions.

The baseline bribery game without the bells and whistles of probability of audit and fine was the ideal instrument through which we could measure the pure effect of ethics education. In Treatment 1 (T1, henceforth), students were introduced to the ethics module over four weeks. In the fifth week, they were invited to participate in the bribery game. In Treatment 2 (T2, henceforth) too, students were introduced to the ethics module over four weeks. However, they were invited to participate in the bribery game only in the eighth week, i.e., good four weeks after the module was completed. Thus, while T1 gave us an immediate measure, T2 gave us a delayed measure of ethics module on ethical conduct. In the control treatment (T0 henceforth), we did not introduce the ethics module to the subjects, but simply recorded their responses in the experimental game in the very first week in order to avoid possible contamination from T1 and T2. Spill over about the nature of the corruption game from T1 to T2 after the fifth week cannot entirely be ruled out, but it is not clear how that information may have affected behavior in T2. Anecdotally, subjects in a treatment were not aware that another group of students were participating in another treatment. Figure 5 provides an overview of the timeline for the treatments.

[Figure 5 here]

A total of 190 students from the MBA program participated in the three treatments of Experiment 2, with 30 pairs in T0, 33 in T1, and 32 in T2. The minimum and maximum amount paid were Rs. 160 and Rs. 640, with an average payoff of Rs. 295 (~ 17.35 USD in PPP terms).

5.2 Results

[Figure 6 here]

Figure 6 illustrates the main results and Tables 3A and 3B report the statistical tests on the amount of bribe demanded, acceptance rates and the beliefs. The average bribe demanded including the zeroes in T0, T1, and T2 are 233.33M, 200M, and 250M, respectively. The distributions of the bribe demand in each treatment are plotted in Figure A3 in Appendix 2. Though it appears that the average bribe in T1 is less than that in T0 and T2, neither of the differences are statistically significant. Similar results emerge when we exclude the zeroes from consideration. However, the percentage of honest POs increases from 20% in T0 to 39.4% in T1, the difference being statistically significant (χ^2 test, p=0.09). In T2, the honesty percentage decreases again to 18.8%, making the difference between T2 and T1 statistically significant (χ^2 test, p=0.07) again. Further, Table 3A reports that the percentage of successful bribery decreased from 60% in T0 to 36.36% in T1, which subsequently increased to 68.75%. The difference between T1 and T0 and that between T1 and T2 are statistically significant (χ^2 tests, p=0.06 and 0.01, respectively).

[Table 3A and 3B here]

Thus, the pattern indicates that there had been no change in the bribe amount across treatments – those who chose to be corrupt in T1 and T2 were just as corrupt as in T0. However, on the extensive margin, ethics module did have an impact on both likelihood of bribe demand and successful bribery; albeit for a short while. Thus, our next set of results are consistent with Hypothesis 3A and 3B from Section 2.

Result 3A: The proportion of POs who demand bribe are lower in T1 than in T0.

Result 3B: The proportion of POs who demand bribe are higher in T2 than in T1, indicating that the immediate effect of ethics education is greater than that in the longer run.

The actual acceptance rates of Cs decrease from T0 to T1, though the difference is not statistically significant. It increases from T1 to T2 and the differences both at bribe 200M and 400M are statistically significant (χ^2 test, p=0.08 and p=0.01, respectively). These results suggest that ethics education may have had some initial effect on the Cs' acceptance behavior, but once again the effects are short lived. One other way of seeing the Citizens' acceptance behavior is by comparing the average minimum bribe acceptable to the Citizens (MBAC) across the treatments. Table A7 reports the regression results – once again we see that MBAC is lower on an average in T1 than in T0, however, the differences are not statistically significant. This gives us our next result, which only partly corroborates Hypothesis 3C.

Result 3C: The actual acceptance rate of Citizens is statistically higher in T1 than in T2.

A further consequence of these outcomes may be seen in the proportion of successful bribery across the treatments. Table 3A shows that successful bribery in T1 is 23.64 percentage points lower than in T0 (χ^2 test, p-value=0.06), and 32.39 percentage points lower than in T2 (χ^2 test, p-value=0.01).

Our experiment is not equipped to tell us why time has a mitigating effect, if not an annulling one, on the effect of ethics education. Business students are extensively trained in Economics, Game Theory and other aspects of business studies, which may make them more self-regarding and hence, more corrupt (Frank and Schulze, 2000; Frank, Gilovich and Regan, 1993). They may have been through other life experiences which may have exposed them to bribery or introduced them to people outside their academic environment, where the ethical norm is likely to be different. All of the mechanisms above can potentially lead to a mitigating effect over time.

Interestingly, as Table 3B indicates, the POs also expect a smaller proportion of the Cs to accept bribe demand of 200M and 400M. While the difference between expected acceptance rate in T0 and T1 is not significantly different for bribe demand 200M, the difference is significantly different for bribe demand 400M (χ^2 test, p=0.03). The expected acceptance rates increase in T2 again for both bribe amount 200M and 400M, but the numbers are not statistically different from T1.

The pattern is similar when it comes to the Cs' belief about how corrupt the POs are. The Cs expect the POs to extract an average bribe of 293.3M, 242M and 318M in T0, T1 and T2, respectively. The difference between T1 and T2 is statistically significant (*t*-test, p=0.07). The Cs also expect more honest POs in T1 than in T0 and T2. The pattern in the Cs' data is similar to what we observed among the POs. The ethics module does have an effect, both in terms of how acceptable paying bribe is for the Cs and their expectations about how corrupt the POs are. However, the impact is again short lived. When measured after four weeks in T2, the effect vanishes. Further, the Cs do not believe that the POs will choose to be any less corrupt in terms of the bribe amount. They only expect a smaller proportion of the POs to demand bribe in T1, compared to T0. This impact too vanishes in T2, when measured four weeks after the intervention. These results are reinforced in the regression results reported in Table A6 in Appendix 2.

6. Comparison of Experiment 1 and Experiment 2

A comparison of results from Experiments 1 and 2 is tricky, as discussed earlier. The two experiments were conducted with two different subject pools, and therefore, the treatment means may not be readily comparable. In fact, the subject pool effects are readily observable with the average bribe demand including the zeroes being 321M and 233M in the baseline treatment of Experiment 1 and 2, respectively. As a result, if we are looking to compare the results of the two experiments, we may compare the normalized treatment effects by computing the percentage changes relative to the baseline treatment in each experiment. In Experiment 2, the likelihood of honest POs increased from T0 to T1 by 19.4 percentage points (χ^2 test, p=0.09), which is 97% of the number of honest POs in T0. The treatment difference between T2 and T0 is more modest and is statistically insignificant. In Experiment 1, the likelihood of honest POs in LP and HP is more than that in the baseline by 26.32 and 7.89 percentage points, respectively. These numbers translate to increases in 166.69% and 49.97% of the baseline number of honest POs in the Control treatment, for LP and HP, respectively. Only the difference between LP and Control is statistically significant.

For average bribe, we normalize the treatment effects by reporting the effects in terms of standard deviation changes from the baseline treatments. In Experiment 1, the average bribe demand excluding the zeroes (including the zeroes) in LP is 0.87 (0.95) standard deviation lesser than in Control, both the difference being statistically significant at 1% level. However, the corresponding differences in Experiment 2 are quantitatively smaller and statistically insignificant. This gives us our final result corresponding to Hypothesis 3D.

Result 3D: The effect of probability of detection and fine on PO's bribe demand behavior is greater than that of ethics module. Thus, monetary interventions have a far greater deterrence effect on bribe demand behavior than non-monetary interventions.

Further, norm changing interventions such as ethics education has little effect on the amount of bribe demanded, but they operate primarily through changing the likelihood of engaging in unethical act. Herberich, List and Price (2011) is the only study we know of that arrives at a similar conclusion – they find that normative appeals affect the extensive margin and price incentive based interventions work on both the intensive and extensive margins for green technology adoption. Future studies in this area can potentially uncover the underlying channels for why that is the case.

An important issue which is unraveled in course of the comparison between Experiment 1 and Experiment 2 relates to rate of successful bribery. In Experiment 1 we do not see a significant decrease in successful bribery in either LP or HP, primarily because monetary interventions induce citizens to accept bribes more often. However, in Experiment 2 we do find a significant decrease in successful bribery in the short run, primarily driven by Citizens' lower acceptance rate in T1. In quantitative terms, the reduction in successful bribery in LP is 22.72% of the successful bribery level in the Control treatment of Experiment 1. But the reduction in successful bribery in T1 is 39.4% of the successful bribery level in T0 of Experiment 2. Clearly, the ethics module has been more effective than monetary interventions so far as reducing overall bribery is concerned, though the effect disappears in the longer run. Once again it is evident that the intrinsic and extrinsic effects operate through different mechanisms – the latter can create perverse incentives, which can end up hurting the ultimate objective of the policy intervention[21].

7. Summary and Conclusion

In developing countries such as India, corruption is a significant impediment to development. Our paper uses experimental data to investigate alternative strategies to combat a class of corruption widely prevalent in many developing countries, namely, harassment bribery. Given that it is difficult to test some of the predictions of economic theory in the context of corruption due to the clandestine nature of the activity, laboratory based experimental games provide us with a useful tool to observe bribery behavior. In two experiments, we study the disincentive effects of raising extrinsic and intrinsic costs of indulging in corruption. In Experiment 1 we compare the two principle levers of anti-corruption interventions, namely probability of audit and amount of fine. In particular, we test the Beckerian prediction, which follows from expected utility framework, that low probability of detection and high fine is a greater deterrent than high probability of audit and low fine for the same level of expected payoff. Indeed, it is the low probability of detection and high fine that serves as a more effective deterrent both in terms of likelihood of bribe demand and bribe amount. In fact, in our experiment, high probability and low fine has no effect, whatsoever, on bribery behavior.

[21] We thank an anonymous referee for bringing this important issue to our attention.

Our results also throw light on a long standing paradoxical result in the literature. The fact that Abbink, Irlenbusch and Renner (2002) find introducing fine has an effect but Schulze and Frank (2003) do not, may now be explained by the following feature of their respective experimental design — fine was higher in the former while probability of audit was higher in the latter. Our result that fine has a greater deterrence than probability of audit can help explain why the results were contradictory in the first place.

Do our conclusions extend to other parameter combinations as well? We believe they do. It is possible that quantitatively the results may vary but qualitatively, a smaller probability of detection with higher fine will be a greater deterrent to corruption than higher probability of detection and smaller fine, ceteris paribus. Notice, this conclusion holds only for risk averse people and we believe that the perpetrators of the class of harassment bribery we studied are indeed characterized by risk aversion. However, our conclusions may not hold for big ticket corruptions and certainly will not hold for organized crime, where the perpetrators may have an entirely different risk profile. In fact, Becker's theory contends that risk lovers, will be deterred more by a higher probability of detection with lower fine. Our paper of course is not equipped to test that hypothesis.

Our second experiment aimed to raise the moral cost of indulging in corruption by subjecting participants to normative appeals through an ethics educational module. We find that ethics education has a small and significant effect in the likelihood of bribe demand but no effect on the bribe amount. Further, the effect is short lived and disappears in four weeks. A guarded comparative analysis of the extrinsic and intrinsic approaches to deter corruption suggests that the former is more effective in deterring bribe demand than the latter. A caveat in this regard: like in any other module driven descriptive intervention, it is hard to say what would have happened if, for the same time span, the treatment intensity of ethics module was increased. Our results indicate a direction and scope for a rich set of future studies in this context.

It is important to note that the two approaches operate through different underlying mechanisms. In general, the extrinsic disincentive based interventions may create perverse effects, which may act against the final objective of the intervention itself. In our experiment, it was rational on the part of the Citizens from a purely strategic sense to accept bribe demand more often in the audit treatments. This rendered the intervention ineffective from the standpoint of overall bribery. From this perspective, absence of such narrow strategic considerations may make normative appeals through ethics campaign relatively more effective. Thus overall, depending on the outcome one is interested in, the great emperor Ashoka may be thought of as both right and wrong.

Our results do not however imply that collective societal approaches aimed at raising intrinsic moral cost and thereby, changing the social norm governing corruption, is not useful. Our preferred interpretation of the result is that in order to make a lasting dent on corruption, ethics education should be administered at regular intervals. Of course, our experiment does not have the means to test such long term effects, which will remain as an important area for future work. Further, the intervention was aimed at people with an average age of 22.5 – an age where the moral compass of an individual has already been formed. Future research may identify the differential effects of normative appeals among children and young adults.

References

Abbink, Klaus, Bernd Irlenbusch, and Elke Renner (2002). An Experimental Bribery Game. Journal of Law, Economics, & Organization 18, no. 2: 428-54.

Abbink, K. et al. (2014). Letting the briber go free: an experiment on mitigating harassment bribes. Journal of Public Economics, 111, 17-28.

Abbink, K., & Serra, D. (2012). Anticorruption Policies: Lessons from the Lab. In Danila Serra, Leonard Watchekon (ed.) New advances in experimental research on corruption, 77-115. Emerald Group Publishing Limited.

Alm, J., McClelland, G. H., & Schulze, W. D. (1992). Why do people pay taxes? Journal of Public Economics, 48(1), 21-38.

Antes, A. L. et al. (2010). Evaluating the effects that existing instruction on responsible conduct of research has on ethical decision making. Academic Medicine: Journal of the Association of American Medical Colleges, 85(3), 519.

Armantier, O., & Boly, A. (2013). Comparing Corruption in the Laboratory and in the Field in Burkina Faso and in Canada. Economic Journal, 123(573), 1168-1187.

Banerjee, Ritwik, "On the interpretation of bribery in a laboratory corruption game: moral frames and social norms," *Experimental Economics*, 2015, pp. 1–28.

Banerjee, Ritwik. (2016). Corruption, norm violation and decay in social capital. Journal of Public Economics, 137, 14-27.

Bardhan, P. (2006). The economist's approach to the problem of corruption. World Development, 34(2), 341-348.

Barr and Serra (2009). The effects of externalities and framing on bribery in a petty corruption experiment. Experimental Economics 12(4) 488-503.

Basu, K. (2011). Why, for a Class of Bribes, the Act of Giving a Bribe should be Treated as Legal. Ministry of Finance, Government of India Working Paper No. 1/2011-DEA.

Becker, G. S. (1968). Crime and punishment: An economic approach. In The Economic Dimensions of Crime (pp. 13-68). Palgrave Macmillan UK.

Becker, G. S., & Stigler, G. J. (1974). Law enforcement, malfeasance, and compensation of enforcers. Journal of Legal Studies, 3(1), 1-18.

Björn Frank, Günther G Schulze, Does economics make citizens corrupt?, In Journal of Economic Behavior & Organization, Volume 43, Issue 1, 2000, Pages 101-113, ISSN 0167-2681, https://doi.org/10.1016/S0167-2681(00)00111-6.

Bruner, D. M. (2009). Changing the probability versus changing the reward. Experimental Economics, 12(4), 367-385.

Cason, Timothy N., Friesen, L., & Gangadharan, L. (2016). "Regulatory performance of audit tournaments and compliance observability," European Economic Review, Elsevier, vol. 85(C), pages 288-306.

Dohmen, T. et al. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. Journal of the European Economic Association, 9(3), 522-550.

Eide, E. (2000). Economics of criminal behavior. Encyclopedia of Law and Economics, 5, 345-389.

Faravelli, M. (2007). How context matters: A survey based experiment on distributive justice. Journal of Public Economics, 91(7), 1399-1422.

Fehr, E., Naef, M., & Schmidt, K. M. (2006). Inequality aversion, efficiency, and maximin preferences in simple distribution experiments: Comment. American Economic Review, 96(5), 1912-1917.

Fischer, C. M., Wartick, M., & Mark, M. M. (1992). Detection probability and taxpayer compliance: A review of the literature. Journal of Accounting Literature, 11, 1.

Robert H. Frank, Thomas Gilovich and Dennis T. Regan. Does studying Economics Inhibit Cooperation? *The Journal of Economic Perspectives* Vol. 7, No. 2 (Spring, 1993), pp. 159-171.

Frey, B. S., & Meier, S. (2003). Are political economists selfish and indoctrinated? Evidence from a natural experiment. Economic Inquiry, 41(3), 448-462.

Friesen, L. (2012). Certainty of punishment versus severity of punishment: An experimental investigation. Southern Economic Journal, 79(2), 399-421.

Grogger, J. (1991). Certainty vs. severity of punishment. Economic Inquiry, 29(2), 297-309.

Harrison, G. W., Lau, M. I., & Rutström, E. E. (2007). Estimating risk attitudes in Denmark: A field experiment. Scandinavian Journal of Economics, 109(2), 341-368.

Haucap, Justus, and Andrea Müller. (2014). Why Are Economists so Different? Nature, Nurture and Gender Effects in a Simple Trust Game, DICE Discussion Paper No. 136.

Herberich, D., List, J., & Price, M. (2011). How many economists does it take to change a light bulb? A natural field experiment on technology adoption. Working Paper, University of Chicago.

Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. American Economic Review, 92(5), 1644-1655.

Holt, C. A., & Laury, S. K. (2005). Risk aversion and incentive effects: New data without order effects. American Economic Review, 95(3), 902-904.

National Crime Records Bureau (2010). Crime in India: Statistics, Ministry of Home Affairs, New Delhi.

Konow, J. (2014). Can Economic Ethics be Taught. Discussion Paper, University of Kiel.

List, J. A. et al. (2001). Academic economists behaving badly? A survey on three areas of unethical behavior. Economic Inquiry, 39(1), 162-170.

Park, C. G., & Hyun, J. K. (2003). Examining the determinants of tax compliance by experimental data: A case of Korea. Journal of Policy Modeling, 25(8), 673-684.

Pommerehne, W. W., & Weck-Hannemann, H. (1996). Tax rates, tax administration and income tax evasion in Switzerland. Public Choice, 88(1-2), 161-170.

Rose-Ackerman, S. (1975). The economics of corruption. Journal of Public Economics, 4(2), 187-203.

Ryvkin, D., & Serra, D. (2012). How corruptible are you? Bribery under uncertainty. Journal of Economic Behavior & Organization, 81(2), 466-477.

Salmon, T. C. and Serra, D. (2017) Corruption, social judgment and culture: An experiment, In Journal of Economic Behavior & Organization, Volume 142, Pages 64-78.

Schulze, G. and B. Frank (2003) “Deterrence versus Intrinsic motivation: Experimental evidence on the determinants of corruptibility.” Economics of Governance 4, 143-160.

Shimshack, J. P., & Ward, M. B. (2005). Regulator reputation, enforcement, and environmental compliance. Journal of Environmental Economics and Management, 50(3), 519-540.

Shleifer, A., & Vishny, R. W. (1993). Corruption. Quarterly Journal of Economics, 108(3), 599-617.

Slemrod, J., Blumenthal, M., & Christian, C. (2001). Taxpayer response to an increased probability of audit: evidence from a controlled experiment in Minnesota. Journal of Public Economics, 79(3), 455-483.

Stafford, S. L. (2002). The effect of punishment on firm compliance with hazardous waste regulations. Journal of Environmental Economics and Management, 44(2), 290-308.

Torgler, B. (2005). Tax morale and direct democracy. European Journal of Political Economy, 21(2), 525-531.

Waples, E. P. et al. (2009). A meta-analytic investigation of business ethics instruction. Journal of Business Ethics, 87(1), 133-151.

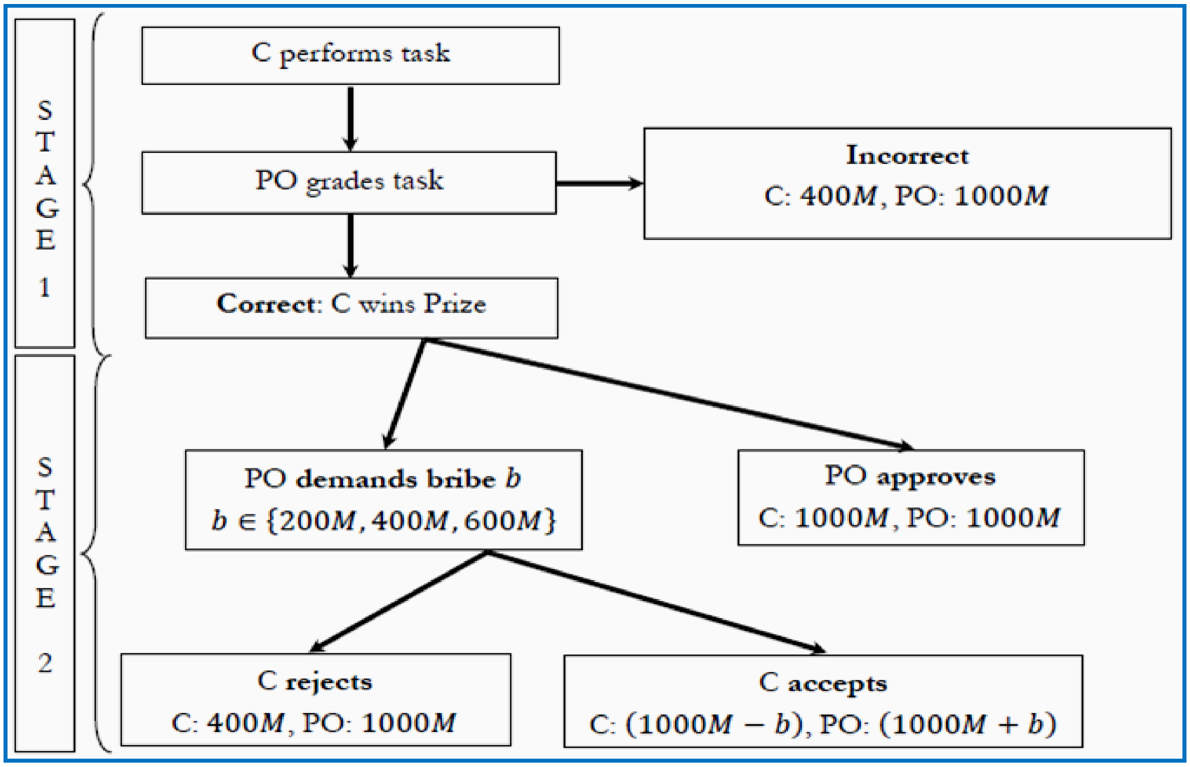


Figure 1: Control Treatment

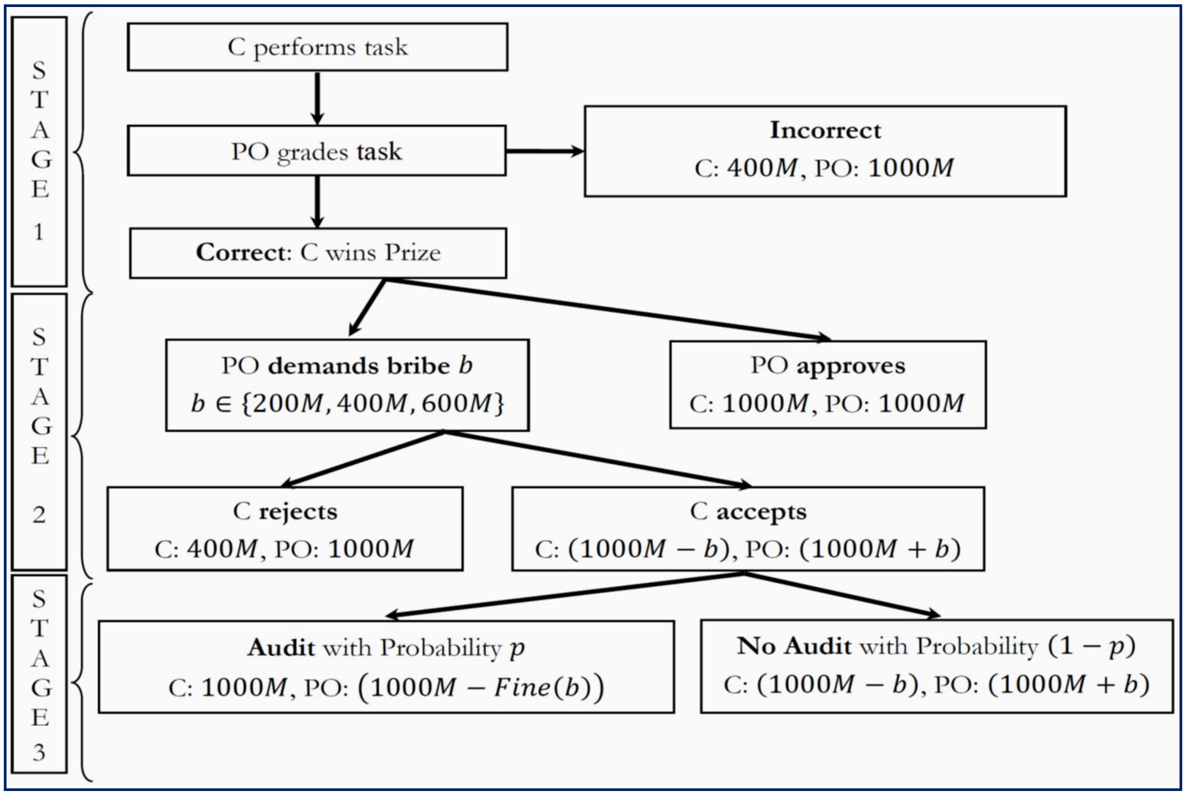


Figure 2: Audit Treatments

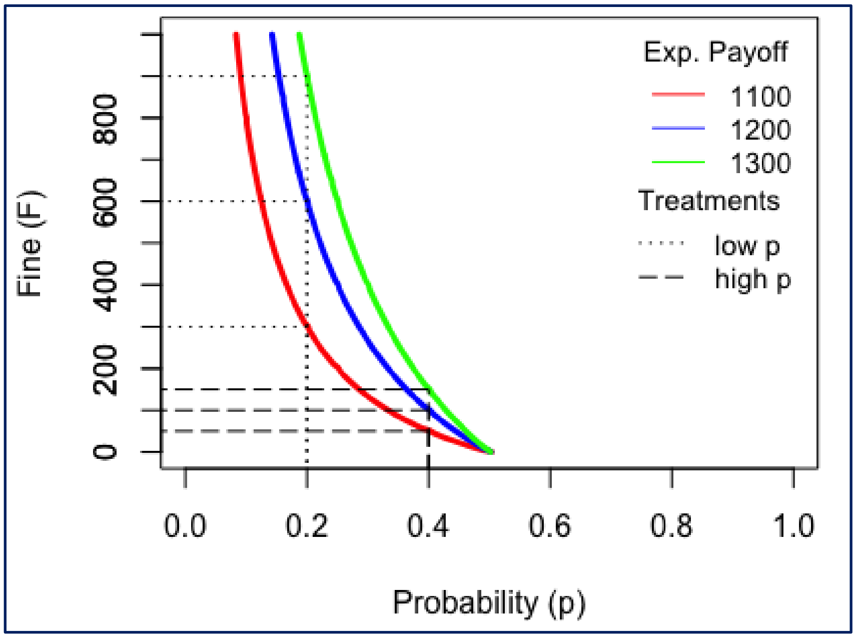


Figure 3: IEPC, Possible Bribe Demands, and Chosen Parameter Values

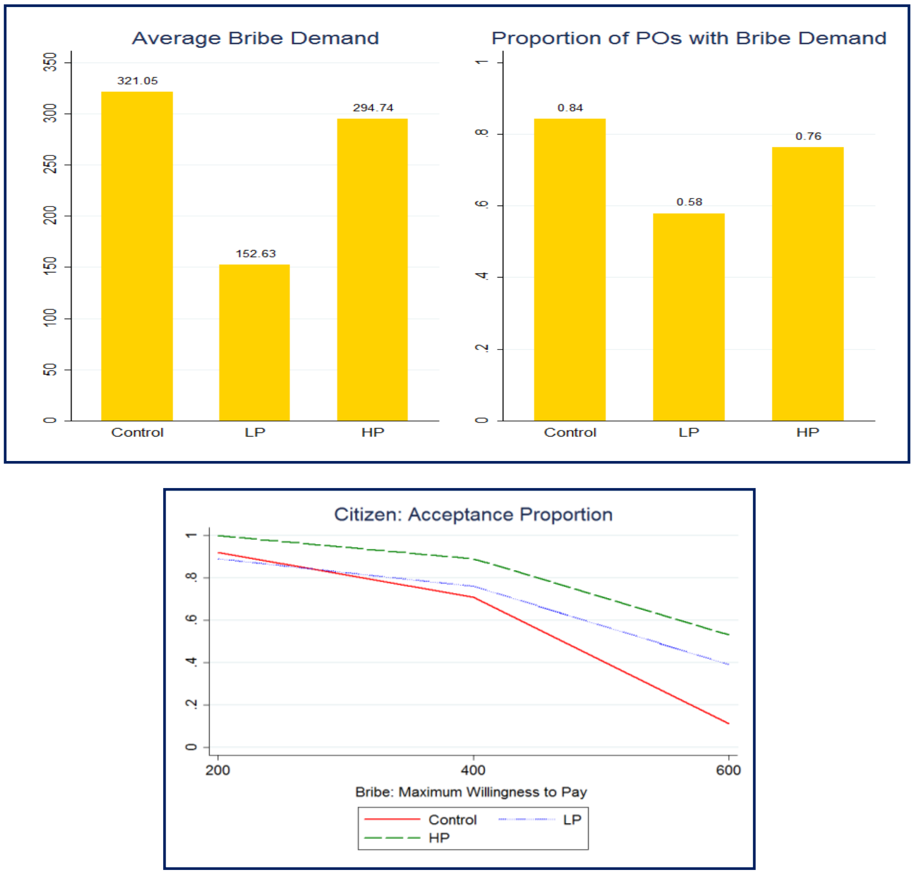


Figure 4: Average Bribe Demand, Proportion of Corrupt POs and Average Acceptance Rate in Experiment 1

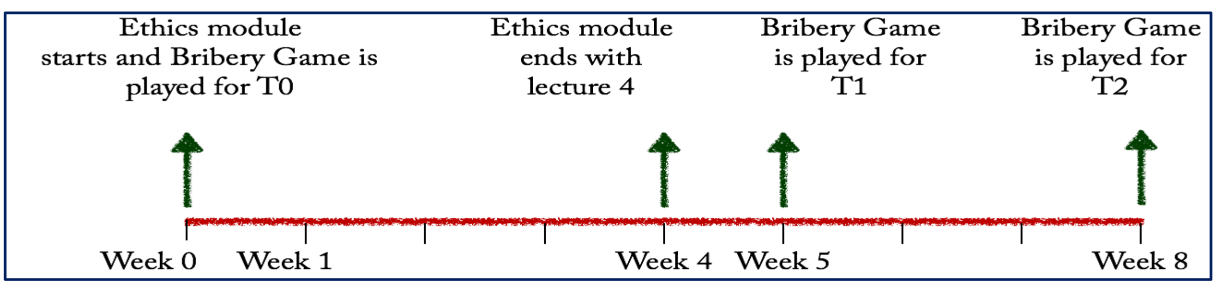


Figure 5: Timeline of Experiment 2

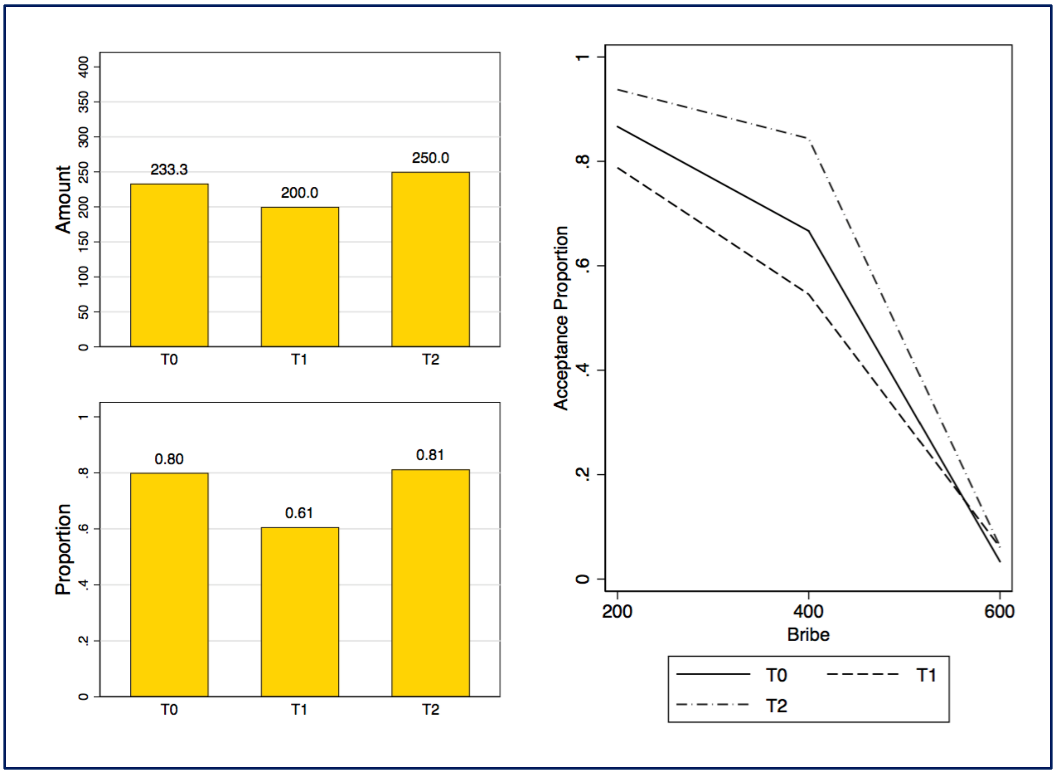


Figure 6: Average Bribe Demand, Proportion of Corrupt POs and Average Acceptance Rate in Experiment 2

Table 1: Fine Schedule and Expected Payoff for Public Official under Audit Treatments

Treatment / Bribe Amount	Low Probability = 0.2	High Probability = 0.4	PO's Expected Payoff
Demand 200M	300M	50M	1100M
Demand 400M	600M	100M	1200M
Demand 600M	900M	150M	1300M

Table 2A: Mean Differences in Decisions across Treatments in Experiment 1

	Control (CTR)	Low Prob (LP)	High Prob (HP)	Difference LP – CTR	Difference HP – CTR	Difference LP – HP
Citizens' Acceptance Rates						
Acceptance Rate for Bribe = 200M (%)	92.11%	89.47%	100.00%	–2.64 (0.69)	7.49* (0.08)	–10.53** (0.04)
Acceptance Rate for Bribe = 400M (%)	71.05%	76.32%	89.47%	5.27 (0.60)	18.42** (0.04)	–13.15 (0.13)
Acceptance Rate for Bribe = 600M (%)	10.53%	39.47%	52.63%	28.94*** (0.00)	42.10*** (0.00)	–13.16 (0.25)
Public Officials' Actual Bribe Demand						
Average Bribe Demand	321.05	152.63	294.74	–168.42*** (0.00)	–26.31 (0.57)	–142.11*** (0.00)
Average Bribe Demand > 0	381.25	263.64	386.21	–117.61*** (0.00)	4.96 (0.89)	–122.57*** (0.00)
Actual Bribe Demand = 0 (%)	15.79%	42.11%	23.68%	26.32*** (0.01)	7.89 (0.39)	18.43* (0.09)
Percentage of Successful Match Between Bribee (PO) and Briber (C)						
Percentage of Successful Bribery	57.89%	44.74%	52.63%	–13.15 (0.25)	–5.26 (0.65)	–7.89 (0.49)

Note: (Average Bribe Demand) [Average Bribe Demand > 0] indicates the average bribe demand (including the zeroes) [excluding the zeroes]. Actual Bribe Demand = 0 (%) indicates the percentage of POs with zero actual demand. For (Control) [Audit] treatment, "Successful Bribery" indicates those transactions in which PO demanded a positive amount of bribe that C agreed to pay, and (therefore the transactions ended successfully) [the transactions were not chosen for audit, which resulted in the transactions ending successfully]. When (averages) [proportions] are compared, bracketed numbers represent *p*-values from (*t*-tests) [*chi-square* tests].

Table 2B: Mean Differences in Beliefs across Treatments in Experiment 1

	Control (CTR)	Low Prob (LP)	High Prob (HP)	Difference LP – CTR	Difference HP – CTR	Difference LP – HP
Citizens' Belief about Public Officials' Bribe Demand						
Expected Bribe Demand	221.05	136.84	242.11	–84.21* (0.06)	21.05 (0.68)	–105.26** (0.03)
Expected Bribe Demand > 0	336.00	305.88	383.33	–30.12 (0.52)	47.33 (0.32)	–77.45 (0.13)
Expected Bribe Demand = 0 (%)	34.21%	55.26%	36.84%	21.05* (0.07)	2.63 (0.81)	18.42 (0.11)
Public Officials' Belief about Citizens' Acceptance Rate						
Acceptance Rate for Bribe = 200M (%)	97.37%	94.74%	97.37%	–2.63 (0.56)	0.00 (1.00)	–2.63 (0.56)
Acceptance Rate for Bribe = 400M (%)	76.32%	78.95%	92.11%	2.63 (0.78)	15.79* (0.06)	–13.16* (0.10)
Acceptance Rate for Bribe = 600M (%)	5.26%	39.47%	42.11%	34.21*** (0.00)	36.85*** (0.00)	–2.64 (0.82)
Second Order Belief: Public Officials' Belief about Citizens' Belief about PO's Bribe Demand						
Average of PO's Belief about Citizens' Expectation about PO's Demand	305.26	221.05	326.32	–84.21** (0.04)	21.05 (0.64)	–105.26*** (0.01)

Note: (Expected Bribe Demand) [Expected Bribe Demand > 0] indicates the average expected bribe demand (including the zeroes) [excluding the zeros]. Expected Bribe Demand = 0 (%) indicates the percentage of POs with expectation of zero bribe demand. When (averages) [proportions] are compared, bracketed numbers represent *p*-values from (*t*-tests) [*chi-square* tests].

Table 3A: Mean Differences in Decisions across Treatments in Experiment 2

	T0	T1	T2	Difference T1–T0	Difference T2–T0	Difference T2–T1
Citizens Acceptance Rates (%)						
Acceptance Rate for bribe = 200M	86.67	78.89	93.75	–7.78 (0.41)	7.08 (0.35)	14.86* (0.08)
Acceptance Rate for bribe = 400M	66.67	54.55	84.38	–12.12 (0.33)	17.71 (0.11)	29.83** (0.01)
Acceptance Rate for bribe = 600M	3.33	6.06	6.25	2.73 (0.61)	3.02 (0.59)	0.19 (0.98)
Public Officials' Actual Bribe Demand						
Average Bribe Demand	233.33	200	250	–33.33 (0.46)	16.67 (0.68)	50 (0.28)
Average Bribe Demand > 0	291.67	330	307	38.33 (0.32)	16.03 (0.63)	22.31 (0.59)
Actual Bribe Demand = 0 (%)	20%	39.4%	18.8%	19.4* (0.09)	1.2 (0.90)	20.6* (0.07)
Percentage of Successful Match Between Bribee (PO) and Briber (C)						
Percentage of Successful Bribery	60%	36.36%	68.75%	–23.64* (0.06)	8.75 (0.47)	32.39** (0.01)

Note: (Average Bribe Demand) [Average Bribe Demand > 0] indicates the average bribe demand (including the zeroes) [excluding the zeroes]. Actual Bribe Demand = 0 (%) indicates the percentage of POs with zero actual demand. When (averages) [proportions] are compared, bracketed numbers represent *p*-values from (*t*-tests) [*chi-square* tests]. We acknowledge that the number of POs for whom Actual Bribe Demand > 0 is small (i.e., 24, 20 and 26 for T0, T1 and T2, respectively) which can potentially invalidate some of the assumptions for a *t*-test. However, a non-parametric test of difference yields similar conclusions. (MW Ranksum Test, *p*-values of the differences are 0.48, 0.79 and 0.66, in the same order as presented in the table).

Table 3B: Mean Differences in Beliefs across Treatments in Experiment 2

	T0	T1	T2	Difference T1–T0	Difference T2–T0	Difference T2–T1
Citizens' Belief about Public Officials' Bribe Demand						
Expected Bribe Demand	293.3	242.4	318	–50.9 (0.25)	25.42 (0.53)	76.3* (0.07)
Expected Bribe Demand > 0	338.5	320	340	–18.5 (0.62)	–1.5 (0.96)	20 (0.59)
Expected bribe Demand = 0	13.3%	24.2%	6.3%	10.9 (0.27)	–7.1 (0.88)	17.9** (0.04)
Public Officials' Belief about Citizens' Acceptance Rate						
Acceptance Rate for bribe = 200M	96.67	84.85	93.75	11.82 (0.11)	–2.92 (0.59)	8.90 (0.25)
Acceptance Rate for bribe = 400M	80	54.55	68.75	–25.45** (0.03)	–11.25 (0.31)	14.20 (0.24)
Acceptance Rate for bribe = 600M	0	6.06	9.38	6.06 (0.17)	9.38* (0.09)	3.32 (0.62)
Second Order Belief: Public Officials' Belief about Citizens' Belief about PO's Bribe Demand						
Expectation of Citizens' Expectation	325.9	304	340.7	–21.93 (0.52)	14.82 (0.64)	36.74 (0.30)

Note: (Expected Bribe Demand) [Expected Bribe Demand > 0] indicates the average expected bribe demand (including the zeroes) [excluding the zeros]. Expected Bribe Demand = 0 (%) indicates the percentage of POs with expectation of zero bribe demand. When (averages) [proportions] are compared, bracketed numbers represent *p*-values from (*t*-tests) [*chi-square* tests].

Appendix 1

Here is a theoretical framework that is based on expected utility, along the lines of Becker. Suppose a PO is deliberating on how much bribe to demand. Consider an initial situation where the probability of audit and the fine are given by p and F, respectively. The expected payoff is given by

$$E(\pi) = p(w - F) + (1 - p)(w + b) = -pF + w + b - pb \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

In terms of our audit treatments, now assume that we move from HP to LP by virtue of an increase in F and a decrease in p, such that the PO stays on the same iso-expected-payoff curve as in equation (1). Totally differentiating equation (1), we obtain

$$d\big(E(\pi)\big) = 0 \Rightarrow -pdF - Fdp - bdp = 0 \Rightarrow -(F + b)dp = pdF$$

$$\Rightarrow \frac{dp}{dF} = -\frac{p}{(F + b)} \ldots\ldots\ldots\ldots\ldots\ldots (2)$$

The PO's expected utility is given by

$$E(U) = pU(w - F) + (1 - p)U(w + b) \ldots\ldots\ldots\ldots\ldots\ldots (3)$$

To analyze the impact of an increase in F (and a corresponding decrease in p so that $E(\pi)$ stays constant) on $E(U)$, we differentiate both sides of equation (3) with respect to F to obtain

$$\frac{dE(U)}{dF} = U(w - F)\frac{dp}{dF} + pU'(w - F)(-1) + U(w + b)\left(-\frac{dp}{dF}\right)$$

Using equation (2) in the above equation we obtain

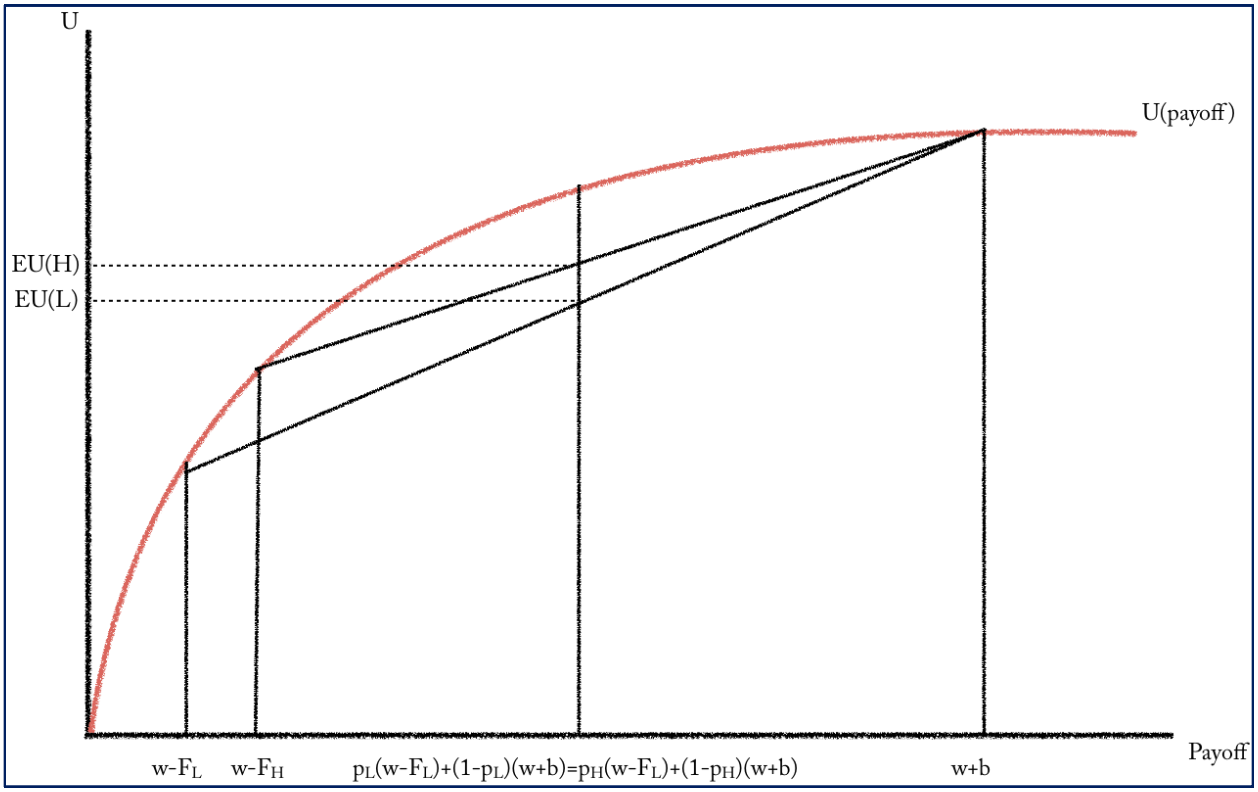
$$\frac{dE(U)}{dF} = -\frac{p}{(F + b)}U(w - F) - pU'(w - F) + \frac{p}{(F + b)}U(w + b)$$

$$\Rightarrow \frac{dE(U)}{dF} = -p\left[\underbrace{U'(w - F)}_{M} - \underbrace{\left\{\frac{U(w + b) - U(w - F)}{(w + b) - (w - F)}\right\}}_{A}\right] \ldots\ldots\ldots\ldots\ldots\ldots (4)$$

If the utility function is concave, the marginal rate of change in utility (given by 'M' in equation (4)) is greater than the average rate of change (given by 'A'). Stated differently, the condition $M > A$ holds only when tangents to a curve lie above the curve, implying $U'' < 0$, i.e., the PO is risk averse. As such, equation (4) indicates that $U'' < 0 \Leftrightarrow \frac{dE(U)}{dF} < 0$. For a risk averse PO, an increase in fine decreases his expected utility more than an increase in probability of audit, given that the expected payoff under the two situations remains unchanged. This is shown in the figure below: the same expected payoff leads to a higher expected utility under a situation of high probability and low fine than the alternative of low probability and high fine.

Finally, as discussed before, the above model does not shed any light on how the PO's expected utility maximizing bribe demand (b^*) will vary across the two policies (LP and HP). An answer to this question will require a more complicated model. However, a similar result (that shows that in comparison to HP, LP is more effective in reducing b^*) will emerge, even in the context of the more complicated model (under reasonable conditions).

Figure A1.1: Expected Utility, Risk Aversion and the Audit Treatments



Appendix 2

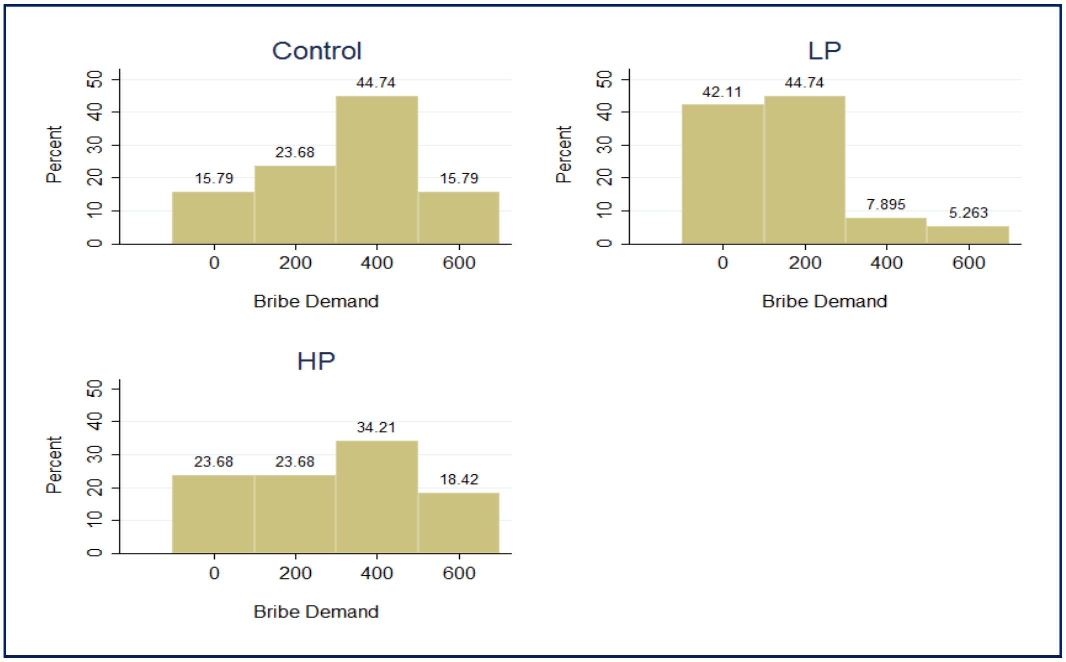


Figure A1: Empirical Distribution of Bribe Demand across Treatments in Experiment 1

Note: We reject the null hypothesis that the distribution of bribe demand between Control and LP are equal (Kolmogorov-Smirnov (KS) test, *p*-value=0.00) and that between LP and HP are equal (KS test, *p*-value=0.01). However, we fail to reject the null that the distribution of bribe demand between Control and HP are equal (KS test, *p*-value=1.00)

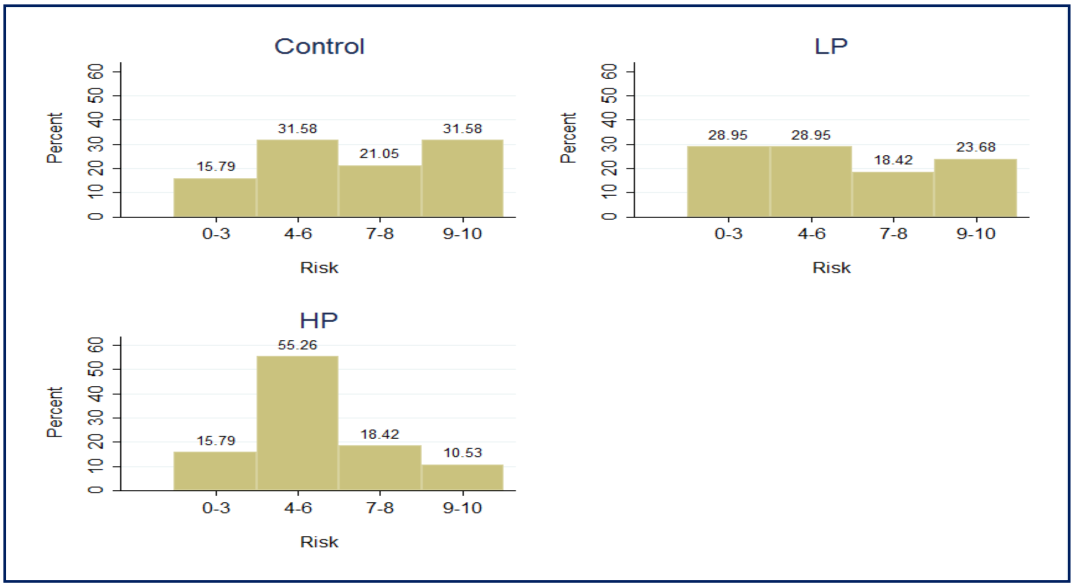


Figure A2: Empirical Distribution of Risk across Treatments in Experiment 1

Note: The figure corresponds to the POs, and not the Cs. For any given pair of treatments, we fail to reject the null hypothesis that the distribution of Risk between the treatments is equal. A KS test between the distributions for (Control and LP) {Control and HP} [LP and HP] yields a *p*-value of (0.79) {0.21} [0.79].

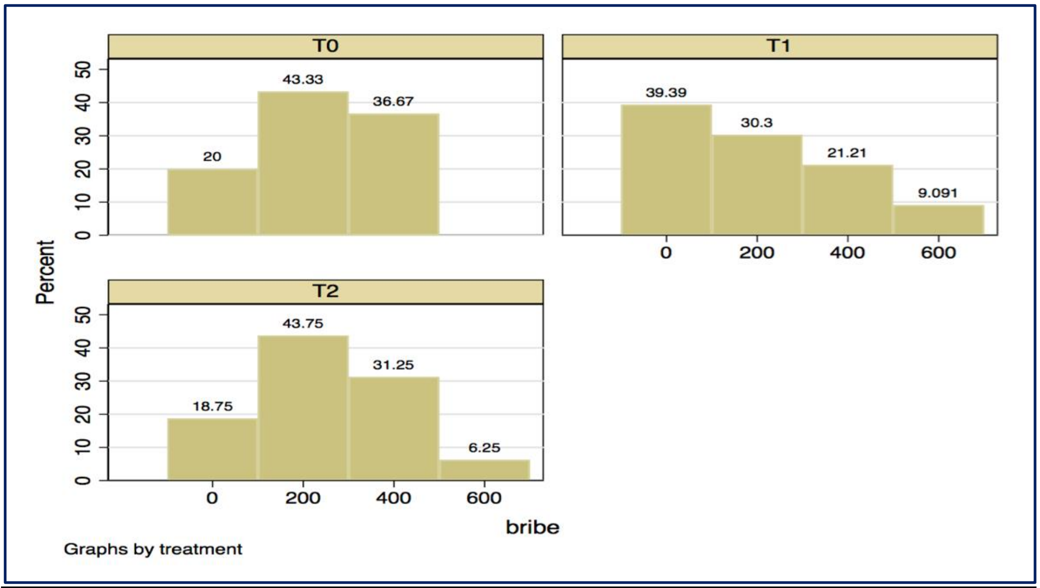


Figure A3: Empirical Distribution of Bribe Demand across Treatments in Experiment 2

Note: We fail to reject the null hypothesis that the distribution of bribe demand between (T1 and T0) {T2 and T0} [T1 and T2] are equal (KS test, *p*-value=0.60) {KS test, *p*-value=0.99} [KS test, *p*-value=0.49].

Table A1: Treatment-wise Summary Statistics for Explanatory Variables in Experiments 1 & 2

		Experiment 1		
Variables	Definition	Control	LP	HP
Female	Binary, 1 if female	0.34	0.37	0.55
Age	Age, in years	22.32	21.47	21.16
LnIncome	Natural log of annual income (in INR)	10.67	10.54	10.43
Score	Average of scores in Class X and XII (in %)	81.36	82.36	84.34
Risk	Response to the question on risk (0 to 10)	6.66	5.68	5.53
	Number of observations	38	38	38
		Experiment 2		
Variables	Definition (same as above)	T0	T1	T2
Female		0.46	0.44	0.44
Age		22.25	22.48	22.61
LnIncome		11.58	11.78	11.77
Score		80.07	86.99	82.39
Risk		6.66	5.91	6.39
	Number of observations	30	33	32

Note: The above table corresponds to data from public officials only.

Table A2: Regression Results for Treatment Marginal Effects in Experiment 1

Dependent variable: Bribe demand	OLS (Continuous)	Negative Binomial (Count)	Logit (Binary)	OLS (Continuous)	Negative Binomial (Count)	Logit (Binary)
LP Dummy	–168.42*** (40.60)	–0.82*** (0.18)	–0.28** (0.12)	–167.69*** (41.48)	–0.77*** (0.17)	–0.33*** (0.01)
HP Dummy	–26.32 (45.05)	–0.10 (0.18)	–0.10 (0.12)	–29.59 (46.24)	–0.09 (0.17)	–0.18 (0.14)
Risk				14.38*** (5.49)	0.08*** (0.03)	0.03** (0.01)
Female				19.12 (35.43)	0.09 (0.16)	0.08 (0.09)
Age				–17.09*** (5.45)	–0.09*** (0.03)	–0.03** (0.01)
LnIncome				3.76 (17.29)	0.004 (0.08)	–0.02 (0.05)
Score				–1.14 (3.07)	–0.01 (0.01)	0.01 (0.01)
Intercept	321.05*** (30.68)			652.31** (330.62)		
	$R^2 = 0.14$	$\chi^2 = 14.62$	$\chi^2 = 6.60$	$R^2 = 0.23$	$\chi^2 = 32.08$	$\chi^2 = 16.82$

Note: Control treatment is used as the base treatment. For each model, N = 114. Standard errors are in the parentheses. ***, **, and * indicate significance at 1% level, 5% level, and 10% level, respectively.

Table A3: Marginal Effects from OLS Regression for Varying Risk Values in Experiment 1

Condition	ALL OBS	≤ 90th Percentile	≤ 80th Percentile	≤ 70th Percentile	≤ 60th Percentile	≤ 50th Percentile
Dependent variable: Bribe demand	Column (1)	Column (2)	Column (3)	Column (4)	Column (5)	Column (6)
LP Dummy	–167.69***	–142.46***	–129.94***	–140.76***	–156.70***	–156.46**
	(41.48)	(44.35)	(45.46)	(52.12)	(51.57)	(62.21)
HP Dummy	–29.59	–26.22	–23.82	–20.92	27.09	33.55
	(46.24)	(49.98)	(53.09)	(54.85)	(59.33)	(75.51)
Risk	14.38***	19.22***	20.42***	14.61	10.50	–0.91
	(5.49)	(6.96)	(7.78)	(10.65)	(10.64)	(13.48)
Female	19.12	46.18	53.22	55.88	28.41	36.93
	(35.43)	(39.31)	(40.74)	(43.56)	(45.37)	(54.81)
Age	–17.09***	–19.78***	–20.19***	–22.40***	–32.63***	–30.87***
	(5.45)	(6.11)	(6.38)	(7.95)	(6.44)	(7.28)
LnIncome	3.76	16.74	24.58	40.47**	52.72**	59.67**
	(17.29)	(17.96)	(19.61)	(20.40)	(21.98)	(26.23)
Score	–1.14	–1.65	–1.67	–2.27	–5.63	–5.64
	(3.07)	(3.29)	(3.35)	(3.71)	(3.54)	(3.82)
Intercept	652.31**	574.76	489.91	440.10	822.75*	728.96
	(330.62)	(386.45)	(405.82)	(459.05)	(443.11)	(490.99)
Observations	114	94	89	80	67	47
R-squared	0.23	0.26	0.26	0.26	0.38	0.38

Note: "ALL OBS" stands for all observations. Each of subsequent models restricts the sample to bottom k^{th} percentile of the survey based Risk measure distribution, $k \in \{90,80,70,60,50\}$. Numbers in the parentheses represent standard errors. ***, **, and * indicate significance at 1% level, 5% level, and 10% level, respectively.

Table A4: Treatment Marginal Effects with Categorized Risk in Experiment 1

Dependent variable: Bribe demand	OLS (Continuous)	Negative Binomial (Count)	Logit (Binary)
LP Dummy	–170.90*** (41.30)	–0.79*** (0.17)	–0.33** (0.13)
HP Dummy	–30.58 (46.57)	–0.09 (0.17)	–0.18 (0.14)
Categorized Risk+	34.06** (15.31)	0.17** (0.07)	0.07* (0.04)
Gender (1, if Female)	20.89 (35.73)	0.10 (0.16)	0.08 (0.09)
Age	–17.31*** (5.49)	–0.09*** (0.03)	–0.03** (0.01)
Log of Monthly Family Income	4.75 (17.20)	0.01 (0.08)	–0.02 (0.04)
Average Exam Score	–1.28 (3.06)	–0.01 (0.01)	0.01 (0.01)
Intercept	662.60** (328.32)		
	$R^2 = 0.22$	Wald $\chi^2 = 30.03$	Wald $\chi^2 = 13.99$

Note: Control treatment is used as the base treatment. For each model, N = 114. Robust standard errors are in the parentheses. ***, **, and * indicate significance at 1% level, 5% level, and 10% level, respectively. + Categorized Risk is defined as 1, 2, 3 and 4 for risk range 0-3, 4-6, 7-8, and 9-10, respectively.

Table A5: Regression Results for Treatment Marginal Effects for Citizens in Experiment 1

Dependent variable: MBAC	OLS (Continuous)	Negative Binomial (Count)	Logit (Binary)	OLS (Continuous)	Negative Binomial (Count)	Logit (Binary)
LP Dummy	63.16 (40.36)	0.35 (0.23)	–0.03 (0.07)	67.98* (41.54)	0.37* (0.22)	–0.04 (0.06)
HP Dummy	136.84*** (33.15)	0.72*** (0.19)	.	121.26*** (39.51)	0.62*** (0.21)	.
Risk				–10.76** (4.92)	–0.06** (0.02)	–0.01 (0.01)
Female				–68.51** (30.07)	–0.33** (0.14)	–0.05 (0.05)
Age				5.05 (4.90)	0.03 (0.02)	0.01 (0.01)
LnIncome				33.77*** (9.55)	0.17*** (0.05)	0.03* (0.02)
Score				–2.45 (2.47)	–0.01 (0.01)	0.002 (0.003)
Intercept	347.37*** (24.66)			173.85 (247.42)		
N	114	114	76+	114	114	76+
	$R^2 = 0.11$	$\chi^2 = 16.36$	$\chi^2 = 0.15$	$R^2 = 0.26$	$\chi^2 = 44.26$	$\chi^2 = 6.47$

Note: "MBAC" stands for maximum bribe acceptable to the Citizens. Control treatment is used as the base treatment. Standard errors are in the parentheses. ***, **, and * indicate significance at 1% level, 5% level, and 10% level, respectively. + Each Citizen in HP was willing to pay a positive amount of bribe. In other words, the indicator variable for HP predicts the binary dependent variable perfectly. As such, all observations from HP were dropped from the regression results.

Table A6: Regression Results for Treatment Marginal Effects in Experiment 2

Dependent variable: Bribe demand	OLS (Continuous)	Negative Binomial (Count)	Logit (Binary)	OLS (Continuous)	Negative Binomial (Count)	Logit (Binary)
T1 Dummy	–33.33 (44.22)	–34.12 (43.99)	–0.19* (0.12)	–24.72 (46.48)	–26.04 (45.77)	–0.16# (0.11)
T2 Dummy	16.67 (40.37)	15.80 (38.30)	0.02 (0.12)	20.58 (42.33)	24.42 (39.62)	0.03 (0.12)
Risk				10.96 (12.38)	11.38 (11.96)	0.04 (0.03)
Female				17.15 (36.28)	12.18 (37.25)	0.03 (0.10)
Age				–1.51 (11.56)	–2.88 (10.85)	–0.01 (0.03)
LnIncome				–1.52 (15.40)	–3.94 (15.42)	0.01 (0.04)
Score				–0.70 (2.77)	–0.47 (2.50)	–0.00 (0.01)
Intercept	233.33*** (27.24)			258.36** (369.21)		
	$R^2 = 0.02$	$\chi^2 = 1.14$	$\chi^2 = 4.28$	$R^2 = 0.03$	$\chi^2 = 2.13$	$\chi^2 = 5.68$

Note: Control treatment is used as the base treatment. For each model, N = 95. Standard errors are in the parentheses. ***, **, * and # indicate significance at 1% level, 5% level, 10% level and 15% level, respectively. Given that we have only around 30 observations in each treatment, the full model for likelihood of bribe demand, including all the explanatory variables, does not have enough statistical power.

Table A7: Regression Results for Treatment Marginal Effects for Citizens in Experiment 2

Dependent variable: MBAC	OLS (Continuous)	Negative Binomial (Count)	Logit (Binary)	OLS (Continuous)	Negative Binomial (Count)	Logit (Binary)
LP Dummy	-47.27 (39.93)	-0.16 (0.14)	0.10 (0.09)	-42.41 (41.75)	-0.17 (0.16)	0.10 (0.08)
HP Dummy	48.75 (34.50)	0.14 (0.10)	-0.05 (0.08)	53.04 (34.73)	0.14 (0.11)	-0.04 (0.07)
Risk				–3.09 (7.24)	–0.02 (0.03)	0.02 (0.01)
Female				–13.38 (30.42)	–0.04 (0.11)	0.04 (0.05)
Age				-23.93** (11.24)	-0.08* (0.04)	0.02 (0.01)
LnIncome				-0.62 (14.96)	0.02 (0.05)	-0.01 (0.02)
Score				–0.09 (0.09)	-0.00 (0.00)	0.01 (<0.01)
Intercept	320.00*** (26.41)			892.58 (276.82)		
N	95	95	95	95	95	95
	$R^2 = 0.07$	$\chi^2 = 6.48$	$\chi^2 = 3.25$	$R^2 = 0.13$	$\chi^2 = 65.84$	$\chi^2 = 8.36$

Note: "MBAC" stands for maximum bribe acceptable to the Citizens. Control treatment is used as the base treatment. Standard errors are in the parentheses. ***, **, and * indicate significance at 1% level, 5% level, and 10% level, respectively.

Appendix 3: Supplementary material (for online publication only)

(The instructions for low-probability-high-fine treatment (LP) are given below. For the high-probability-low-fine treatment (HP), we replaced the LP parameters with the HP parameters (see Table 1). The instructions for control treatment (CTR) include Stages 1 and 2, but exclude Stage 3. For Experiment 2, only the instructions for the control treatment were used. Though we used the term fine in the main paper, when we ran the experiment we had used penalty in the Instruction.)

Instructions for All Participants

1.1 Introduction

Welcome to an experiment on decision making! If you read the following instructions carefully, you can, depending on your decisions and the decisions of other participants, earn a considerable amount of money. Communication with the other participants is strictly prohibited during the experiment. If you have any questions, please raise your hand, and we will come to you. This is an anonymous experiment and you will not know either the real identity of anyone else or the choices they make. To facilitate payments, we have assigned a unique identity number to each of you.

Earnings

During the experiment we will reward you in terms of a currency which we term as "Mohor". All Mohors you earn in the experiment will be converted into (Indian) Rupee at the end of the experiment. The conversion rate is: 1 Mohor = Rs 0.40. The amount of money you earn will depend on your decisions as well as on the decisions other participants make.

Overview

The experiment consists of an interactive situation between two participants and a survey. Please start by carefully reading the below instructions for the interactive situation. You will receive the instructions for the survey at the end.

1.2 The Interactive Situation

Each of you is randomly matched with another participant seated in another room. Between the two of you, one participant plays the role of a "Citizen", while the other participant plays the role of a "Public Official". The three stages are described below.

Stage 1

First, the citizen performs a task. He/She is asked to count the number of occurrences of the letter "A" from a random sequence of alphabets. For example, the number of occurrences of the letter “A” in the sequence "DEWAABKACCQAJ" is 4. Five sequences will be provided, and a citizen is asked to correctly count the number of occurrences of the letter “A” in at least three sequences. Each citizen will have five minutes to complete the task.

If a citizen is unable to complete the task in five minutes, then the experiment ends, and the citizen receives 400 Mohors for participation and the public official receives 1000 Mohors as salary.

On the contrary, if the citizen is able to count (correctly or incorrectly) the number of “A”s in at least three sequences in five minutes, then the citizen’s answers will be graded by the matched public official. If the public official determines that the citizen did not correctly solve at least three sequences, then the experiment ends, and the citizen receives 400 Mohors for participation and the public official receives 1000 Mohors as salary. On the contrary, if the public official determines that the citizen has correctly solved at least three sequences, then the citizen not only receives 400 Mohors for participation, but he/she wins an additional prize of 600 Mohors, and the experiment moves to **Stage 2**.

Stage 2

As indicated above, the experiment moves to Stage 2 when a public official determines that the matched citizen has won a prize of 600 Mohors, by correctly solving the at least three sequences. However, the citizen will receive the prize only when the public official *approves* the prize. The public official will now make *ONE* of two following choices:

(a) *Approve* the citizen's prize, OR

(b) *Demand a bribe* in return for approving the citizen's prize.

If the public official chooses to approve the citizen's prize (choice (a)), the experiment ends, in which case the citizen receives the prize of 600 Mohors ***plus*** the participation fee of 400 Mohors and the public official receives 1000 Mohors as salary. If a public official chooses to demand a bribe (choice (b)), then he/she must also indicate the amount of bribe he/she is asking for. The public official may ask for a bribe of either 200 Mohors, or 400 Mohors, or 600 Mohors.

If a public official chooses to demand a bribe in return for approving the citizen's prize (choice (b)), then the citizen will make *ONE* of two following choices:

(i) *Accept* the demand for bribe, OR

(ii) *Reject* the demand for bribe. prize.

If the citizen rejects (choice (ii)), then the citizen receives only the participation fee (and not the prize) and the public official receives 1000 Mohors as salary (and not the bribe). If the citizen accepts (choice (i)), then the game goes to **Stage 3**.

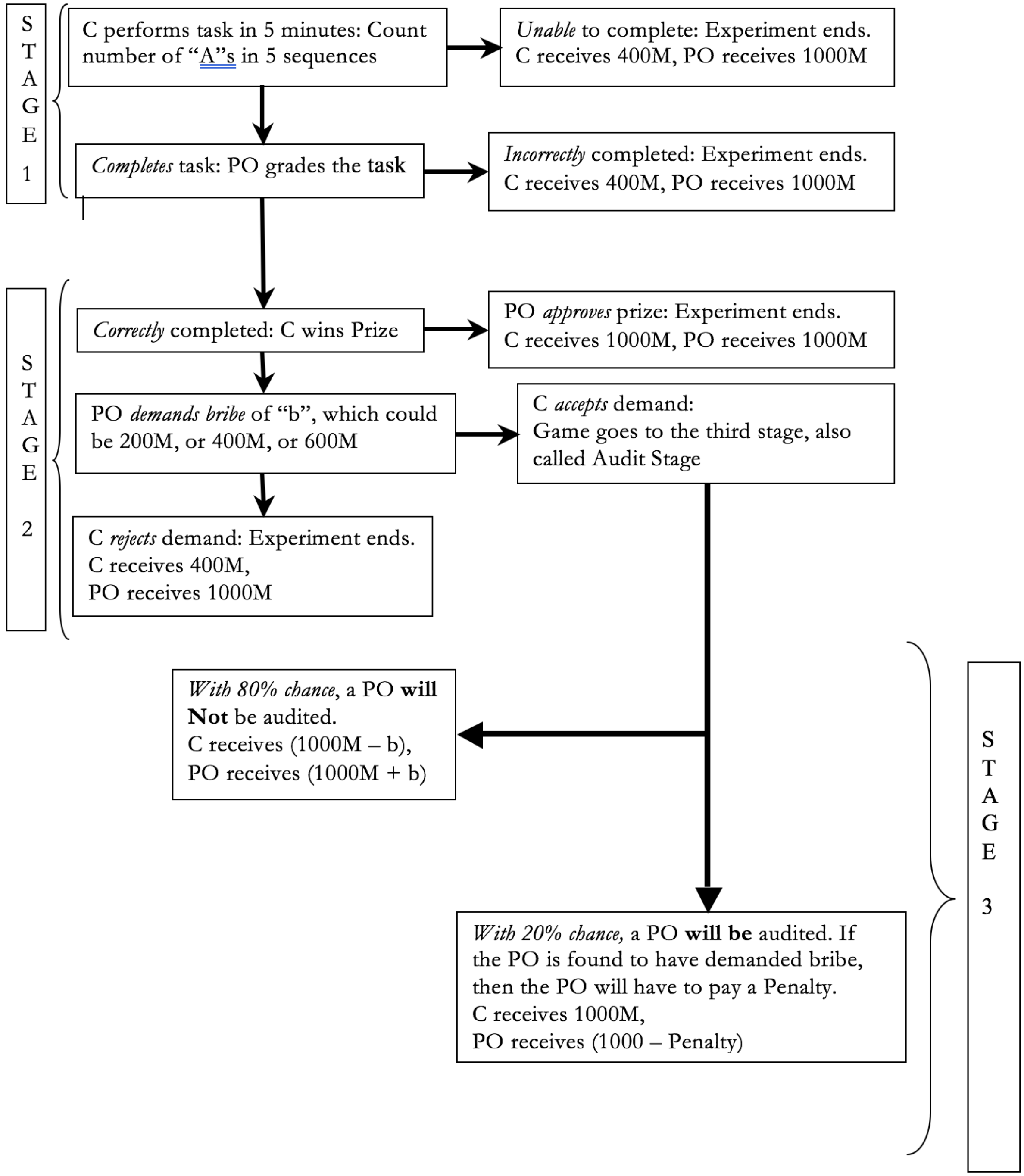
Stage 3

In Stage 3, which we call the Audit stage, there is a 20% chance that an audit will be conducted on a public official. Thus, the audit is randomly conducted and you do not know in advance who will be audited and who will not. If an audited person is found to demand bribe then a fine is imposed on him. Please take a look at the table below to see what the fine amounts are.

If a public official is caught having demanded a bribe (chance 20%) then a fine P is imposed on his earning. Public Official earns (1000 – Penalty) Mohors and Citizen earns 1000 Mohors. If he is not audited (chance 80%) then he earns (1000 + Bribe) Mohors and the citizen earns (1000 – Bribe) Mohors.

Bribe amount (in Mohors)	Chance of **NOT** facing an audit	Chance of facing an audit	Penalty (in Mohors)
200	80%	20%	300
400	80%	20%	600
600	80%	20%	900

The below chart depicts the stages. C = Citizen, PO = Public Official, M = Mohor



Instructions for Citizens

Registration Number: JU2016C

Each of you in this room has been assigned the role of a Citizen. Each of you has been matched with another participant (seated in another room), who has been assigned the role of a Public Official.

What you need to do

Please read the questions below and write your answers in the space provided. Your answers will be matched with the Public Official you are paired with and both of earnings will be determined.

Question 1 for Citizen (The Task): Please count, within five minutes, the number of times the letter "A" appears in the five sequences. Please record your answers in the blank spaces provided. Note that in order to win the prize of 600 Mohors, a citizen's answers must be correct for at least three sequences. You will be provided the task on a separate sheet.

Question 2 for Citizen: Read the example below. After that, fill in the blanks for 2(A)-2(C). To follow the example, quiz and subsequent earnings below is the table which shows the chances of facing the audit for you and the fine involved.

Bribe amount (in Mohors)	Chance of **NOT** facing an audit	Chance of facing an audit	Penalty for Public Official (in Mohors)
200	80%	20%	300
400	80%	20%	600
600	80%	20%	900

Example: Suppose you successfully complete the task and win the prize of 600 Mohors. However, the public official (PO) demands a bribe of 400 Mohors to approve the prize. You accept to pay the bribe. Now, at the audit stage, the PO is audited **(chance of which is 20%)** and he/she is caught having demanded bribe. In that case, you will receive (400 + 600) = 1000 Mohors. Since there is a 600 Mohors fine for demanding a bribe of 400 Mohors, the PO will receive 1000 – 600 (penalty amount from the above table) = 400 Mohors. On the contrary, if you reject the bribe demand, then you get 400 Mohors (Participation fee only) and the public official gets 1000 Mohors.

In the same manner, if the PO had demanded a bribe of 600 Mohors and he/she was caught in the audit stage, then he/she would have received (1000 – 900) = 100 Mohors.

Quiz: Suppose you successfully complete the task and win the prize of 600 Mohors. However, the public official demands a bribe of 600 Mohors to approve the prize. You accept to pay it.

A) At the audit stage the PO is not caught (chance of which is ________________%).

Your earning is ________________. The Public Official's earning is _____________.

B) Now suppose that at the audit stage PO is caught (chance of which is ________%).

The public official gets __________________. Your earning is _____________.

C) Suppose you reject to pay the bribe.

PO's earning is ____________________.

Your earning is ____________________.

<u>Question 3 for Citizen</u>: If you believe you have successfully completed the task, please indicate how much bribe you think will be demanded from you by the public official. Think carefully before mentioning what you think the public official will do. You will receive an additional 25 Mohors if your answer matches with what the public official's decision.

I think the public official will (Please **<u>underline</u>** any ONE of the four choices below):

- Approve the prize
- Demand a bribe of 200 Mohors
- Demand a bribe of 400 Mohors
- Demand a bribe of 600 Mohors

<u>Question 4 for Citizen</u>

If the public official indeed demands any of the bribe amounts (200 Mohors or 400 Mohors or 600 Mohors), then you need to decide whether you accept the demand or reject the demand. If you reject, you will receive 400 Mohors as participation fee and the public official will receive 1000 Mohors as salary. If you accept, you will receive (400 Mohors + 600 Mohors – bribe) and the public official will receive either (1000 Mohors + bribe) or (1000 Mohors – penalty), depending on whether the public official is caught during the audit or not. Please give your decision Accept or Reject for each bribe demand.

Amount of Bribe Demanded (in Mohors)	Your Earning if you *Accept* the Demand **<u>and</u>** the PO is not caught	Your decision (Write Accept or Reject)
200	1000 – 200 = 800	
400	1000 – 400 = 600	
600	1000 – 600 = 400	

If you are done, please submit this worksheet to the experimenter. We will now conduct the audit stage with Public Officials.

Instructions for Public Officials

Registration Number: JU2016P

Each of you in this room has been assigned the role of a Public Official. Each of you has been matched with another participant (seated in another room), who has been assigned the role of a Citizen.

What you need to do

Please read the questions below and write your answers in the space provided. Your answers will be matched with the answers of the citizen you are paired with and thereby your earnings will be determined.

Question 1 for Public Official: Read the example below. Then, fill in the blanks for 1(a)-1(d).

Example: Suppose the citizen successfully completes the task and win the prize of 600 Mohors. However, you (public official) demand a bribe of 400 Mohors to approve the prize. The citizen accepts to pay the bribe. An audit **(chance of which is 20%)** in stage three reveals that you have demanded a bribe. In that case, the citizen will receive (400 + 600) = 1000 Mohors. Since there is a 600 Mohors fine for demanding a bribe of 400 Mohors (consult the table below), you will receive 1000 – 600 (penalty amount from the table) = 400 Mohors. On the contrary, if the citizen rejects the bribe demand, the citizen receives 400 Mohors (Participation fee only) and you receive 1000 Mohors.

In the same manner, if you had demanded a bribe of 600 Mohors and the audit stage revealed that, then you would have received (1000 – 900) = 100 Mohors.

Quiz: Suppose the citizen successfully completes the task and win the prize of 600 Mohors. However, you demand a bribe of 600 Mohors to approve the prize. The citizen accepts to pay it.

D) At the audit stage you are not caught (chance of which is ______________%).

The citizen's earning is ______________. Your earning is _________________.

E) Now suppose audit reveals you demanded bribe (chance of which is ________%).

The citizen's earning is ______________. Your earning is _________________.

F) Suppose the citizen rejects to pay the bribe.

The citizen's earning is ______________. Your earning is _________________.

Instruction for Public Official: **Please respond to the questions below.**

Please take a look at the citizen's answers and compare those answers with the answer-key provided to you. Are the citizen's answers correct (for at least three sequences)? Please **underline** any ONE of the two choices:

YES NO

Question 2 for Public Official

If you have underlined "YES", then the citizen has won the prize. Would you now like to approve the prize or demand a bribe? Take a look at the table below and after that, please **underline** any ONE of the four choices that follow.

Bribe amount (in Mohors)	Chance of **NOT** facing an audit	Chance of facing an audit	Penalty for Public Official (in Mohors)
200	80%	20%	300
400	80%	20%	600
600	80%	20%	900

- Approve the prize
- Demand a bribe of 200 Mohors
- Demand a bribe of 400 Mohors
- Demand a bribe of 600 Mohors

Question 3 for Public Official

Please mention whether you think the citizen will accept/reject your bribe demand. You will receive an additional 25 Mohors if your answer matches with the citizen's decision (please put a **tick** mark against your choice).

- Suppose you demand a bribe of 200 Mohors: Citizen will Accept / Reject it.
- Suppose you demand a bribe of 400 Mohors: Citizen will Accept / Reject it.
- Suppose you demand a bribe of 600 Mohors: Citizen will Accept / Reject it.

Question 4 for Public Official: A few minutes back, you graded the citizen's task using the answer-key. Therefore, you now know whether the citizen has successfully completed the task. However, the citizen does not yet know whether he/she successfully completed the task. Assume that the citizen believes that he/she has completed the task successfully, i.e., the citizen believes that his/her answers for at least three sequences are correct.

If the citizen believes that he/she has successfully completed the task to win the prize, then he/she may start thinking whether you (the public official) will approve the prize or demand a bribe. Please indicate how much bribe the Citizen believes you will ask for. Think carefully before mentioning what you think the citizen believes you will do. You will receive an additional 25 Mohors if your answer matches with the citizen's belief.

I think the citizen believes that I will (Please **underline** any ONE of the four choices below):

Approve the prize	Demand a bribe of 200 Mohors	Demand a bribe of 400 Mohors	Demand a bribe of 600 Mohors

Audit Stage
At stage three, the chance that you will be audited is 20%. To figure out whether you will be audited or not, please write any number between 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10 in the space given below. After we collect our answer sheets, we will throw a ten sided dice. Chance of occurrence of any one number is 10%. We will throw the dice twice to randomly obtain two numbers. If your chosen number happens to be one of the two randomly chosen numbers then you will be audited. In this way we ensure that the chance of you being audited is 20%.

Question 5 for Public Official: My choice of number is ________________ (any number between 1 and 10, including 1 and 10). If you write any other number, your entire earnings will be taken away.

Please hand in the response sheets now. We will now randomly throw the ten sided dice to figure out whether you will be audited.