Economics of Misbehavior
Economics of Misbehavior

by Ritwik Banerjee

A PhD thesis submitted to School of Business and Social Sciences at Aarhus University in partial fulfillment of the requirements of the PhD degree in Economics and Business.

May 2014
Preface

Perhaps the most important global event of the period during which I conducted my research for the PhD thesis was the Purple Revolution, which started in, but was not limited to, the Middle East. The revolution was as much about democracy as about corruption and injustice. The Purple Revolution had its fair share of influence around the world - India was no exception. As political scams and scandals broke one after the other in India during the period, a major mass movement soon emerged, led by a platform which came to be known as India Against Corruption, with corruption as the single point agenda. This process of social churning was an interesting backdrop and provided much of the critical thoughts and questions that shaped my research for the thesis.

The thesis was written between September 2010 and May 2014 during my graduate studies in Economics at Aarhus University. Four years of dedication could not have, by itself, led to this fruition without inspirations from a number of people, perhaps more than I can acknowledge here. First, I thank the Department of Economics and Business at Aarhus University for providing me with an excellent research environment and support and for generously providing me with grants for conference, workshops, courses and experiments.

I will like to express my sincere gratitude to my thesis advisor Alexander Koch, whose guidance has been extremely valuable for me in the last four years. I have greatly enjoyed the discussions with him over the years - the most important takeaway from which has been the need to be precise and objective. I am also grateful to Torben M. Andersen, my secondary advisor, for creating the possibility for me to pursue PhD studies at Aarhus University and also for guiding me along the way.

I also take this opportunity to thank Elizabeth M. King and Peter Orazem for providing me with the opportunity to visit World Bank for the period between January 2013 and June 2013 and making my stay academically enriching. Thanks is specially due to Joydeep Bhattacharya and Nabanita Datta Gupta for their unstinting support and encouragement. I have always found them extending a helping hand in times of need. I also thank my friends in Aarhus for making my PhD studies an enjoyable experience and and those back in India for their encouragement and motivation.

Finally, I thank my parents - Tridibesh and Santana Banerjee, my sister - Rituparna and my wife - Ritika, for being sources of perennial support.

Ritwik Banerjee
May 2014
Updated Preface

The predefense was held on August 4, 2014. The assessment committee consists of Klaus Abbink, Monash University, Marco Piovesan, University of Copenhagen and Tor Eriksson, Aarhus University. I am thankful to the members of the committee for their careful reading of my dissertation and for their constructive comments and suggestions. I have incorporated some of them in the present version of the dissertation while others remain for future research.

Ritwik Banerjee
May, 2014
Contents

Summary

7

Chapter One
Student and Teacher Attendance: The Role of Shared Goods in Reducing Absenteeism 11

Chapter Two
On the Interpretation of Bribery in a Laboratory Corruption Game:
Moral Frames and Social Norms 37

Chapter Three
Corruption, Norm Violation and Decay in Social Capital 79

Chapter Four
Does bureaucracy attract more corrupt people? An experimental evidence from India 137

Chapter Five
Awareness programs and change in taste based caste bias 156
Summary

The cost of corruption is hard to overstate, especially in the context of developing countries. Not only does it impede growth and hinder poverty reduction by directly affecting economic activity, it also adversely affects social capital and thus social cohesion. Loosely defined as “abuse of entrusted power for private gain”, corruption thus remains an important issue that ought to be studied and understood, in order to devise effective policy measures to counter it. The following essays attempt to understand different facets of unethicality.

In the first essay titled “Student and Teacher Attendance: The Role of Shared Goods in Reducing Absenteeism” (joint work), we study the problem of teacher absenteeism in developing countries - something which is considered very unethical and greatly hampers student learning. We advance a theoretical model which demonstrates that, if teacher and student attendance generate a “shared good”, then teacher and student attendance will be mutually reinforcing. We use data from Northwest Frontier Province of Pakistan and provide empirical evidence supporting the proposition above. Controlling for the endogeneity of teacher and student attendance, we find that the most powerful factor raising teacher attendance is the attendance of the children in the school, and the most important factor influencing child attendance is the presence of the teacher. One policy implication which comes out of this is that in order to reduce teacher absenteeism, policy makers should find innovative ways by which the attendance of children can be raised. This also indicates that policies such as midday meals, which are seen as student specific interventions and are found to increase student attendance, may end up reducing teacher absence.

In the second, third and fourth essay, I study corruption through the lenses of behavioral economics using experimental data. Despite the fact that corruption is widely pervasive, the number of people who chooses not to be corrupt is astonishingly high. It is clear that the phenomenon of corruption (or lack of it) can hardly be explained by assuming that people maximize their narrowly defined self-interest. The replacement of \textit{homo-economicus} with \textit{homo-behavioralis} is thus the most plausible route to explain the observed patterns. Corruption has now been studied using experimental methods for some time now, mainly because data on corruption is otherwise not observable. However, past studies on laboratory corruption games have not been able to find consistent evidence that subjects make “immoral” decisions. A possible reason, and also a critique of laboratory corruption games, is that the experiment may fail to trigger the intended immorality frame in the minds of the participants, leading many to question the very \textit{raison d’être} of laboratory corruption games. To test this idea, in the second chapter, I compare behavior in a harassment bribery game with a strategically identical but neutrally framed ultimatum game. The results show
that fewer people, both as briber and bribee, engage in corruption in the bribery frame than in the alternative, suggesting that moral costs are indeed at work. To provide further support that the bribery game does measure moral costs, I elicit the shared perceptions of appropriateness of the actions or social norm, under the two frames. I show that the social norm governing the bribery game frame and ultimatum game frame are indeed different and that the perceived sense of social appropriateness plays a crucial role in determining the actual behavior in the two frames. This study, in a way, validates the use of experimental games as a tool for studying corruption.

Having examined the validity of experimental games, in the third chapter I investigate the causal link between corruption and trust and then analyze the mechanisms through which corruption affects social capital, using a lab based corruption game. In one treatment, subjects play harassment bribery game while in the control they play a strategically identical but differently framed ultimatum game (UG). Each treatment is followed by a trust game. Besides comparing the trust behavior, I also analyze the association between behavior in a trust game and response to the trust question in the World Value Survey. I find that not only do people trust less following the bribery game but that trust is also negatively related to the amount of bribe demanded. I also find that people who play the bribery game also tend to expect lower trustworthiness than their counterpart in the ultimatum game. This in turn explains why they trust less in the first place. We then turn to the link between the WVS-trust question and trust in experimental trust games and find that the former measures trust and expectations about other’s trustworthiness. However, the response to WVS-trust is stable while behavior in trust game, particularly the expectation component of the behavior, is susceptible to short term fluctuations.

In the fourth chapter (joint work), we study whether corrupt people self select themselves in professions where the scope of corruption is high. In this experimental study, corruption takes the form of embezzlement of resources and we compare corruption propensities among private sector job aspirants and among students aspiring to join Indian bureaucracy at the highest echelons. We find that at the intensive margin, aspirant civil servants indulge in corruption more than private sector aspirants. They underpay and also over report the number of tokens required more than their private sector counterparts. The findings suggest that, contrary to the widely held view, bureaucracy in India attracts more corrupt people, possibly due to the corrupt rent seeking potential therein.

In the fifth and the final chapter (joint work), we present some experimental evidence on the predictions offered by Becker (1971), that relative employment of the discriminated social group will improve if there is a decrease in the level of prejudice for the marginally discriminating employer, in the context of caste in India, with management students (potential employers in the near future) as subjects. First, we measure caste prejudice and show that awareness through a TV social program
reduces implicit prejudice against the lower caste and the reduction is sustained over time. Second, we find that the treatment reduces the prejudice levels of those in the left tail of the prejudice distribution - the group which can potentially affect real outcomes as predicted by the theory. And finally, we find that a larger share of the treatment group subjects exhibit favorable opinion about reservation in jobs for the lower caste.
Chapter One
Student and Teacher Attendance: The Role of Shared Goods in Reducing Absenteeism

(Published in *Economics of Education Review*)

Ritwik Banerjee*, Elizabeth M. King*, Peter F. Orazem*, and Elizabeth M. Paterno*

1Aarhus University, Denmark
2World Bank, USA
3Iowa State University, USA
4Ministry of Community and Social Services, Government of Ontario

Abstract

A theoretical model is advanced that demonstrates that, if teacher and student attendance generate a shared good, then teacher and student attendance will be mutually reinforcing. Using data from the Northwest Frontier Province of Pakistan, empirical evidence supporting that proposition is advanced. Controlling for the endogeneity of teacher and student attendance, the most powerful factor raising teacher attendance is the attendance of the children in the school, and the most important factor influencing child attendance is the presence of the teacher. The results suggest that one important avenue to be explored in developing policies to reduce teacher absenteeism is to focus on raising the attendance of children.

JEL: I21, I28

Keywords: Absenteeism; Teacher attendance; Student attendance; Shared good; Northwest Frontier Province; Pakistan

*ranerjee@econ.au.dk; Address: Department of Economics and Business, Fuglesangs Allé 4 8210 Århus V, Denmark. We are grateful for the help and advice of Harold Alderman, Mae Chu Chang, Guilherme Sedlacek, and the late Sar Khan in designing the data collection and to Rafiq Jaffer and members of the Northwest Province Education Management Information System for the implementation of the surveys. The opinions and conclusions expressed in this paper are those of its authors and do not necessarily reflect positions of the World Bank and its member governments.
1 Introduction

One of the most vexing obstacles on the path to universal literacy is teacher absenteeism in developing countries. Poor countries struggle to pay for enough teachers for their schools, but resources are wasted when teachers shirk their responsibilities. Teacher absenteeism rates have been found to be quite high in developed countries averaging 19% for primary school teachers (Chaudhury et al., 2006; Das et al., 2007). The 11% absenteeism rate in Peru is the lowest of the rates reported, but is still more than double the rates in developed countries. Furthermore, the national averages mask large variations within countries. In India where teachers are absent 25% of the time, rates are as high as 42% in the state of Jharkhand (Kremer et al., 2005).

One factor that seems to explain high teacher absenteeism is the lack of adequate supervision which makes it possible for teachers to shirk without penalty. This problem is particularly acute in rural areas where supervision might require significant travel, but even in cities teacher absenteeism can be a serious problem. Contributing to the problem are generous leave policies that allow teachers to miss 10% or more of the class days. Myriad other duties take teachers away from the classroom, including training, meetings with superiors, and administrative responsibilities. In the countries surveyed by (Chaudhury et al., 2004a), official leaves and work obligations accounted for between 25% and 86% of the teacher absences.

However, even in countries with generous leave policies, there is tremendous variation in attendance rates among teachers. Some will use all their allotted leave while others do not use the leaves they are allowed. A number of hypotheses have been advanced to explain this variation in teacher attendance, including pay, working conditions, opportunities for alternative employment outside teaching, and traveling distance to school. Especially among female teachers, family responsibilities can also lead to absences. In addition, inadequate supervision and monitoring insulate teachers from accountability for their performance (Majumdar, 2001). Studies show that while 35 out of 600 rural private schools in India reported incidence of teacher suspension due to absence or negligence, only 1 out of 3,000 teachers in rural public schools is suspended for the same reason (Kremer et al., 2005).

Econometric evidence of the determinants of teacher absenteeism in developing countries is limited, and research has failed to generate consistent results. (Chaudhury et al., 2004b, 2006) estimated teacher absenteeism regressions for six countries and found no variable to be consistently significant across the six regressions. Of the 22 variables (excluding the constant term and survey wave dummies) used in their multi-country regressions (Chaudhury et al., 2006), only six were statistically significant despite nearly 35,000 observations. Barely half of those variables had coefficients of the same sign in more than one-half of the countries. Even teacher salaries do not have consistent effects on teacher attendance, no doubt because salaries are set by civil service rules and not performance. Nevertheless, two results emerge consistently in the estimates. Monitoring, as measured by the school’s proximity to the Ministry of Education and the frequency of recent school
inspections, appears to raise attendance, as does having students with more educated parents.

Several countries have experimented with programs that improve school and teacher monitoring. In rural EDUCO schools in El Salvador, a community organization was contracted by the central education agency to be responsible for hiring and firing teachers and for closely monitoring their performance. Jimenez and Sawada (1999) found that, compared with non-EDUCO schools, teacher absences and student absences in EDUCO schools were lower. In the state of Rajasthan in India, schools that were required to provide a photograph of the teacher and students using a digital camera with a time/date feature improved attendance Duflo et al. (2012). The program resulted in an immediate improvement in teacher attendance. Teacher absenteeism was halved in the treatment schools, dropping from an average of 44% to 21%. Average time a child was exposed to a teacher increased by 30% and student performance increased by 0.17 standard deviations.

In Andhra Pradesh province in India, teachers on fixed-term, renewable contracts lacking the professional training normally required for a civil service posting were hired and randomly posted to 100 government schools. Pay was on par with private school salaries, but only one-fifth the pay of the civil service teachers. Nevertheless, attendance by contract teachers was 84%, significantly higher than the 73% attendance by civil service teachers Muralidharan and Sundararaman (2013). Duflo et al. (2011) report on a similar program, Kenya’s Extra Teacher Program which provided funds to hire additional teachers on a contract basis. Despite being paid 25% of the civil service rate, the contract teachers attended more frequently and their students performed better in comparison to the civil service teachers. Meanwhile, civil service teachers in schools that were allocated additional staff actually reduced their attendance from the already low rate of 58% to 45%!

The policy implication being drawn from these studies is that strong or high-stakes incentives such as pay-for-performance may be needed to discourage shirking. However, the Indian and Kenyan experience suggested that teachers on term contracts were the most responsive to such contracts. In fact, many of the teachers on temporary contracts in India and Kenya hoped to move to civil service contracts in the future, the teachers whose attendance did not improve.

Our study examines another avenue for improving teacher attendance—that teacher attendance depends on the attitudes and behaviors of students themselves, as reflected by student attendance. The intriguing possibility we pursue in this paper is that, even in the absence of close monitoring and extrinsic incentives, teachers attend because their students show up and students attend because their teachers show up. We couch this possibility in the context of a matching process between teachers and students in which both parties gain utility when both show up, but where both have outside options that also provide utility. In that context, teachers who believe their students are only weakly committed to attend will have more absences, and children who believe their teacher is prone to shirking will also shirk. In the Chaudhury et al. (2004b,a) analysis, for example, the attributes of the children and their parents were at least as important as the attributes of the

---

1Muralidharan and Sundararaman (2013) evaluated a second program that offered a modest bonus to teachers whose students improved on standardized tests. Test scores rose, but teacher attendance did not change.

14
teachers or the school. That result should not be surprising in that teacher attendance rates can vary dramatically within a country among teachers with standardized contracts, qualifications, curricula, and working conditions.

We illustrate the performance of this matching process using a unique data set composed of a representative sample of primary teachers and a random sample of their students collected in the North West Frontier Province of Pakistan in 1994-95. We find that by far the most important factor for teacher attendance is a higher probability that their students attend, and that the most important factor for child attendance is also a higher likelihood that the teacher will appear. The implication is that policies crafted to increase child attendance, such as those that conditional transfer programs that require a specified level of child attendance, will have the collateral benefit of raising the attendance of the teachers in local schools. Likewise, policies such as those that give the community some power to monitor teacher behavior in local schools will have the associated benefit of better student attendance.

The next section reviews some stylized facts about teacher and student attendance in Pakistan. The following section lays out a simple model that generates a sequence of equilibria where child and teacher attendance are positively correlated due to a match between the teacher and the child and/or his parents. The theory motivates an empirical strategy that is outlined in Section 4. Section 5 gives a detail description of the data while Section 6 lays out the results.

2 Teacher and Student Attendance in the Northwest Frontier Province

This study is based on information collected from a sample of 257 primary schools selected from a population of over 20,000 schools in NWFP, Pakistan. Sixty-eight percent of the selected schools were situated in rural areas, with school management divided between government schools (82%), mosque schools (4%), and private schools (14%), in proportion to their presence among the universe of schools. The data collection occurred during the spring of 1994. Because of the small number of mosque schools and their similar attendance patterns to government schools, we merge the mosque and government schools into one group.

Teacher attendance in random spot checks of the schools averaged 84% in the government and mosque schools and 93% in the private schools. The higher private teacher attendance occurs despite the fact that civil service salaries were twice that of private school teachers. Technically, government teachers would face salary cuts or disciplinary actions if they were absent 16% of the time. This is true even though leave policies in Pakistan are quite generous. Teachers can take as many as 25 days of casual leave per year without loss of pay.

Disciplinary actions against absent teachers are rare because the recorded teacher absences understate the true incidence. Table 1 lists the average absence rate recorded on the official school
registry during the months when the spot checks took place. The official rate is lower than the spot-check absenteeism rate in both public and private schools, but the gap is larger in government schools. In fact, average days of reported casual leave taken by civil service teachers is less than the casual days taken by private school teachers. None of the government teachers exceeded 25 days of casual leave which would have compromised their pay while almost 10% of the private school teachers passed the 25 day threshold\(^2\). Given the actual absenteeism rate for government teachers is over twice that private school teachers, the official registry almost certainly overstates how often teachers are present.

Corroborating evidence is found in the parents responses to a question asking whether their children’s schooling is harmed by absent teachers. The proportion of parents responding in the affirmative is 79% in private schools, but 93% in public schools. When teachers themselves were asked whether absent colleagues was a problem, 62% of civil service teachers responded affirmatively compared to 35% of private school teachers. These large differences would not occur given the lack of differences in official attendance unless the official attendance registry information was faulty.

The teachers do seem to take student attendance with more diligence than they exercise on the attendance of their colleagues. Spot-check and attendance registry data were quite consistent on average and school by school. The apparent reason is that there is an incentive not to report teacher absence as it can lead to income loss, but there is no reason to not to report student absences. As a result, we will use the registry data to measure student absence over the two month period that the 257 schools were being monitored, but we will use the spot-check attendance to measure teacher attendance.

Table 2 divides the teachers and students into high and low attenders by gender groups and then reports the means of various factors that might be thought to influence the attendance decision. For teachers, there are only modest differences in the means between attending and nonattending male teachers, but there are surprisingly large differences in these factors for female teachers given the groups either attended or did not attend school on a randomly chosen day. For male teachers, the main factor associated with higher attendance was whether the teacher had supervisory responsibilities or other official duties in the school outside the classroom. For female teachers, however, there were numerous factors that were correlated with spot-check attendance. Women teachers were less likely to attend if they were married, had more children and had more children of preschool age. Consistent with the finding that civil service teachers attend less regularly elsewhere, the average salary of nonattending female teachers is higher than of attending teachers, but the nonattending teachers were also disproportionately in government schools. Women teachers attended more regularly in schools with more teachers. But the key stylized facts that motivates our study is that the only factor that is consistently correlated with both male and female teacher attendance is the attendance of the children in the classroom: 32 percentage points higher child

\(^2\)There are numerous other legal justifications for teacher absences including official functions, illness, maternity leave, training, and earned leave based on years of service.
attendance when male teachers attend and 24 percentage point higher attendance when female teachers attend.\(^3\)

We have more variation in child attendance because the measure is based on two months of attendance registry data. When we compare children at the upper and lower tails of attendance, there are no household attributes that significantly affect the attendance of girls. Differences are more pronounced for boys who attend more regularly if their fathers are present and more educated, if their family is wealthier, if they have fewer siblings, the school is private and/or has better infrastructure, and is one kilometer farther from home. Only the last result is inconsistent with typical results. However, the only consistent difference for both boys and girls is that both attended more frequently in classrooms where their teacher attended more frequently: 37 percentage point higher teacher attendance for attending boys and 15 percentage point higher teacher attendance for attending girls.

These stylized facts suggest that teacher and student attendance are closely integrated. To model the teacher-student relationship, the next section draws on the framework Becker (1973) and Becker (1974a) advanced to model marriage as well as the extensions by Manser and Brown (1980) and McElroy and Horney (1981). As with a marriage, teachers and students cooperate to produce a “shared good” which raises utility for both parties.\(^4\) The key feature of a shared good is that it cannot be produced without the participation of both teachers and students. In the classroom context, examples of a shared good are the satisfaction from the learning that takes place, a productive mentor-mentee relationship, and the status conferred by the community to teachers and students of a well-run school.\(^5\) As we will see, the existence of a shared good explains why teacher and student attendance are jointly determined, leading to interesting policy prescriptions that could lead to improved teacher performance and student outcomes in developing countries.

### 3 Model

The economic models that have been applied to frame issues regarding teachers include those that treat teachers’ educational background and years of experience as inputs into a production function for student achievement;\(^6\) those that describe the accountability relationship between the teacher and the community (made up of local parents and their children and/or their representatives in

---

\(^3\)Note that it is routine in multi-teacher schools that present teachers take over the classes of absent teachers as well as teaching their own classes, and so children still get some form of schooling services if their own teacher is absent, and so it is not automatically the case that absent teachers mean that the children will be absent as well.

\(^4\)The concept of a “shared good” is different from that of a “public good” in the sense when a number of households are created through corresponding matches in the marriage market, a “public good” may be enjoyed by multiple households whereas a “shared good” is enjoyed only within a household.


\(^6\)Glewwe (2002) reviews the educational production literature.
school administration) and those that focus on the labor supply behavior of teachers as being one of income optimization constrained by the disutility from effort exerted. While the last model is frequently used to examine why individuals enter or remain in the teaching profession, the model applies more generally to teacher choices regarding how much effort to expend.

One aspect of being a teacher that none of these models emphasizes is that teachers may derive utility from their professional practice beyond the salary they receive. Certainly many professions offer hedonic returns in the form of pride of accomplishment or prestige among one’s peers, but a unique aspect of being a teacher is that the process of instruction and its ultimate output is intrinsically a product of the joint efforts by the teacher and by the student. We stress this feature of the production process because the concerns about frequent teacher absences and of irregular student attendance is underpinned by the belief that the teaching-learning process requires the direct interaction between teacher and student. Thus, when either teacher or student is absent, learning production suffers. Furthermore, the frequency and duration of this interaction contributes to learning because it allows the two parties to know each other. A teacher doesn’t start off knowing a student’s endowments (e.g., ability to learn, ability to focus, persistence), and so it may take repeated interactions to understand how best to teach a class. Similarly, a student may need a few classes to get to know and adapt to a teacher’s style of teaching. This way of framing the teacher-student relationship recognizes that interaction, cooperation and mutual adjustment are all aspects of the teaching-learning process.

To capture this joint and cooperative production of learning, we note that teachers and students are two parties in a mutually beneficial contractual relationship. Let $l_T$ and $l_C$ be the hours the teacher ($T$) and the child ($C$) attend such that $l_T, l_C \in [0, 1]$. The teacher or student shirk by their absence, setting $l_T < 1$ or $l_C < 1$. We assume that the shared good is produced using a constant elasticity of substitution production process using student and teacher time. The amount of the shared good, $G$, produced is

---

7Podgursky and Springer (2007) present a review of compensation options to resolve the principal agent problem in education.

8Two early examples of empirical models of teacher labor supply include Theobald (1990) and Stinebrickner (1998).

9There are other interpretations of the joint good. Yamamura (2011) points to a common attitude of trust in others as inducing a lower dropout rate in Japan. The distinction is that our good is produced by the education process whereas the good already exists in the Yamamura model.

10A referee pointed out that the productivity of teacher time will depend on the heterogeneity of abilities and skills in the students. Letting $\sigma$ represent the variance of learning potential among children, we can redefine the coefficient governing the marginal product of teacher time as $\beta_T = \beta(H_T, \sigma)$ where $H_T$ represents the teacher’s abilities which positively affects teacher productivity and $\sigma$ negatively influences the productivity of the teacher’s time. The effect of rising heterogeneity in student learning competencies can be evaluated as the impact of a smaller $\beta_T$. In empirical applications, one could incorporate student heterogeneity by using a random coefficients framework or by incorporating terms in $l_T^\rho$ and $l_T^\rho * \sigma$ into the specification for the shared good $G$. We do not emphasize the role of student heterogeneity in our current application as our grade K-2 children will not vary much in ability, but it is useful to recognize how such heterogeneity affects school management. Lower teacher productivity in the face of student heterogeneity in ability explains why schools are organized in grades with children divided into relatively homogeneous skill groups defined by expected competencies that may be further subdivided by tracking Duflo et
\[ G = \left( \beta_T l_T^p + \beta_C l_C^p \right)^{1/\rho} \]

**Teacher’s attendance decision**

The teacher’s total utility is derived from the shared good, \( U(G) \), and from goods purchased with income derived from teaching and other activities. A nonshirking teacher will be paid \( W_T^{NS} = w_T \).

For simplicity, we assume that an observed absence results in forfeiture of income for the period the teacher is absent\(^{11} \). On the other hand, time away from school has value of \( w_0T \) to the teacher, either because the teacher has an alternative job or because the teacher values time in home production. We assume that for the \((1 - l_T)\) share of time the teacher is absent, s/he earns \((1 - l_T)w_0T\) working away from school. Consequently, if the teacher shirks and is caught, wage income is \( W_T^{SC} = w_T l_T + w_0T(1 - l_T) \).

As is often the case in developing countries, teachers’ attendance is monitored with error. In section 2, we found that the official attendance register understates the incidence of absenteeism and so only some of teacher shirking is observed and subject to official punishment. On the other hand, one teacher’s absence imposes costs on fellow teachers who have to monitor their own classes and those of the absent teachers. That suggests that teachers may face informal sanctions by their colleagues if they attend too infrequently, even if they are not subjected to official penalties from their supervisors. To address these elements, we specify a parameter \( 0 \leq \alpha \leq 1 \) that indicates alternatively the probability the teacher is caught shirking by supervisors or the discount the shirking teacher applies to his earnings due to the expected sanctions imposed by his/her peers\(^{12} \).

If the teacher shirks and is not caught or does not face peer sanctions, his income is \( W_T^{SNC} = w_T + w_0T(1 - l_T) \). Consequently, using \( \alpha \) as the expected proportional penalty from shirking, the shirking teacher’s expected income will be \( W_T^S = (1 - \alpha(1 - l_T))w_T + w_0T(1 - l_T) \). The nonshirking wage is a special case of this formulation, as setting \( l_T = 1 \) implies that the expected income is \( w_T \).

Assuming the child’s time in school is given, the teacher’s decision is to select \( l_T \) so as to maximize expected utility,

\[
\max_{l_T} EU_T = (1 - \alpha(1 - l_T))w_T + w_0T(1 - l_T) + U \left( \left( \beta_T l_T^p + \beta_C l_C^p \right)^{1/\rho} \right)
\]

\(^{11}\)As noted above, teachers do not forfeit pay until their casual leave exceeds 25 days and they may qualify for other leaves for illness, official duties, maternity, training and earned leave, and so one can view a teachers use of legal leaves as equivalent to unobserved absence in the context of the model.

\(^{12}\)We are grateful to a referee for alerting us to this alternative interpretation of \( \alpha \) as a discount applied to earnings due to peer pressure.
The teacher’s first-order condition is
\[ MU_T(G) + \alpha w_T = U'(G) \left( \beta_T l_T^\rho + \beta_C l_C^\rho \right)^{1 - \frac{1}{\rho}} \beta_T l_T^{\rho - 1} + \alpha w_T \geq w_{0T} \quad (1) \]

The left-hand-side of the inequality is the marginal utility the teacher derives from devoting time to school which equals the gain from the additional shared good generated from additional teacher school attendance plus the expected loss of teacher income if the teacher spent that time shirking. The right-hand-side is the return to devoting that time increment to outside work or time in the home. The teacher sets \( l_T = 0 \) and would never attend if the inequality is violated, meaning that the utility from spending any time at school is dominated by spending that time outside school. At the other extreme, the teacher sets \( l_T = 1 \) and always attends if the marginal utility from the shared good plus expected income from attendance is greater than the value of time away from school. More generally, condition (1) holds with equality and so the teacher will spend at least some time shirking. Shirking increases as the anticipated cost from peer pressure or supervisory discipline decreases, as the value of time out of school increases, and as teacher salary and the marginal utility of the shared good decreases. Assuming an interior solution, the teacher’s equation governing time allocated to school will be of the form:
\[ l_T^* = l_T(l_C, \alpha, w_T, w_{0T}) \quad (2) \]

**Child’s attendance decision**

The child’s attendance choice, or that of the parent acting on behalf of the child, involves selecting \( l_C \) so as to maximize expected utility. The child has an opportunity cost of time spent in school, \( w_{0C} \), that represents the larger of either the value of time spent in home production or the wage in the local market for child labor. The child time allocation problem, assuming \( l_T \) is fixed, is
\[ \max_{l_C} EU_C = w_{0C}(1 - l_C) + U \left( \left( \beta_T l_T^\rho + \beta_C l_C^\rho \right)^{\frac{1}{\rho}} \right). \]

The child’s first order condition is:
\[ MU_C(G) = U'(G) \left( \beta_T l_T^\rho + \beta_C l_C^\rho \right)^{\frac{1}{\rho}} \beta_T l_T^{\rho - 1} \beta_C l_C^{\rho - 1} \geq w_{0C} \quad (3) \]

If the inequality is violated, the value of child time outside school exceeds the value in school and so \( l_C = 0 \). On the other hand, if the marginal utility of the shared good exceeds the opportunity cost of time, then the child attendance decision is \( l_C = 1 \). In general, there will be an interior decision so (3) holds with equality. The child’s reduced form equation governing time allocated to school will be of the form:
\[ l_C^* = l_C(l_T, w_{0C}) \quad (4) \]

The teacher’s and student’s reduced form equations represent two equations in the unknowns \( l_C \)
and $l_T$. From (2) and (4), it is clear that teachers’ attendance will depend on the elements of the child’s attendance function and students’ attendance will depend on the arguments in the teacher’s attendance. The equilibrium solution requires that the expected and the actual attendance are the same so that children fully anticipate the teacher’s attendance and teacher fully anticipates the child’s attendance. In that case, $l_T^* = l_T$ and $l_C^* = l_C$. Furthermore, the amount of the shared good $G$ in (1) is the same as in (3). Rearranging the conditions (1) and (3) yields

$$\frac{(w_{0T} - \alpha w_T)}{\beta_T l_T^{p-1}} = \frac{w_{0C}}{\beta_C l_C^{p-1}} \quad (5)$$

Applying our maintained assumption of an interior solution in (1), $MU_T > 0$ which implies that the numerator on the left-hand-side must be positive. That in turn requires that the value of a teacher’s time away from school must exceed the expected lost wage from being caught shirking. Because both sides of equation (5) are positive, we know that any factor that raises teachers’ attendance will also increase students’ attendance, and so there will be a positive correlation between their attendance decisions. We will test these predictions in the empirical work below.

## 4 Empirical Strategy

We are interested in estimating equations (2) and (4) from our theory. The linear approximation to the functional form in (2) is

$$l_T = \gamma_T^T l_C + \gamma_T^T w_T + \gamma_T^T w_{0T} + \gamma_T^T \alpha + \varepsilon_T \quad (6)$$

The specification shows that teacher attendance will depend on student attendance and measures of the teacher’s expected return from spending time inside and outside of class. The teacher’s compensation $w_T$ includes the teacher’s salary and also working conditions such as the availability of furniture in the school, commuting distance from home to school, and whether the teacher has head teacher or other supervisory responsibilities besides teaching. The teacher’s incentive to attend school will also reflect the likelihood of facing costs from shirking, $\alpha$. Our review of spot-check versus official attendance suggest that the likelihood of official penalties from shirking are greater in private schools and virtually nonexistent in government schools, and so we would expect greater attendance in private schools. Likelihood of unofficial sanctions by peers increase with the number of other teachers in the school\(^{13}\). Attendance will also vary negatively with factors that raise the

\(^{13}\)Fewer teachers can make collusive arrangements to share attendance, and more teachers would mean more chances of being observed absent. Olson (1965) argued that coordination between individuals in a large group become difficult. If we think of shirking as an objective which is shared by most teachers, then this objective is more likely to be fulfilled when the number of teachers in the school are small where possibilities of coordination
value of teacher time outside school, holding constant the salary paid in the school. These include
the teacher’s human capital (age and education) and the factors that alter the value of time in
household production (gender, marital status, and number of young children).

The linear approximation to (4) provides an equation explaining child attendance

\[ l_C = \gamma^C_T l_T + \gamma^C_0 w_0C + \varepsilon_C \] (7)

The child’s attendance will depend on the teacher’s attendance and on the value of the child’s
time outside relative to the value inside school. The relative value of time inside versus outside
class depends on school amenities such as furniture, class size, and teacher attributes and on child
age, gender, health and home attributes (parent and sibling demographics, wealth, and distance
from school).

The theory explains why teacher and student attendance will be jointly determined, and as
such, estimates of (6) and (7) are properly viewed as tests of the correlation between teacher and
student attendance conditional on other teacher attributes and other student attributes that might
be expected to affect school attendance. However, one cannot interpret the estimates of \( \gamma^T_C \) or \( \gamma^C_T \)
as the causal effect of child attendance on teacher presence or the impact of teacher attendance on
child time in school. The existence of a shared good between teachers and students only implies a
positive correlation between \( l_C^T \) and \( l_T^C \).

Our conditional correlations may be clouded by an unknown common factor that is not included
as a control in the regressions. Therefore, one could contend that a positive sign on \( \gamma^T_C \) and \( \gamma^C_T \)
is really due to a common missing factor in the regressions (6) and (7). Therefore, we propose two
alternative strategies to test for evidence of a positive correlation between teacher and student
attendance consistent with a shared good.

First, consider the projection of the teacher’s and child’s attendance on their own exogenous
variables

\[ l_T = \theta^T_0 w_T + \theta^T_0 w_0T + \theta^T_0 \alpha + \mu_T \] (8A)
\[ l_C = \theta^C_0 w_0C + \mu_C \] (8B)

The other strategy is to use (8A) and (8B) to suggest instruments for \( l_C \) and \( l_T \) in (6) and (7),
yielding the structural equations

failure are small. Olson effect holds in our data. Nevertheless, one of the authors recalls a series of school visits in
which the one teacher schools had one teacher present; the two teacher school had one teacher present and the three
teacher school had one teacher present.


22
\[ l_T = \delta_T^C \mathbb{E}_C(l_C \mid w_{0C}) + \delta_T^T w_T + \delta_0^T w_{0T} + \delta_0^T \alpha + \zeta_T \] (9A)

\[ l_C = \delta_C^T \mathbb{E}_T(l_T \mid w_{0T}, w_T, \alpha) + \delta_0^C w_{0C} + \zeta_C \] (9B)

Estimating (8A,B) and (9A,B) simultaneously provides unbiased estimates of the coefficients \( \delta_T^C \) and \( \delta_T^C \) which must be positive to be consistent with the prediction from the theory. The identifying assumption is that child household attributes do not affect teacher attendance and that teacher household attributes do not affect child attendance. Beyond these strategies, it is difficult to imagine a convincing instrument that would demonstrate a causal link between child and teacher attendance. For example, a weather shock like an unusually heavy rainy season would affect transportation to school for both students and teachers. A drought would create income shocks for both teachers and the children. An infectious disease would limit attendance of both students and teachers. A definitive test of a causal link between the two groups’ attendance rates would almost surely require an experimental design that exogenously raises student attendance without affecting teachers or that changes teacher absenteeism with altering student incentives to attend. An example is a girls scholarship program in Kenya that gave high scoring girls free tuition in a randomly drawn sample of schools. Performance by those girls improved, and their teachers also attended more regularly than did control students and teachers Kremer et al. (2009).

5 Data

Survey Characteristics

A sample of 257 schools forms the basis of the analysis. At each school, a teacher from each of the first three grades was selected at random. In schools with fewer than four teachers in the first three grades, all teachers were sampled. A total of 650 teachers were interviewed from 257 schools to obtain information on teacher household and school attributes. Information on teachers’ absence was obtained from two sources- the official attendance register kept at the school and a spot check of teacher attendance conducted during unannounced visits to the school.

In each of the teachers’ classes, two children were selected at random for the child attendance sample. For each child, daily attendance from the attendance register was obtained as well as a spot check on the same day as the teacher. For each child, an interviewer was dispatched to the household to administer a detailed survey of household demographics and socioeconomic status.
Dependent Variables

As discussed above in Section 2, the discrepancy between official teacher attendance from school registry data and the spot check attendance measures obtained by the surveyors suggest that official registry data underreports actual teacher absences. Hence we shall use the spotcheck measure as the dependent variable for teachers’ attendance equations. However no such discrepancy was found between children’s official attendances and their spotcheck measures, presumably because there is no incentive on the part of the teacher to under report child absence. Because the registry data represents two months of observed child behavior and appears reliable, we use it as the dependent variable measuring students’ attendance.

Independent Variable Selection

I. The theory suggested three factors that influence whether a teacher shows up in school on a given day. We assign variables from the surveys that reflect these variables. Note that some may fit in more than one category. For teachers, the key drivers are the incentives to attend, the opportunity cost of time in school and the probability of being caught while shirking.

- \( w_T \): the incentives to attend. The most obvious incentive to attend is the teacher’s salary. However, the teacher may still be paid, even if the teacher shirks. School furnishings such as desks, chairs, and blackboards affect the quality of the work environment. Some teachers are given administrative responsibilities beyond their teaching which can also affect the quality of the job.

- \( w_{0T} \): the opportunity cost of time in school. One likely source of higher opportunity cost of attendance will be due to household responsibilities. Teachers who are married and who have children under age six have greater value of time in the home. Commuting time from the teacher’s home to school raises the cost of attendance. Finally, noting that the teacher’s salary is also included in the regressors, teacher’s with greater endowments of skill that are valued outside school will have alternative earning possibilities outside school. We include the teacher’s age and education.

- \( \alpha \): The probability of being caught shirking. Schools with more teachers have greater opportunity to pass on responsibilities to another teacher, but there are also more difficulties in establishing collusive agreements on shirking. In addition, teachers may face more informal sanction from their colleagues when there are more peers who can observe attendance and impose penalties. Private schools are reputed to have closer supervision and no constraints on dismissal which increases the costs of shirking.

II. For the child’s attendance equation, the key issue is the relative value of child time out of school versus in school.
• \( w_{0C} \): the opportunity cost of child time. A child’s productivity outside relative to inside school will reflect the child’s age, health and gender. Having more siblings in the home can affect both ability to pay for schooling and represent an additional need for child time. The presence and abilities of parents affect the ability of the household to produce without using children to produce. Similarly, household wealth, measured as the first principal component of a vector of household asset measures, indicates the ability of the household to afford devoting child time to school\(^{14}\). Finally, commuting time from home to school increase the cost of devoting child time to school. School quality is indexed by school furnishings and whether the school is under private versus government management.

6 Results

Our model suggests that there will be a strong correlation between teachers’ attendance and that of their students. The simple correlation is 0.47, although one might suspect that is due to common school attributes. We report results of the teachers’ attendance equations in Table 3 and the coefficients of students’ equations are given in Table 4. If one were to assume that teachers view student attendance as an exogenously given, like other school characteristics then equation (8A) may be estimated simply using an ordinary least square with the student attendance as a part of the repressors. Similarly if the students take teachers attendance as a given school characteristic then OLS on (8B) will give us the required estimates. These results are reported in the first column of Table 3 and Table 4 respectively. The coefficient of student attendance on teachers’ attendance is quite large at 0.74 and highly significant. That of teachers’ attendance on student attendance is also quite large and statistically significant at 0.27. However, this equation is misspecified if we worry that teacher and student attendance may be subjected to common shocks that create the appearance of a correlation when it is not in fact present. The use of exogenous attendance also appears to bias other coefficients so that private school actually reduces teacher attendance while the presence of children under 6 in the home raises teacher attendance, results that seem implausible. We investigate the interrelationship between child and teacher attendance taking more care of their jointness in columns 2 and 3. The second column in each table reports the coefficients from ordinary least squares (OLS) estimation of equations (8A) and (8B). The key take away from those estimates is how little of the variation in teacher and child attendance can be explained by their own household and school information. The teachers’ demographic household and school information explains only 8% of the variation in teacher attendance while the children’s demographics, household attributes and school characteristics only explain 6% of the children’s attendance variation. The

\(^{14}\)We follow Filmer and Pritchett (2001) in aggregating a large vector of household attributes into a single measure. The thirteen household attributes include measures of the quality of home construction; availability of telephone, water, sewer and electricity; household human capital including measures of occupation and literacy; and possession of various household appliances and electronics.
error term from the teachers’ regression will be orthogonal to teachers’ demographic, household
and school attributes while the error terms from the children’s regression will be orthogonal to the
children’s demographic, household and school attributes. Nevertheless, the teacher and child errors
are correlated at 0.41. The unobserved heterogeneity (to the econometrician) in child and teacher
attendance are significantly positively correlated.

Equations (9A) and (9B) provide a structural model that will allow us to directly estimate the
effect of predictable teacher attendance on child attendance and predictable child attendance on
teacher attendance. We estimate the system jointly using seemingly unrelated regression. The
results are reported in the third column of Tables 3 and 4, respectively. However, we first estimated
each of the equations independently to derive a measure of the attendance variation that can be
explained when we add the predicted child attendance to the teachers’ equation and the predicted
teachers’ attendance to the children’s equation. The increase in the R² is dramatic. Adding
predicted child’s attendance to the teachers’ equation raises the percent of explained variation in
teachers’ attendance from 8% to 28%. Adding predicted teacher attendance in to the model
explaining the students’ attendance raises the percent of explained variation from 6% to 25%. It
seems apparent that attendance is a joint decision between teachers and students, consistent with
the prediction suggested by equation (5).

The key parameter in Table 3 is that on children’s attendance that is predicted on knowledge
of the children’s home environments and characteristics of the school. The coefficient of 0.87 is
highly significant and suggests that 100% child attendance will raise the teachers’ attendance by
87 percentage points. At sample means, the elasticity is 0.83, and so a 10% increase in child
attendance raises teacher attendance by 8.3%. These specifications also generate more reasonable
estimates of the other factors. Private school teachers now attend more regularly, as do teachers
without preschool children at home. Attendance in schools with more teachers is higher and this is
consistent with the posited link to peer pressure to attend. Duties outside the classroom (which
includes activities like manning polling booths at elections, representing the school at district level
meetings) tend to increase attendance as it is impossible to avoid detection if absent under these
circumstances. The negative significant coefficient for monthly salary stems from the fact that
public schools offer almost twice the salary of the private schools but experiences more absenteeism
than the private counterparts.

The key parameter in Table 4 is that on teachers’ attendance that can be predicted from
knowledge of the teacher’s household attributes and characteristics of the school. The coefficient is
0.39 and highly significant with an implied elasticity of 0.41. A 10% increase in teacher attendance
increases child attendance by 4%. Students are significantly more likely to attend if the school is
a private one with the corresponding coefficient being significant at 4.04. Private schools, which

\footnote{A more accurate relationship between monthly income and attendance is given by the correlation between the
two conditioned on whether a school is private or not. For a private school the correlation is 0.07, for public schools
it is 0.02 implying that the relationship is positive but weak.}
are typically better managed than the public ones, have a higher tuition fee and thus the families may be more compelled to send their children to school regularly\textsuperscript{16}. Apart from teacher attendance and private the presence of the parents appear to have some influence on the child’s decision to attend or not. However there are no other factors that matter more for teacher attendance than the attendance of their students, and there is no large influence on child attendance than the consistent presence of the teacher. Policies that encourage the attendance of one will increase the attendance of the other.

For example, conditional transfer programs have been shown to increase how regularly children attend school in Latin America. Our results suggest that conditional cash transfers aimed at raising children’s attendance will also increase how regularly teachers attend. Alternatively, the use of date-time digital cameras increases the attendance of teachers in rural India. Our results suggest that children in those communities will attend more regularly as well.

For the other variables, there are several avenues where policy could increase teacher attendance suggested by the results in Table 3. Family responsibilities appear to exacerbate teacher absenteeism, given the negative effects of teacher marital status and children on attendance. Attendance is greater in larger schools, presumably because of better monitoring of teacher attendance due to greater possibility of being noticed if absent. As hypothesized above, it may also be because it is more difficult to collude among a large group and ensure that absence goes unreported. Attendance is significantly higher in private schools where monitoring is greater and punishment for absenteeism more severe. Finally, teachers who have more administrative responsibilities attend more regularly as well. Our results show that policies that raise teacher attendance will increase time children spend in school.

For the other variables, there are several avenues where policy could increase teacher attendance suggested by the results in Table 3. Family responsibilities appear to exacerbate teacher absenteeism, given the negative effects of teacher marital status and children on attendance. Attendance is greater in larger schools, presumably because of better monitoring of teacher attendance due to greater possibility of being noticed if absent. As hypothesized above, it may also be because it is more difficult to collude among a large group and ensure that absence goes unreported. Attendance is significantly higher in private schools where monitoring is greater and punishment for absenteeism more severe. Finally, teachers who have more administrative responsibilities attend more regularly as well. Our results show that policies that raise teacher attendance will increase time children spend in school.

The regressions in Table 4 do not provide clear policy avenues to raise children’s attendance. It is clear that children attend private school more regularly, and other studies have demonstrated that in Pakistan, even low fee private schools generate better cognitive outcomes than government schools Alderman et al. (2001). Experimental work has shown that a girls school scholarship program that

\textsuperscript{16}Thaler (1980) discusses that people who pay for tickets to a basketball match are more likely to attend it under inclement weather than people who have received the tickets for free.
tied a tuition payment to attendance raised girls’ attendance significantly in Balochistan Province of Pakistan (Alderman et al., 2003; Kim et al., 1999), and perhaps that is an avenue that could be explored in future research.

7 Conclusion

For the other variables, there are several avenues where policy could increase teacher attendance suggested by the results in Table 3. Family responsibilities appear to exacerbate teacher absenteeism, given the negative effects of teacher marital status and children on attendance. Attendance is greater in larger schools, presumably because of better monitoring of teacher attendance due to greater possibility of being noticed if absent. As hypothesized above, it may also be because it is more difficult to collude among a large group and ensure that absence goes unreported. Attendance is significantly higher in private schools where monitoring is greater and punishment for absenteeism more severe. Finally, teachers who have more administrative responsibilities attend more regularly as well. Our results show that policies that raise teacher attendance will increase time children spend in school.

The proposed impetus for the interaction between teacher and student in this study is a shared good which serves as an additional motivation for both the teacher and the student to attend. The existence of the shared good means that weaker incentives than those being tried by governments may suffice to ensure better teacher attendance as long as students come to school regularly. While the shared good is unobservable, its existence creates a positive correlation in the error terms of equations explaining student and teacher attendance based solely on their own attributes and those of the school. We verify that prediction using data on teacher and student attendance from primary schools in Northwest Frontier Province of Pakistan. We also find evidence consistent with a second prediction: a dramatic increase in the fit of models of teacher attendance when we add predicted child attendance, and similarly, a dramatic increase in the explained variation in child attendance when predicted teacher attendance enters the model.

References


Table 1: Comparison of Spot-check and Official Teacher and Student Absences in 267 Schools in Northwest Frontier Province, Pakistan

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Government</th>
<th>Private</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teachers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spot-check absence(^a)</td>
<td>13.1%</td>
<td>16.3%</td>
<td>7.1%</td>
</tr>
<tr>
<td>Official absence(^b)</td>
<td>8.1%</td>
<td>8.7%</td>
<td>4.1%</td>
</tr>
<tr>
<td>Causal leave(^c)</td>
<td>9.6 days</td>
<td>9.4 days</td>
<td>10.9 days</td>
</tr>
<tr>
<td>% causal leave&gt;25 days(^c)</td>
<td>2.7%</td>
<td>0.0%</td>
<td>9.8%</td>
</tr>
<tr>
<td>Total leave</td>
<td>12.4 days</td>
<td>12.2 days</td>
<td>13.4 days</td>
</tr>
<tr>
<td>Proportion of parents saying their children are hurt by absent teachers</td>
<td>91%</td>
<td>93%</td>
<td>79%</td>
</tr>
<tr>
<td>Proportion of teachers saying absent colleagues are a burden</td>
<td>56%</td>
<td>62%</td>
<td>35%</td>
</tr>
<tr>
<td>Monthly Salary (rupees)</td>
<td>1951</td>
<td>2117</td>
<td>1008</td>
</tr>
<tr>
<td>Children</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spot-check absence(^a)</td>
<td>15%</td>
<td>16%</td>
<td>12%</td>
</tr>
<tr>
<td>Official absence(^b)</td>
<td>15.3%</td>
<td>12.3%</td>
<td>16.3%</td>
</tr>
</tbody>
</table>

\(^a\) Absence noted by the survey taker on the day of the school visit.
\(^b\) Absentee rate reported on the official registry during the month of the spot-check.
\(^c\) Leave taken during the previous academic year as reported on the official school registry.
Table 2: Teacher and student attributes by frequency of attendance, Primary grades K-2, Northwest Frontier Province, Pakistan

<table>
<thead>
<tr>
<th>Teacher Attributes</th>
<th>Male Teacher Attendance</th>
<th>Female Teacher Attendance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absent</td>
<td>Present</td>
</tr>
<tr>
<td>Age</td>
<td>32</td>
<td>33</td>
</tr>
<tr>
<td>Married</td>
<td>0.8</td>
<td>0.7</td>
</tr>
<tr>
<td>Highest Academic Qualification</td>
<td>3.7</td>
<td>3.8</td>
</tr>
<tr>
<td>Total number of children</td>
<td>8.0</td>
<td>7.3</td>
</tr>
<tr>
<td>Number of children &lt; 6</td>
<td>1.3</td>
<td>1.2</td>
</tr>
<tr>
<td>Commute to school</td>
<td>32</td>
<td>29</td>
</tr>
<tr>
<td>Monthly Salary (Rupees)</td>
<td>2097</td>
<td>2134</td>
</tr>
<tr>
<td>Added responsibilities</td>
<td>0.24</td>
<td>0.42</td>
</tr>
<tr>
<td>Number of teachers</td>
<td>4.1</td>
<td>4.6</td>
</tr>
<tr>
<td>Furniture</td>
<td>0.35</td>
<td>0.43</td>
</tr>
<tr>
<td>Private</td>
<td>0.02</td>
<td>0.09</td>
</tr>
<tr>
<td>Children’s Attendance</td>
<td>0.57</td>
<td>0.89</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Children Attributes</th>
<th>Boys Attendance Group</th>
<th>Girls Attendance Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lower Quartile</td>
<td>Upper Quartile</td>
</tr>
<tr>
<td>Age</td>
<td>7.4</td>
<td>7.1</td>
</tr>
<tr>
<td>Healthy</td>
<td>0.97</td>
<td>0.93</td>
</tr>
<tr>
<td>Dad Present</td>
<td>0.91</td>
<td>0.98</td>
</tr>
<tr>
<td>Mom Present</td>
<td>0.95</td>
<td>0.96</td>
</tr>
<tr>
<td>Education of dad</td>
<td>3.3</td>
<td>4.4</td>
</tr>
<tr>
<td>Education of mom</td>
<td>0.36</td>
<td>0.48</td>
</tr>
<tr>
<td>Wealth</td>
<td>4.84</td>
<td>6.42</td>
</tr>
<tr>
<td>Siblings</td>
<td>4.2</td>
<td>3.6</td>
</tr>
<tr>
<td>Distance from School</td>
<td>1.2</td>
<td>2.4</td>
</tr>
<tr>
<td>Siblings</td>
<td>4.1</td>
<td>3.5</td>
</tr>
<tr>
<td>Furniture</td>
<td>0.3</td>
<td>0.5</td>
</tr>
<tr>
<td>Private</td>
<td>0.09</td>
<td>0.25</td>
</tr>
<tr>
<td>Teacher’s Attendance</td>
<td>0.61</td>
<td>0.98</td>
</tr>
</tbody>
</table>

*statistic of the test of differences in the means between high and low attenders. *significance at the .10 level; **significance at the .05 level.
Table 3: OLS and SUR estimates of teachers' spot-check attendance

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>27.9**</td>
<td>91.1**</td>
<td>27.7*</td>
</tr>
<tr>
<td></td>
<td>(7.59)</td>
<td>(6.40)</td>
<td>(15.41)</td>
</tr>
<tr>
<td>Male (=1 if gender is Male)</td>
<td>-0.79</td>
<td>2.87</td>
<td>-1.53</td>
</tr>
<tr>
<td></td>
<td>(2.47)</td>
<td>(2.50)</td>
<td>(2.10)</td>
</tr>
<tr>
<td>Age</td>
<td>1.01</td>
<td>0.738**</td>
<td>1.014**</td>
</tr>
<tr>
<td></td>
<td>(0.17)</td>
<td>(0.17)</td>
<td>(0.23)</td>
</tr>
<tr>
<td>Highest academic qualification</td>
<td>-0.67</td>
<td>-2.13*</td>
<td>-0.408</td>
</tr>
<tr>
<td></td>
<td>(1.23)</td>
<td>(1.23)</td>
<td>(1.20)</td>
</tr>
<tr>
<td>Married</td>
<td>-12.5**</td>
<td>-12.5**</td>
<td>-7.56*</td>
</tr>
<tr>
<td></td>
<td>(2.60)</td>
<td>(2.60)</td>
<td>(3.75)</td>
</tr>
<tr>
<td>Number of small children (less than the age of six)</td>
<td>2.94**</td>
<td>1.041</td>
<td>-1.54 *</td>
</tr>
<tr>
<td></td>
<td>(0.97)</td>
<td>(0.96)</td>
<td>(0.92)</td>
</tr>
<tr>
<td>Commuting time</td>
<td>-0.01</td>
<td>-0.014</td>
<td>-0.014</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>Monthly salary</td>
<td>-0.008**</td>
<td>-0.004</td>
<td>-0.006**</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>Other responsibilities</td>
<td>3.36*</td>
<td>7.539**</td>
<td>4.098**</td>
</tr>
<tr>
<td></td>
<td>(1.93)</td>
<td>(1.92)</td>
<td>(1.55)</td>
</tr>
<tr>
<td>Total number of teachers</td>
<td>0.344</td>
<td>0.297</td>
<td>0.357*</td>
</tr>
<tr>
<td></td>
<td>(0.27)</td>
<td>(0.28)</td>
<td>(0.20)</td>
</tr>
<tr>
<td>Furniture (includes desk, chair, blackboard)</td>
<td>3.56</td>
<td>5.001</td>
<td>-0.184</td>
</tr>
<tr>
<td></td>
<td>(3.20)</td>
<td>(3.23)</td>
<td>(2.93)</td>
</tr>
<tr>
<td>Private</td>
<td>-2.42</td>
<td>7.12*</td>
<td>6.98**</td>
</tr>
<tr>
<td></td>
<td>(4.12)</td>
<td>(4.05)</td>
<td>(3.56)</td>
</tr>
<tr>
<td>Student Attendance (actual in column 1, instrumented in column 3)</td>
<td>0.74**</td>
<td>-</td>
<td>0.869**</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.04)</td>
<td></td>
</tr>
<tr>
<td>R-sq</td>
<td>0.28</td>
<td>0.08</td>
<td>0.28</td>
</tr>
<tr>
<td>Sample Size</td>
<td>926</td>
<td>926</td>
<td>772</td>
</tr>
</tbody>
</table>

† Note: Standard errors in parentheses are corrected for clustering at the school level. In the seemingly unrelated regression, repeated observations on the same teacher were weighted by a factor of 1/n where n is the number of children in the student sample that have that teacher. Col (1) gives the OLS with exogenous students' attendance, col (2) gives the OLS estimates and col (3) gives the SUR conducted jointly with column 2 in table 4. **: significant at alpha=0.05 ; *: significant at alpha=0.10
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>57.9**</td>
<td>79.38**</td>
<td>43.3**</td>
</tr>
<tr>
<td></td>
<td>(5.70)</td>
<td>(6.18)</td>
<td>(5.68)</td>
</tr>
<tr>
<td>Male (=1 if gender is Male)</td>
<td>-0.69</td>
<td>-1.57</td>
<td>1.59</td>
</tr>
<tr>
<td></td>
<td>(1.25)</td>
<td>(1.39)</td>
<td>(1.28)</td>
</tr>
<tr>
<td>Age</td>
<td>-0.53</td>
<td>-0.77</td>
<td>-0.086</td>
</tr>
<tr>
<td></td>
<td>(0.40)</td>
<td>(0.44)</td>
<td>(0.35)</td>
</tr>
<tr>
<td>Healthy</td>
<td>-0.46</td>
<td>-1.08</td>
<td>-2.559</td>
</tr>
<tr>
<td></td>
<td>(2.45)</td>
<td>(2.74)</td>
<td>(2.47)</td>
</tr>
<tr>
<td>Distance from school</td>
<td>0.1</td>
<td>0.18</td>
<td>0.053</td>
</tr>
<tr>
<td></td>
<td>(0.18)</td>
<td>(0.21)</td>
<td>(0.29)</td>
</tr>
<tr>
<td>Siblings</td>
<td>-0.51*</td>
<td>-0.44</td>
<td>0.039</td>
</tr>
<tr>
<td></td>
<td>(0.30)</td>
<td>(0.34)</td>
<td>(0.29)</td>
</tr>
<tr>
<td>Dad present</td>
<td>1.5</td>
<td>6.80**</td>
<td>1.122</td>
</tr>
<tr>
<td></td>
<td>(2.75)</td>
<td>(3.05)</td>
<td>(2.51)</td>
</tr>
<tr>
<td>Mom present</td>
<td>8.44**</td>
<td>7.22*</td>
<td>6.266</td>
</tr>
<tr>
<td></td>
<td>(3.53)</td>
<td>(3.95)</td>
<td>(4.06)</td>
</tr>
<tr>
<td>Education of dad</td>
<td>-0.01</td>
<td>-0.022</td>
<td>0.182</td>
</tr>
<tr>
<td></td>
<td>(0.13)</td>
<td>(0.14)</td>
<td>(0.13)</td>
</tr>
<tr>
<td>Education of mom</td>
<td>0.004</td>
<td>0.102</td>
<td>-0.006</td>
</tr>
<tr>
<td></td>
<td>(0.24)</td>
<td>(0.27)</td>
<td>(0.26)</td>
</tr>
<tr>
<td>Wealth</td>
<td>0.013</td>
<td>0.108</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>(0.16)</td>
<td>(0.18)</td>
<td>(0.17)</td>
</tr>
<tr>
<td>Furniture (includes desk, chair, blackboard)</td>
<td>-2.52</td>
<td>-2.08</td>
<td>-2.21</td>
</tr>
<tr>
<td></td>
<td>(1.76)</td>
<td>(1.97)</td>
<td>(1.99)</td>
</tr>
<tr>
<td>Private</td>
<td>5.678**</td>
<td>8.23**</td>
<td>4.039**</td>
</tr>
<tr>
<td></td>
<td>(1.76)</td>
<td>(1.96)</td>
<td>(2.13)</td>
</tr>
<tr>
<td>Teacher attendance (actual in column 1, instrumented in column 3)</td>
<td>0.270**</td>
<td>-0.387**</td>
<td>-0.387**</td>
</tr>
<tr>
<td></td>
<td>(5.70)</td>
<td></td>
<td>(0.02)</td>
</tr>
<tr>
<td>R-sq</td>
<td>0.25</td>
<td>0.06</td>
<td>0.25</td>
</tr>
<tr>
<td>Sample Size</td>
<td>970</td>
<td>970</td>
<td>772</td>
</tr>
</tbody>
</table>

† Note: Standard errors in parentheses are corrected for clustering at the school level. In the seemingly unrelated regression, repeated observations on the same teacher were weighted by a factor of 1/n where n is the number of children in the student sample that have that teacher. Col (1) gives the OLS with exogenous teachers’ attendance, col (2) gives the OLS estimates and col (3) gives the SUR conducted jointly with column 2 in table 3. **: significant at alpha=0.05; *: significant at alpha=0.10
Table 5: Summary Statistics of Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Max</th>
<th>Min</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>30.6</td>
<td>0.27</td>
<td>15</td>
<td>60</td>
<td></td>
<td>Age of the teachers</td>
</tr>
<tr>
<td>Male</td>
<td>0.66</td>
<td>0.02</td>
<td>0</td>
<td>1</td>
<td>D</td>
<td>=1 if the teacher is Male</td>
</tr>
<tr>
<td>Highest academic qualification</td>
<td>3.72</td>
<td>0.03</td>
<td>1</td>
<td>6</td>
<td>C</td>
<td>Six categories included</td>
</tr>
<tr>
<td>Married</td>
<td>1.61</td>
<td>0.03</td>
<td>1</td>
<td>2</td>
<td>C</td>
<td>Two Categories</td>
</tr>
<tr>
<td>Number of small children</td>
<td>0.92</td>
<td>0.04</td>
<td>0</td>
<td>6</td>
<td></td>
<td>Less than the age of four</td>
</tr>
<tr>
<td>Commuting time</td>
<td>30.9</td>
<td>1.02</td>
<td>0</td>
<td>300</td>
<td></td>
<td>In minutes</td>
</tr>
<tr>
<td>Monthly salary</td>
<td>1896</td>
<td>20.5</td>
<td>0</td>
<td>3366</td>
<td></td>
<td>In Rupees</td>
</tr>
<tr>
<td>Other responsibilities</td>
<td>0.38</td>
<td>0.01</td>
<td>0</td>
<td>1</td>
<td>D</td>
<td>=1 if other responsibilities are assigned</td>
</tr>
<tr>
<td>Number of Teachers</td>
<td>4.92</td>
<td>3.94</td>
<td>1</td>
<td>26</td>
<td></td>
<td>Total teachers in the primary school</td>
</tr>
<tr>
<td>Furniture</td>
<td>0.5</td>
<td>0.35</td>
<td>0</td>
<td>1</td>
<td>D</td>
<td>=1 if privately owned</td>
</tr>
<tr>
<td>Private</td>
<td>0.17</td>
<td>0.38</td>
<td>0</td>
<td>1</td>
<td>D</td>
<td>=1 if the school is not owned by the state</td>
</tr>
<tr>
<td>age</td>
<td>7.41</td>
<td>1.38</td>
<td>2</td>
<td>27</td>
<td>D</td>
<td>Age of the student</td>
</tr>
<tr>
<td>Male</td>
<td>0.67</td>
<td>0.47</td>
<td>0</td>
<td>1</td>
<td>D</td>
<td>=1 if gender is Male</td>
</tr>
<tr>
<td>Healthy</td>
<td>0.95</td>
<td>0.22</td>
<td>0</td>
<td>1</td>
<td>D</td>
<td>=1 if student is healthy</td>
</tr>
<tr>
<td>Distance from school</td>
<td>1.35</td>
<td>3.05</td>
<td>0</td>
<td>32</td>
<td></td>
<td>In kilometer</td>
</tr>
<tr>
<td>Siblings</td>
<td>3.88</td>
<td>1.94</td>
<td>0</td>
<td>14</td>
<td></td>
<td>Number of brothers and sisters</td>
</tr>
<tr>
<td>Dad present</td>
<td>0.91</td>
<td>0.28</td>
<td>0</td>
<td>1</td>
<td>D</td>
<td>=1 if present</td>
</tr>
<tr>
<td>Mom present</td>
<td>0.94</td>
<td>0.25</td>
<td>0</td>
<td>1</td>
<td>D</td>
<td>=1 if present</td>
</tr>
<tr>
<td>Education of dad</td>
<td>4.62</td>
<td>5.05</td>
<td>0</td>
<td>16</td>
<td></td>
<td>Index of wealth created from household interview</td>
</tr>
<tr>
<td>Education of mom</td>
<td>0.94</td>
<td>2.46</td>
<td>0</td>
<td>14</td>
<td></td>
<td>Index of wealth created from household interview</td>
</tr>
<tr>
<td>Wealth</td>
<td>5.44</td>
<td>3.9</td>
<td>0.55</td>
<td>28.62</td>
<td></td>
<td>Index of wealth created from household interview</td>
</tr>
</tbody>
</table>

| Teachers’ Attendance          | 86.9  | 31.1      | 0   | 100 | D    | =100 if the Teacher is present on two spot-checks |
| Children’s Attendance         | 84.7  | 18.9      | 33.33| 100 |      | =100 if the child attends full time on the attendance register |

† Variable Type: 'D' refers to Dummy and 'C' refers to Categorical. Highest Academic Qualification: 1 is 6th Grade Pass, 2 is 8th Grade Pass, 3 is Matric, 4 is FA/FSc, 5 is BA/BSc, 6 is MA/Msc. Other responsibilities: tasks other than teaching e.g. administrative work of local govt., work of polling officer, work at behest of political parties etc. Married: 1 is Never married, 2 is Married, Widowed.
Chapter Two
On the Interpretation of Bribery in a Laboratory Corruption Game: Moral Frames and Social Norms

Ritwik Banerjee*
Aarhus University, Denmark
(Second R&R at Experimental Economics)

Abstract

Past studies on laboratory corruption games have not been able to find consistent evidence that subjects make “immoral” decisions. A possible reason, and also a critique of laboratory corruption games, is that the experiment may fail to trigger the intended immorality frame in the minds of the participants, leading many to question the very raison d’être of laboratory corruption games. To test this idea, we compare behavior in a harassment bribery game with a strategically identical but neutrally framed ultimatum game. The results show that fewer people, both as briber and bribee, engage in corruption in the bribery frame than in the alternative, suggesting that moral costs are indeed at work. To provide further support that the bribery game does measure moral costs, we elicit the shared perceptions of appropriateness of the actions or social norm, under the two frames. We show that the social norm governing the bribery game frame and ultimatum game frame are indeed different and that the perceived sense of social appropriateness plays a crucial role in determining the actual behavior in the two frames. Finally, we comment on the external validity of behavior in lab corruption games.

Keywords: Corruption, Framing Effects, Social Norms, External Validity

JEL Classification: C91 C92 D03

*rbanerjee@econ.au.dk; tel: +45 50291492. I thank Goutam Gupta and Amitabha Chatterjee for letting me conduct the experiment at Jadavpur University and Presidency University, respectively. I am grateful to Abhishek Das for his excellent assistance while conducting the experiment. I am also thankful to Alexander Koch, John A List, Klaus Abbink, Roel van Veldhuizen, Arnab Mitra, Joydeep Bhattacharya participants at TIBER XII and the 24th Jerusalem School in Economic Theory and two anonymous referees for their comments and suggestions. All remaining errors are mine. The grant for the experiment came from the Department of Economics and Business, Aarhus University.
1 Introduction

Laboratory experiments in Economics, despite their widespread use today, have often been subjected to (unfair) criticism on account of lacking realism and generalizability. In particular critics have questioned the external validity of the results and have argued that inconsequential size of stakes, desire to maintain a positive self image and a lack of an appropriate moral frame, among other things, confound the inferences drawn from them (for a general critique of lab experiments see Levitt and List (2007)). However, lab experiments have since received considerable defense as economists have concurred that the possibility of controlled variations in lab environments, thereby generating causal inferences, is a valuable source of knowledge (Falk and Heckman, 2009).

This is reflected in the increasing popularity of lab based corruption games as tools to identify, measure and understand the otherwise clandestine phenomenon called corruption. These tools allow us to manipulate the institutional (Alatas et al. (2009), Abbink et al. (2012), Banuri and Eckel (2012b)), social (Banerjee et al., 2013) and cultural contexts (Cameron et al. (2009), Barr and Serra (2010)), which in turn help us draw causal inferences. However, it is not entirely surprising that laboratory corruption games have been received with a fare share of skepticism too, primarily on the following two grounds. First, subjects maintain a positive self image in lab and therefore affect external validity and second, experimenters are unable to impose an appropriate moral frame in an artificial lab environment. While it is possible that the desire to maintain self image may be a confound, any observed measure of corruption in the lab should be interpreted as a lower bound of behavior in the field. Furthermore, if it is the treatment effect, as opposed to the absolute level, that is of interest, then a laboratory corruption game is a useful and relatively inexpensive tool for the researchers. Finally, Armantier and Boly (2013) in a clever and convincing study find evidence that results from lab based corruption games are indeed generalizable outside the lab and that the generalizability holds across cultures.

However, the other important feature of lab based corruption games is yet to be established on firm grounds. Do subjects in the game actually make decisions under the moral frame that the experimenter intends to impose? In other words, for an act to be interpreted as an unethical act we must find compelling evidence of a moral cost at play in the exercise of that act. This important feature, interpreted in terms of the treatment effect between the corruption frame and a neutral frame, has been found to be missing (Abbink and Hennig-Schmidt, 2006) or weak (Barr and Serra, 2009) in the past studies.

---

1For instance, Camerer (Forthcoming) in his reply to Levitt and List (2007) shows that the overwhelming majority of lab experiments do indeed generalize to comparable field settings. Kessler and Vesterlund (Forthcoming) emphasize that it is the qualitative effect (i.e. direction of effect), as opposed to quantitative effect (i.e. precise size of the effect), which more generalizable.

2Levitt and List (2007) warn us that lab experiments may not reveal so much when it comes to identifying deep structural parameters. Then one needs to take recourse to field experiments despite the fact that it is often prohibitively expensive as List (2012) found out with a $103,000 field experiment on corruption.

3Two distinct questions emerge from this: One, do people operate under the same moral environment in the two
While all studies incorporating such a game show that subjects do not opt for the highest material pay off, it is still not clear whether they act under the influence of the psychological frame of morality that the experimenter intends to impose. In this paper we revisit the question whether moral costs are at play in a lab based corruption game. We implement a design where in one treatment a harassment bribery game is implemented. In the counterfactual, a strategically identical but differently framed ultimatum game is implemented in which the psychological reference point of the subjects is changed by a subtle change in entitlement.

As a second test of whether the appropriate moral frame is at work in these corruption games, we elicit measures of how socially appropriate actions in each of the two treatments are. To do so we follow Krupka and Weber (2012) and elicit commonly held beliefs about social norms through a coordination game designed to elicit second order beliefs in an incentive compatible way. Socially held beliefs about what is the right thing to do potentially play a vital role in determining behavior. The author recalls an instance where a friend from India underwent a radical change in his driving practices after a four week travel to the United States, as he followed the newly internalized norms (e.g. sticking to lanes, not honking) which he had acquired. His action was being dictated, at least in part, by what was considered appropriate in the US. The strategy of eliciting measures of social appropriateness therefore enables us to see whether participants assign different moral standards for the two strategically identical but differently framed environments. This, in turn, helps us explore in very concrete terms, how the elicited measures of social appropriateness affect actual behavior, if at all.

In our between subject design “citizens” and “public officials” play a real effort harassment bribery game. A citizen performs a task and earns a prize if successful. A public official, however, may want a bribe in order to let the citizen have her prize. The citizen may then accept or reject the demand for a bribe. In the counterfactual treatment, Participant A (analogous to the Citizen) - upon successfully completing the task - earns the right to go to the second stage of the game. At the second stage, Participant B (analogous to the Public Official) plays an ultimatum game, with the same stake size as the prize of the bribery frame, and decides how much to share with Participant A, if anything. The latter can then accept or reject Participant B’s offer. Though frames and is this why one does not observe framing effects? Two, if yes what is the appropriate counterfactual of a corruption experiment that will potentially yield evidence of moral cost at work? Abbink and Hennig-Schmidt (2006) speculate that neutral frames are insufficient to induce an alternative behavioral norm.

We follow Fehr and Fischbacher (2004) to define social norm for our purpose: “The standards of behavior that are based on widely shared beliefs how individual group members ought to behave in a given situation”. There have been various other definition of social norms in the literature: Pareto noted that “... people have opinions about how they should or should not behave. They also have opinions about how others should or should not behave.” Ostrom (2000) further emphasized on the mutually shared aspect of social norms and defined it as “...shared understandings about actions that are obligatory, permitted or forbidden”.

Harassment bribery is a form of bribery where a public official asks for bribe from a citizen who is entitled to a service that the official is obligated to provide. This form of bribery is very common in developing countries where citizens are entitled to government services but either they have to pay a bribe in order to obtain them or avoid inordinate procedural delays. Such services include issuance of a passport or a driver’s license (given that the candidate has passed the driving test).
strategically identical, these two treatments differ in terms of the sense of entitlement among the subjects and this feature, we argue, is crucial in generating the framing effects that we see. We then relate corruption behavior in the framed lab experiment to stealing behavior, in order to understand what the behavioral primitive measured by corruption in bribery game actually is. Finally, we elicit the social appropriateness of the two frames through a second experiment with a between-subject design and then examine whether the strategically identical choice in one is considered as more inappropriate than in the other.

Our results show that retained shares in the ultimatum game are higher than the bribes demanded under the corruption frame. This points to evidence of a psychological moral cost at work in the bribery frame - something which been largely missing in past studies. A caveat though is that in unlike past studies, ours is a two-player game where corruption takes the form of a harassment bribe and does not generate negative externalities to others. Besides, our experiment is conducted in India which is highly corrupt and not in western countries with low corruption norms.

In addition to difference in actual behavior, social appropriateness measures differs considerably for the same strategies across the two treatments as well. Strategies are considered more socially inappropriate in the corruption frame than in the alternative. Finally, we show that the actually observed actions can be explained by the elicited social appropriateness measures. Overall, the findings suggest that laboratory corruption games indeed measure unethical behavior and thus lend themselves as cheap diagnostic tools which can help analyze various public policy interventions.

The contributions of this paper are the following. First, the paper makes a contribution to the framing literature by studying the effect of difference in entitlement in two strategically identical frames. Second, the treatment effects we find indicate that the experimenters successfully impose the appropriate moral frame in laboratory corruption games. This is indeed good news for laboratory experiments in general and lab based corruption games in particular. Our results, combined with the results obtained by Armantier and Boly (2013), will go a long way in convincing the skeptics of lab based corruption games. Third, we measure in clear categorical terms the amorphous concept of norms governing an unethical act and contrast it with that of a strategically identical but differently framed game. Fourth, to the best of our knowledge, this is one of the first attempts to answer how actual corruption behavior is related to socially held perceptions about what is morally the right thing to do. Notice that this result is distinct from other prevalent explanations of why people obey norms - e.g. fear of sanctions upon violation of norms (Fehr and Fischbacher, 2004), to obtain self esteem (Bernheim, 1994) or for informational advantages (Banerjee, 1992).

This paper is divided into six sections. Following the introduction, Section 2 briefly reviews the literature of framing effects as well as that of social norms. In Section 3 the experimental designs of the two main treatments are laid out along with the design for elicitation of social appropriateness. Section 4 presents and discusses the main results while the summary and conclusion is given in Section 5.
2 Related Literature

The literature, on how perceptions about what is morally the right thing to do affect actual corrupt behavior, is still nascent. Not only are such perceptions vulnerable to common factors which might also affect corrupt behavior, but conversely corrupt behavior might determine such perceptions as well. Thus, economists have turned to cross country studies to relate cultural norms to corruption\(^6\) (see for instance Treisman (2000), Serra (2006), Fisman and Miguel (2007)), though lack of comparable data greatly limits that scope.

This limitation has been tackled in the recent years by introducing controlled variations in the experiments but again with mixed success (for a review of laboratory corruption games see Serra and Wantchekon (2012) and for corruption and culture see Banuri and Eckel (2012a)). Cameron et al. (2009) conduct comparable experiments in countries with different corruption levels and Barr and Serra (2010) study corruption behavior among students from different countries but both fail to find across the board effect of home country corruption indices on behavioral responses in the lab. However, social norms in these studies are proxied by the aggregate indices of corruption, which measures (however inaccurately) perceptions of prevalence rather than that of appropriateness.

From a different perspective, behavior has been analyzed through the lens of contagion effect or the tendency to conform to behavior of others (see for instance Bicchieri and Xiao (2009), Krupka and Weber (2009), Cason and Mui (1998) for dictator games and Innes and Mitra (2013) for deception games). Conceptually the prevalence of unethical acts is very different from social norms which are a set of commonly held, unwritten normative codes governing our actions. Despite these codes not binding, they impose a substantial cost on people by way of inflicting moral sanctions for following a contrarian way.

Accordingly in our view social norms, culture, prevalence though often approximated with each other, are very different. This approximation is tackled in this paper by directly eliciting measures of social norm or appropriateness using a coordination device introduced by Krupka and Weber (2012). We use this device to elicit the measure of social appropriateness for the two frames we study and then predict cross treatment behavior\(^7\).

In the context of laboratory corruption games the role of framing has been analyzed by Abbink and Hennig-Schmidt (2006) and then subsequently by Barr and Serra (2009). Abbink and Hennig-Schmidt (2006) rely on using value free or “neutral” vocabulary to create a context free environment and pit it against a “loaded” environment. Their hypothesis is that the moral frame in a “loaded”

---

\(^6\)Though here we mainly discuss the experimental literature, several theoretical explanations have been advanced about why people may conform to social norms. While Banerjee (1992) explained preponderance of adherence to social norms in terms of informational advantage of others, Fehr and Fischbacher (2004) and Bernheim (1994) explained the conformity to social norms in terms of fear of sanction in case of violation and desire to gain social esteem, respectively.

\(^7\)Though Reuben and Riedl (2013) acknowledge that this tool is clearly a superior method of eliciting normative views than those based on questionnaires, only a handful of studies (for example Gächter et al. (2013)) have used this.
environment will lead to less corruption. However, they do not find a framing effect and attribute it, among other things, to the fact that even with a neutral frame, the game captures the quintessential features of corruption.

Barr and Serra (2009) later point to several other plausible reasons why no framing effect was found - mismatch between the created briber-bribee environment and real life experiences of subjects, triggering of role play rather than moral compunctions induced by “artificial” environments, punishment strategies in both neutral and loaded versions, etc. However, despite correcting for the majority of these limitations in their own study, Barr and Serra (2009) find mixed results as far as framing effects are concerned. For low externalities, private citizens are not found to be more likely to offer bribes under a corruption frame but a statistically significant difference is found in the mean bribe offered. However, for high externalities, and also for the pooled sample, private citizens are more likely not to offer a bribe in the corruption frame than in the abstract frame; though no significant differences were found in the mean bribe offered. Finally, there was no framing effect found in the public official’s decision to accept or reject bribes. To summarize, despite some instances, framing effects have been found to be largely absent.

Barr and Serra (2009) claim that their absence of framing effects in the role of the “public servant” is “consistent with the petty corruption frame appearing artificial to student subjects”. One wonders why students, who can identify themselves as bribe offering private citizens, cannot identify with bribe receiving public officials. Further, the conclusion of Cooper et al. (1999) that familiarity with the “underlying structure” of the game is important in order to find effects may hold true even without the subject population having real life experience with the exact roles which are being examined in the game. Therefore, the literature on framing demands a further investigation on why framing effect may or may not be found.

This paper argues that the reasons for the negative result of Abbink and Hennig-Schmidt (2006) and somewhat weak effects found in Barr and Serra (2009) lie elsewhere. In order to find the effects of “psychological and social factors” the ideal control treatment is not merely a neutrally framed environment but one in which the expectations of the agents are suitably changed. The use of neutral language does not alter the psychological reference point as subjects by and large identify with the underlying structure, thereby resulting in no evidence of framing effects. The ideal counterfactual should in fact be one in which the reference frame itself is altered keeping the strategic choices identical. This is ensured not just by using a neutral language but also by altering the sense of entitlement which in turn helps change the reference frame. We argue that this is crucial in any study which seeks to investigate the role of a particular frame.

---

8 The lack of framing effects has also been found in dictator games (Dreber et al., 2013).
9 Barr and Serra (2009) does find that the share of public servants who refuse to take a bribe is consistently higher in the corruption frame than in the abstract frame but the differences are not significant.
3 Experimental Design

We implement a harassment bribery game\(^{10}\) with a real effort task. Real effort task has been found to induce a sense of “ownership” (see for instance Hoffman et al. (1994), Ruffle (1998)) - this is crucial for the sense of harassment to be triggered. Successful completion of the real effort task in the bribery game treatment leads to a prize for a citizen but a public official may demand a bribe in order to let her have the prize. Of course the citizen may refuse or accept the demand depending on whether the value of the prize outweighs the (material and moral) cost of bribe. Then we exploit the structure of the harassment bribery game and posit it as an ultimatum game where successful completion of the real effort task merely leads to qualification for Player A to play the next stage of the game. In the next stage Player B splits an amount (equivalent to the prize) between himself and Player A and the latter decides whether to accept it or not. It is now clear that despite the change in the sense of entitlement, the two games remain strategically identical.

It is important to note that we do not run the treatment design equivalent to that of Abbink and Hennig-Schmidt (2006) and Barr and Serra (2009) i.e. we do not run a treatment which relies entirely on neutral language. This is because the fact that the treatment effect in this context is fairly weak is a robust finding of two studies.\(^{11}\) Finally our experiment is entirely a pen and paper one.

Due to the fact that the result from the second experiment is used to make inferences from the first, we begin by laying out the designs of the two experiments and then go on to the analysis of the results.

3.1 First Experiment

3.1.1 Harassment Bribery Game Treatment

Figure 1(a) lays out the bribery game (BG henceforth). A citizen (C) has fifteen minutes to perform a real effort task\(^{12}\). The task is graded by a matched public official (PO) who is seated in another room. If a citizen scores 10 or above she “passes” the test and is entitled to a prize of 400 Mohars (M400). If she fails to obtain at least 10 then she earns only the participation fee of M200.

---

\(^{10}\)Abbink et al. (2012) implements a version of harassment bribery game but it was not known to the author when this experiment was conceived.

\(^{11}\)There is a fourth possibility of a complete 2x2 design - one where the ultimatum game treatment is framed in terms of loaded language but passing the test only qualifies one to the next stage of the game. Since a citizen does not deserve a prize, the question of bribe becomes irrelevant. One can however introduce an ultimatum game at the second stage with a neutral word for bribe such as transfer. With a mix of loaded terms (citizen, public official) and neutral terms (transfer instead of bribe), it is not clear what the treatment will measure.

\(^{12}\)A version of the matrix or box task, introduced by Mazar and Ariely (2006), was used. In the task subjects had fifteen minutes to find two numbers in a 3 by 3 matrix which added up to 10. For example a matrix may have 4.55 and 5.45 which adds up to 10. There were twenty such matrices. The task was specifically chosen to ensure that subjects knew how many they had solved correctly. Thus, they knew whether they had won the prize and any demand for a bribe might have been considered unfair.
However, even if the citizen scores 10 or above, the public official may demand a bribe of amount $b$ in order to let the citizen have her prize. Notice that the bribe of amount $b$ is an extract from the value of the citizen’s entitlement, which in this case is $M400$. Any $b > 400$ is likely to be rejected by a reasonable citizen. So, we let $b \in \{0, 20, 40...400\}$. Upon receiving a bribe demand for $b$, a citizen may accept it or reject it. If she accepts the demand for a bribe, she earns $M(200 + 400 - b)$ whereas if she rejects it, she earns the participation fee of $M200$ only.

Public Official on the other hand gets a participation fee of $M200$ and a salary of $M400$ for the task of approving the citizens. Thus, if he demands a bribe $b$ and his demand is accepted, then his earnings are $M(200 + 400 + b)$. If his demand is rejected however, his earnings are $M(200 + 400)$ i.e. $M600$ only. Since the official earns at least as much as the citizen, an act of bribery cannot be explained by alternative explanations such as inequity aversion. This is a one shot game which reflects the standard harassment bribe situation where a briber and a bribee meet only once.

Since our interest is in emulating harassment bribe situations, the real effort task which a citizen undertakes is carefully calibrated such that she is more likely to cross the threshold score of 10$^{13}$. At the same time the task itself is crucial as it induces a sense of entitlement and hence a sense of harassment in her.

### 3.1.2 Ultimatum Game Treatment

Figure 1(b) lays out the ultimatum game treatment (UG henceforth). In this treatment Participant A (analogous to citizens in BG) performs the real effort task which in turn is graded by a randomly matched Participant B (analogous to public officials in BG). If Participant A (P-A henceforth) solves 10 or more boxes correctly within the stipulated fifteen minutes, then she qualifies for the next round of the game in which Participant B (P-B henceforth) decides to keep $Mx$ for himself and make an offer $M(400 - x)$ to P-A. Thus in this case P-A’s performance in the test only determines whether she proceeds to the next part of the game or not$^{14}$. If she does proceed to the next stage, then she can decide whether to accept or reject the offer in which cases she receives $M(200 + 400 - x)$ and $M200$, respectively. If P-B’s offer is accepted then P-B gets $M(200 + 400 + x)$, if it is rejected, then he only gets $M(200 + 400)$, i.e. $M600$. In this way the treatment represents a strategically identical version of BG but there is a subtle change in the sense of entitlement.

This alternative frame is designed to change the sense of guilt or the psychological moral cost of indulging in a corrupt act$^{15}$. The naming of subjects as “Player 1” and “Player 2” instead of “public official” and “citizen” and using words like “transfer” instead of “bribe” further induce

---

$^{13}$Notice that letting citizens pay a bribe when they have failed the test is bribery for sure but does not amount to harassment bribery. The ultimatum game parallel applies only for the latter and thus, we were specially interested in it - hence the calibration. In this treatment 95% of the participants scored at least 10.

$^{14}$Calibration of the task ensured that 92.5% of the subjects qualified for the next round.

$^{15}$A referee rightly points out that the UG frame not just eliminates the moral cost of indulging in corruption but replaces it with the moral cost associated with ultimatum games. It is governed by its own moral frame driven by a set of non-monetary motivations which are distinctly different from the one in the BG frame.
1(a) BG
Citizen takes a test
Public official grades
Performance above 10
PO decides whether to ask for a bribe and if yes how much to ask for.

$\{0, 20, 40, \ldots, 400\}$

Accept
Reject

C: $200 + 400 - b$
P: $600 + b$

C: $200$
P: $600$

1(b) UG
Participant A takes a test
Participant B grades
Performance above 10
P-A qualifies for the next part of the game where P-B decides how much to share with P-A.

$\{0, 20, 40, \ldots, 400\}$

Reject
Accept

P-A: $200 + 400 - b$
P-B: $600 + b$

P-A: $200$
P-B: $600$
neutrality. Additionally the expectations of the subjects and their psychological reference points, are now different owing to a change in the entitlement since UG triggers social norms about sharing which BG triggers social norms of not taking a bribe. A comparison of behavior in BG and UG can help us answer whether or not moral costs are at play in BG.

3.1.3 External Validity

At the end of both treatments, the Acceptance/Rejection decisions are handed out in the BG and UG treatments, subjects in the role of P-Bs/public officials are asked to mentally compute their earnings in Mohars and then convert them into Rupees at a previously declared exchange rate. Several examples of how to calculate the earnings are provided at this point. Each subject then finds an envelope underneath his desk, consisting of Rs. 10 currency notes\textsuperscript{16}. Subjects are asked to take out the money that they have earned and leave the rest in the envelope. After these instructions, the experimenters leave the room. The subjects then take their earnings and leave the room one by one\textsuperscript{17}.

Subsequently we match the amount of money left in the envelopes with the response sheets and exit surveys. This allows us to see if a subject has taken more money than his earnings. The objective of this exercise is to identify two behavioral correlates - first, between the decision to seek a bribe and the decision to steal more money from the envelope. Second, between the amount demanded as bribe and the amount of money stolen. Thus its a unique setting where corruption behavior in a framed lab behavior can be juxtaposed with stealing behavior.

3.1.4 Experimental Procedure

The experiment were conducted in Kolkata, India, with student subjects recruited from Jadavpur University and Presidency University. The age and background profiles of these two institutions are very similar. A total of two hundred and forty students were recruited and each subject participated in only one role of one treatment. Sessions were randomly assigned UG and BG treatments and subjects were randomly assigned roles.

Each session consisted of twenty subjects with ten in each role. Eight such sessions were conducted with four each for UG and BG treatments. The subjects were asked to report to two different rooms physically separated by a fair distance in order to maintain anonymity. The instructions for the respective roles were read out (See Appendix 4). Several examples were worked out and earnings of each role were spelt out both in Mohars and Rupees. Answer sheets from the

\textsuperscript{16}The information that an envelope filled with money is placed underneath the desk is revealed only at the end when they are about to receive their earnings. So, this does not have any impact on the earlier behavior.

\textsuperscript{17}One referee points out that this in the true sense is not equivalent to observing the subjects in the field. Though they were not observed by the experimenter, they were still in a lab. We acknowledge that this is not the best design to capture external validity of unethical behavior and perhaps that is why the results are, at best, weak. However, subjects felt unobserved and it was not obvious to them that the remaining number of notes would be matched back to their responses.
Citizens or Participant As were then handed out to the matched Public Officials or Participant Bs who graded them with the help of a solution manual provided to them. While the Citizens and P-As filled out the short exit survey, their answer sheets were graded by the POs/P-Bs. After grading, they wrote their decisions about bribe or transfer in the response sheets. These response sheets were subsequently returned to the other room for acceptance or rejection by Citizens/Participant As. Following this, the Public Officials and P-Bs filled out their exit survey.

The Citizens or the Player As were paid in cash. The Public Officials or Player Bs were asked to calculate their earnings after letting them have the response sheets back. Then they were asked not to communicate among themselves and to take their earnings from the envelopes which were placed underneath their desks, following which the experimenters left the room.

The sessions lasted for approximately one hour. The instructions were read out in English and then paraphrased in Hindi and Bengali. All payoffs for this experiment were stated in terms of a fictitious currency called *Mohar*. At the end of the experiment participants exchanged their earnings for Rupees at the rate 100 *Mohars*=Rs. 50. The earnings ranged from Rs. 100 to Rs. 450 with an average of Rs. 203 (~ 6 USD in Purchasing Power Parity terms).

### 3.2 Second Experiment

#### 3.2.1 Design

This experiment elicits social appropriateness of the UG and BG treatments using a tool, developed by Krupka and Weber (2012). The subjects are rewarded if their appropriateness ratings match the ratings given by most other people in the room. The reward, attached to the mode of the distribution of ratings, helps elicit a second order belief about the societal (defined as the others in the classroom) viewpoint. We interpret this as the social norm governing the situation.

Subjects are shown either the situation in the BG treatment or the one in the UG treatment. Then they were asked to rate each action for each role as one of the following - “very socially inappropriate”, “socially inappropriate”, “somewhat socially inappropriate”, “value neutral”, “somewhat socially appropriate”, “socially appropriate” and “very socially appropriate”. The responses were later converted into numerical scores of -3, -2, -1, 0, +1, +2 and +3 respectively.

Each subject is paid a participation fee of Rs. 110. The action space of the public official/P-B is divided into several categories depending on the level of bribe. In the BG situation subjects are asked to report their social appropriateness rating for the actions of the public official as well as for actions of the citizen in regard to accepting or rejecting the bribe demand for all possible bribe categories. One category from the public official’s strategy space and one from the citizen’s strategy

---

18 A unit of exchange used in medieval India, an etymological history of which can be traced back to the persian word *mohr*, meaning seal.

19 While both the situations in BG and UG are governed by injunctive norms, i.e. what one should or should not do, the nature injunction that works for the two frames are very different.

20 Figure 5 and Table 4 in Appendix 2 lays out the categories.
space are randomly selected. The modal response of the appropriateness rating for the randomly selected categories is noted. If a subject’s appropriateness rating for the randomly selected category for either the public official or the citizen is the same as the modal response, then she is paid double the participation fee, i.e. Rs. 220. The description of the UG situation is identical except that subjects are asked to rate the actions of P-A and P-B.

For the second experiment, a different set of subjects were invited to a classroom. A session described either the BG environment or the UG. Each session comprised of around 20 subjects and four such sessions were conducted. A total of 40 subjects participated in the norm elicitation of BG and another 40 in that of UG. We implemented an entirely between subject design in order to avoid potential experimenter demand effect and confounds arising from subjective assessments of the situations i.e. none of the subjects who participated in the UG and BG treatments, took part in the social norm treatment. The instructions were read out in English and then paraphrased in Hindi and Bengali. Several examples were given to demonstrate the idea of social appropriateness.

4 Results

Since we were interested in the phenomenon of harassment bribery, the general design of the experiment and the task in particular were so calibrated that the majority of the subjects crossed the threshold. Three subjects in the UG and two in the BG treatment scored less than 10. However 38 pairs of subjects in each treatment fulfilled this criteria as one P-A was wrongly classified as having crossed the threshold.\(^{21}\) In the following discussion (but not in the actual experiment) we use the word *ultim* to denote the UG treatment equivalent of the bribe i.e. the amount that P-B proposes to keep for himself.

4.1 Framing Effects (Experiment 1)

We find significant differences across treatments on both in terms of the proportion of subjects who demanded a bribe/*ultim* and the amount of bribe demanded. We also find important differences in the acceptance rate between the two treatments. As Fig. 2(a) shows, all the subjects playing the role of Participant B propose a split with a positive amount for themselves while only 78.9% of the Public Officials demand a bribe. Table 3 in Appendix 1 reports the *p*-values, clustered over sessions and rejects the null of equality of proportions (\(\chi^2\) test, *p*-value=0.07). The mean *ultim* proposed was \(\text{M}244.2\) and this was significantly higher than the mean bribe demanded which was \(\text{M}206\) (Mann-Whitney test, *p*-value=0.08).

\(^{21}\)Presumably because the P-B realized that he could increase his own earnings by letting the P-A pass.
In Fig. 2(b) we investigate the acceptance rate of the Citizens and Participant Bs. While 97% accepted the division proposed in the UG treatment, only 76% accepted to pay the bribe (if one was demanded). The difference is statistically significant ($\chi^2$ test, clustered $p$-value=0.03) as reported in Table 3 in the Appendix.\(^{22}\) Furthermore, though the average ultim proposed was larger than the average bribe demanded, only one offer was rejected in UG (where ultim proposed was 400); on the other hand bribes which were rejected had a mean of M237. Thus, when posited in terms of bribe much lower amounts were rejected in the BG treatment.\(^{23}\)

We shall explain the difference in the treatment effects in terms of the difference in the fairness perception of the two treatments in Section 4.3.

### 4.2 External Validity

Findings, reported in Table 1, are inconclusive but indicate a (weak) positive correlation. The correlation between the binary choice decision to ask for a bribe and steal is 0.22 and that between bribe amount and stolen amount is 0.3. While statistically insignificant, it is similar to the correlations reported in other studies which look at external validity. For example, Baran et al. (2010) report correlation of 0.29 between field setting and lab while Englmaier and Gebhardt (2011) and Franzen and Pointner (2013) report 0.4 and 0.2, respectively. In fact Franzen and Pointner (2013) conclude by saying that despite the weak or moderate association, their results should still be viewed as the glass half full (i.e. lab and field behaviors are more or less similar).

\(^{22}\)In Section 4.4, we control for it in the regressions and still find that the acceptance rate is significantly higher in the UG treatment than in the BG.

\(^{23}\)Lack of sufficient number of observations make mean comparison test redundant.
A further probe reveals that there is a strong negative correlation between stealing and acceptance to pay a bribe (Spearman Corr. Coeff=-0.61, p-value=0.00). This indicates that we picked up an effect where the probability of stealing increases with the refusal on the part of the citizen to pay a bribe, i.e. unethical behavior at one level crowds out that in another. We interpret these correlations as that the lab corruption game is appealing to the same unethical behavior among subjects as the stealing\(^{24}\).

Table 1: External validity

<table>
<thead>
<tr>
<th>Variable 1</th>
<th>Variable 2</th>
<th>Correlation Coefficient</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1(bribe&gt;0 )</td>
<td>1(Stealing&gt;0)</td>
<td>0.22</td>
<td>0.177</td>
</tr>
<tr>
<td>Successful bribe extraction</td>
<td>1(Stealing&gt;0)</td>
<td>-0.61***</td>
<td>0.000</td>
</tr>
<tr>
<td>Bribe amount</td>
<td>bribe&gt;0 Stolen amount</td>
<td>0.30</td>
<td>0.150</td>
</tr>
</tbody>
</table>

Note: *** denotes significance at the 1% level, ** at 5% level and * at 10% level. Spearman correlation coefficient for binary variables. 1(.) denotes identity variable.

4.3 Social Norms (Experiment 2)

Our discussion above in Section 3 conjectured that the two treatments, though strategically identical, were governed by different perceptions of social appropriateness. In particular we argued that an act of demanding a bribe may be considered more socially inappropriate than that of proposing to keep a comparable amount in UG. Further, accepting to pay a bribe would be considered more socially inappropriate in the former than accepting a proposed division in the latter.

The social appropriateness of UG was elicited from 40 subjects and that of BG from another 40. Figure (4) in Appendix 2 gives a visual description of how the distribution of norm rating changes across the treatments. Table 4 lays out the full distribution of the ratings corresponding to each action of Participant B and the Public Official in UG and BG, respectively. It also presents the mean ratings for both treatments and the p-values reported by the Mann-Whitney ranksum test for the difference in mean ratings.

While in the following analysis we focus on the mean appropriateness ratings, replacing the mean with median or mode does not make a difference to the results. The results show that for low levels, bribe demanded is considered socially more appropriate than comparable amount of ultim. However for medium and higher levels, bribe is seen to be socially more inappropriate than ultim. This clearly indicates that social sanction against a small bribe amount is nearly absent. In the UG treatment an equal split ranks high on the social appropriateness scale and is significantly

\(^{24}\)Direct stealing of cash and deliberately failing to return excess money are two alternative ways in which unethical behavior can be examined on the field however one may note that these two methods appeal to different degrees of unethicality. Our conjecture is that the latter will result in fewer observations in the false-positive domain. Also note that none of our other results hold when stealing is used as the dependent variable.
higher than that of a bribe demand of M200 (MW test, p-value = 0.00). No statistical difference is observed for higher levels of ultims or bribes however.

While the left hand panel of Figure (3a) shows norm ratings for the two treatments and their difference, the right hand panel shows the distribution of actual behavior. A quick glance confirms the co-movement between the difference in the norm ratings demonstrated by the area plot in gold on the left and the difference in actual behavior given by the gold bars. The non-parametric correlation between the difference in the mean norm rating and the difference in actual behavior is positive and significant (Spearman Corr. Coeff=0.78, p-value=0.01), reaffirming the co-movement discussed above.

The mean social appropriateness rating of accepting the proposed split is shown in Fig.(3b). Accepting to pay a small bribe is considered socially appropriate but the inappropriateness steadily increases with the bribe amount. For the UG treatment the appropriateness rating for accepting an equal split is highest and then it declines though it remains significantly higher than that of the BG treatment for equal amounts. To summarize, the two frames elicit different perceptions of appropriateness - equal sharing is considered very appropriate in UG but demanding no or a small amount of bribe is considered appropriate in BG, though the two games are strategically identical. It is also worth noting that the fact that small bribes elicit high appropriateness measures implies that there is little social sanction associated with smaller levels of corruption in India where the experiment is conducted. This may give us a clue about why lack of social sanction against petty bribery helps it perpetuate in other similarly placed countries.

4.4 Regression results

The main result, laid out in Table 2, is consistent with the result from the mean difference test discussed above as we see that the bribe amount in the BG treatment is M42 lower than the ultim demanded in the UG treatment even after controlling for caste, age, gender and log of household income (for a study of how cheating behavior depends on demographic characteristics, see Bucciol et al. (2013)). This difference is significant as indicated in specification (2). Furthermore, the marginal effects from the estimated probit model for acceptance of a bribe/ultim in (3) reveal that the probability of accepting to pay a bribe is 0.29 lower than that of accepting to share an ultim. Specification (4) finds that the probability of acceptance is lower by 0.29 in the BG treatment even after controlling for the effect due to bribe/ultim amount, which turns out to be marginal.\footnote{We thank an anonymous referee for suggesting this specification.}

Besides identifying the treatment effect, it is also important to show how actual behavior is dependent on perceived social appropriateness. In Section 4.3 we have discussed how the difference in norm measures in UG and BG are significantly correlated with the difference in actual behavior. In column (5), the frequency of observations in each category (the categories are given in Fig.4 in
(a) Comparing actual demand of Public Official and Participant B with norms

In Figure 4(a), the left hand side panel compares the mean norm rating for equivalent amounts in UG and BG. The y-axis represents the range of possible norm ratings, i.e. -3 represents very socially inappropriate and 3 represents very socially appropriate. The shaded area represents the difference. The right hand side panel compares actual frequency of choices for equivalent amounts in UG and BG. The shaded bars represent the difference in frequencies between the two treatments for comparable amounts.

(b) Comparing actual acceptance of Citizens and Participant A with norms

Figure 4(b) compares the norm ratings corresponding to acceptance of a demand between UG and BG. The shaded area represents the difference.
Appendix 2) is regressed against the respective elicited social appropriateness measure. Notice that despite the between subject design in eliciting social appropriateness, we do find clear evidence that the frequency of observation in each category increases with social appropriateness, indicating that views of how appropriate an action is, dictate how many or how few people go for it. This effect is, in some ways, a weaker version of earlier evidence of contagion effects as found in Innes and Mitra (2013). While contagion effects demonstrate herd behavior and hint at diminishing marginal cost at work by following others, our results show that there is indeed a set of normative codes governing our actions. This works quite independently of the information about what others do.

Table 2: Regression Results

<table>
<thead>
<tr>
<th>Amount demanded (ultim/bribe)</th>
<th>Acceptance (ultim/bribe)</th>
<th># of obs. in each bin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
<td>Probit</td>
</tr>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Bribery Game Treatment</td>
<td>-38.2***</td>
<td>-42.46**</td>
</tr>
<tr>
<td></td>
<td>(16.38)</td>
<td>(18.26)</td>
</tr>
<tr>
<td>Controls</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Amount demanded</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>SocApp</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>68</td>
<td>68</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.85</td>
<td>0.19</td>
</tr>
</tbody>
</table>

Note: *** denotes significance at the 1% level, ** at the 5% level and * at the 10% level. The reported $R^2$ for probit models is Pseudo $R^2$. The robust standard errors reported in the parentheses are clustered by sessions (except model (1)). The Bribery Game Treatment takes value 1 for the bribery game (BG) and 0 for the ultimatum game (UG). Control variables: gender, age, caste, log of household income, constant. The variable caste takes value 0 for upper caste and 1 for backward castes. Column (3) and (4) report the marginal effects following probit estimations. In column (5), the number of observations in each category (see Figure (4) in Appendix 2 for the categories and the mean norms in each category) is regressed on the mean social norm rating for each category.

5 Discussion and Conclusion

The canonical structure of a laboratory corruption game essentially implements splitting of a surplus which is generated when those in positions of power (e.g. public officials) unlawfully favor
those in need of services (e.g. firms or citizens), for instance look at Abbink et al. (2002), Cameron et al. (2009), Barr and Serra (2010) and more recently Gneezy et al. (2013). To this skeletal, parameters are then altered, layers added and context changed. The harassment bribery game, which we implement in this paper, remains true to this canonical structure where a surplus is split. We compare the behavior of subjects in the harassment bribery game and a strategically identical but differently framed ultimatum game.

Our results show that the two different moral frames trigger different responses among individuals. Not only do less people indulge in bribery, the average bribe demanded turns out to be significantly lower in the bribery game treatment than the average amount proposed for oneself in an ultimatum game. Furthermore, the patterns of acceptance to pay the bribe and to accept the split proposal vary widely as well. This points to the fact that moral costs are indeed at work when subjects operate under a corruption frame.

We then explain the behavioral differences observed in the two frames in terms of a measure of social appropriateness associated with each action in the two frames. Social appropriateness or commonly held perceptions about what is the right thing to do, successfully explain the pattern of the actual choices made in the two frames. Also, according to elicited social appropriateness the bribery frame is considered more inappropriate than the counterfactual.

In the light of the fact that past studies have not found decisive evidence of moral costs at play in laboratory corruption games, it will be interesting to see if the norm elicitation exercise can detect any significant difference in perceived norms for the loaded and neutral frames, as used in the past by Abbink and Hennig-Schmidt (2006) and Barr and Serra (2009). If not then this may further explain their inability to find consistent framing effects. In this paper we alter the psychological reference point of the subjects across the treatments through a subtle change in the sense of entitlement, while keeping the treatments otherwise strategically identical, and we find strong treatment effects. We then argue that in order to see evidence of moral cost at work, the appropriate counterfactual is one which uses neutral language and suitably alters the reference point of the subjects. It is the latter which helps create the alternative moral framework.

Having said that, we recognize that there are other differences between our set up and those of Abbink and Hennig-Schmidt (2006) and Barr and Serra (2009), which can potentially affect the results. For instance they conduct their experiments in western societies where corruption norms are very different from those in India where our experiment was conducted. Although the game we implement is sufficiently generic and true to the fundamental canonical structure of corruption games, it is not identical to the one which they implement. We implement a two person harassment bribery game unlike theirs which is a collusive bribery design. Barr and Serra (2009) in fact has a third person who is affected by a corrupt transaction. Despite the differences, we are tempted to believe that the hook of our results comes from a successful creation of an alternative moral framework, induced by the difference in the sense of entitlement. Interestingly, entitlement effects have also been observed in other strands of the literature, for instance Hoffman et al. (1994), Oxoby
and Spraggon (2008) and Ruffle (1998) in dictator games and Hoffman et al. (1996) in ultimatum games. However, just like in corruption games, merely changing the frame by a different language has been found to have no effect in dictator games (Dreber et al., 2013). Hence, our conclusion that in order to see a framing effect in corruption games, the very psychological reference point should be altered, holds true elsewhere in the literature as well.

The successful creation of an alternative moral framework is also indicated by the different societal perceptions of what is morally the right thing to do in the two environments. A clear evidence of moral cost at play along with evidence of moral indictment from society shows that laboratory corruption games do measure what they are set out to. This also ensures that this tool can be used to measure the change in unethical behavior by changing a relevant policy variable.

Our attempt to relate behavior in a lab with that in a field is met with limited success. Imaginative designs in the future may be able to do that more clearly. Furthermore, we find evidence of lack of social sanction against petty bribery. Besides institutional inadequacies, this may explain why the bad equilibrium of pervasive corruption is difficult to escape in many societies.

Finally, it is often argued, especially in the theoretical models which analyze corruption equilibrium in a rational actor framework, that our pursuit of ethicality is driven largely by self interest. Within the framework of this premise it becomes difficult to explain why the vast majority among us choose legitimate actions when a large number of dubious but materially better off alternatives are present. It seems unlikely that such a voluntary abdication of unethical behavior will be part of the resulting equilibrium without some norm induced behavior. One may conjecture that the complex web of laws and rules, which govern the modern society, may not be sustained if we are to act only with strategic selfishness vis-a-vis these laws. In other words, our actions are disproportionately influenced by what we and others think is morally the right thing to do. Our experiment bears a testimony to that.

References


Armantier, Olivier and Amadou Boly, “Comparing Corruption in the Laboratory and in the Field in Burkina Faso and in Canada,” *The Economic Journal*, 2013, pp. n/a–n/a.


## Appendix 1

### Table 3: Mean Difference

<table>
<thead>
<tr>
<th></th>
<th>Mean BG</th>
<th>Mean UG</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PO/B’s Decision to ask for a bribe/ultim</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>0.790</td>
<td>1.000</td>
<td>0.21*</td>
</tr>
<tr>
<td>p-value ($\chi^2$ test)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clustered p-value ($\chi^2$ test)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>C/A’s Decision to accept the offer</strong></td>
<td></td>
<td></td>
<td>0.211**</td>
</tr>
<tr>
<td>N</td>
<td>0.763</td>
<td>0.975</td>
<td></td>
</tr>
<tr>
<td>p-value ($\chi^2$ test)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-value (Clustered $\chi^2$ test)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Bribe/Transfer Amount (Full sample)</strong></td>
<td></td>
<td></td>
<td>77***</td>
</tr>
<tr>
<td>N</td>
<td>162.6</td>
<td>244.2</td>
<td></td>
</tr>
<tr>
<td>p-value $t$-test</td>
<td></td>
<td></td>
<td>0.000</td>
</tr>
<tr>
<td>p-value clustered $t$-test</td>
<td></td>
<td></td>
<td>0.001</td>
</tr>
<tr>
<td>p-value Rank Sum tests</td>
<td></td>
<td></td>
<td>0.002</td>
</tr>
<tr>
<td><strong>Bribe/Transfer Amount (Restricted sample)</strong></td>
<td></td>
<td></td>
<td>38.2</td>
</tr>
<tr>
<td>N</td>
<td>206</td>
<td>244.2</td>
<td></td>
</tr>
<tr>
<td>p-value $t$-test</td>
<td></td>
<td></td>
<td>0.023</td>
</tr>
<tr>
<td>p-value clustered $t$-test</td>
<td></td>
<td></td>
<td>0.135</td>
</tr>
<tr>
<td>p-value Rank Sum tests</td>
<td></td>
<td></td>
<td>0.084</td>
</tr>
</tbody>
</table>

Note: *** denotes significance at the 1% level, ** at the 5% level and * at the 10% level from clustered $t$-test. Clustering was done for each session. The acceptance decision reported above is conditioned on being asked for a bribe/transfer. The restricted sample consists of only those who demanded a bribe.
Appendix 2

Figure 4: Distribution of norm rating for each category of ultim/bribe in UG and BG

Note: The figure maps the entire distribution of the social norm ratings for UG and BG treatments. For example, the green box plot corresponds to the norm ratings for a bribe/ultim demand between 60 and 100 (from the legend). The social norm ratings for that amount of ultim (bribe) may be read from the left (right) panel. The diamond marks the median of the distribution of ratings for each category. The norm ratings of the UG treatment is an inverted U with the equal split considered as most appropriate. The norm ratings of the BG treatment is a monotonic decline indicating that social appropriateness goes down with the bribe amount.
Table 4: Frequency distribution of norm ratings

<table>
<thead>
<tr>
<th>Bribe/Ultim</th>
<th>-3</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>Mean UG</th>
<th>-3</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>Mean BG</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td>55</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>2.5</td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td>17.5</td>
<td>0</td>
<td>-1.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10-50</td>
<td>35</td>
<td>20</td>
<td>10</td>
<td>5</td>
<td>12.5</td>
<td>5</td>
<td>12.5</td>
<td>-0.95</td>
<td>7.5</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>2.5</td>
<td>85</td>
<td>2.38</td>
<td>0.00</td>
</tr>
<tr>
<td>60-100</td>
<td>12.5</td>
<td>32.5</td>
<td>10</td>
<td>5</td>
<td>10</td>
<td>30</td>
<td>0</td>
<td>-0.43</td>
<td>7.5</td>
<td>0</td>
<td>12.5</td>
<td>15</td>
<td>10</td>
<td>35</td>
<td>20</td>
<td>1.05</td>
<td>0.00</td>
</tr>
<tr>
<td>110-150</td>
<td>5</td>
<td>20</td>
<td>15</td>
<td>12.5</td>
<td>17.5</td>
<td>20</td>
<td>10</td>
<td>0.18</td>
<td>0</td>
<td>10</td>
<td>25</td>
<td>10</td>
<td>20</td>
<td>35</td>
<td>0</td>
<td>0.45</td>
<td>0.03</td>
</tr>
<tr>
<td>160-200</td>
<td>0</td>
<td>5</td>
<td>2.5</td>
<td>25</td>
<td>10</td>
<td>30</td>
<td>27.5</td>
<td>1.4</td>
<td>0</td>
<td>30</td>
<td>17.5</td>
<td>27.5</td>
<td>12.5</td>
<td>7.5</td>
<td>5</td>
<td>-0.35</td>
<td>0.00</td>
</tr>
<tr>
<td>210-250</td>
<td>0</td>
<td>0</td>
<td>25</td>
<td>5</td>
<td>20</td>
<td>22</td>
<td>27.5</td>
<td>1.23</td>
<td>15</td>
<td>35</td>
<td>32.5</td>
<td>5</td>
<td>12.5</td>
<td>0</td>
<td>0</td>
<td>-1.35</td>
<td>0.00</td>
</tr>
<tr>
<td>260-300</td>
<td>5</td>
<td>25</td>
<td>17</td>
<td>0</td>
<td>30</td>
<td>17.5</td>
<td>5</td>
<td>-0.03</td>
<td>27.5</td>
<td>32.5</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1.88</td>
<td>0.00</td>
</tr>
<tr>
<td>310-350</td>
<td>25</td>
<td>50</td>
<td>15</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1.9</td>
<td>52.5</td>
<td>47.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-2.53</td>
<td>0.01</td>
</tr>
<tr>
<td>360-390</td>
<td>52.5</td>
<td>35</td>
<td>7.5</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>-2.3</td>
<td>97.5</td>
<td>2.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-2.98</td>
<td>0.00</td>
</tr>
<tr>
<td>400</td>
<td>97.5</td>
<td>2.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-2.98</td>
<td>92.5</td>
<td>2.5</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-2.68</td>
<td>0.29</td>
</tr>
</tbody>
</table>

Mann Whitney Rank Sum test reported p-value of the mean difference. Shaded cells for each category in each treatment denote the modal ratings.
Appendix 3

Figure 5: Diagnostic tests for the regression in column (5) in Table (2)

A small sample regression, like the one in Column (5) in Table 2, produces biased estimates. To the extent that the distribution of the data can help, the quantile plots of the frequency of observations in each bin and the mean of social norms rating for each bin are given below. The distribution of residuals from the regression in Column (5) is also included. The result of the outlier test, based on inter-quartile range, shows the presence of one outlier. Re-estimating the specification given in column (5) and excluding the outlier does not lead to any significant difference.
Appendix 4 (Instructions)

1 Instructions (Public Officials in BG)

1.1 Introduction

You are now taking part in an economic decision making study. During the experiment you can earn money by receiving a fictitious currency called “Mohar”. We will pay you 200 Mohars for participating but you can earn additional money depending on the decisions you and the others make. All Mohars that you earn in the experiment will be exchanged into Rupees at the end of the experiment. The exchange rate is:

$$1\text{ Mohar} = Rs \ 0.50$$

The experiment consists of several parts and a survey. Please follow the instructions carefully. If you have any questions please let us know by raising your hand. Your question will be answered by us in private.

Please note that communication between participants is strictly prohibited during the experiment. Further instructions will be provided at the beginning of each step of the experiment.

We will give each one of you an identity number. Please do not lose your identity number. This entire experiment is anonymous. We will sometimes form pairs of two participants. The matching of two participants has been randomly determined in advance. You will never be informed of the identity of the participant with whom you have been matched.

Please raise your hand if you have any question.

1.2 The Game

We have divided the total number of participants in this experiment session randomly into two equal groups: Citizens and Public Officials. All the participants in this room are Public Officials. All the Citizens are located in another room.

In this part each one of you is a “Public Official” and you are paired with a “citizen” who is in the other room. Each one of you will receive a booklet filled out by the citizen you are paired with. Nobody will ever be informed of the identity of the citizen he/she is paired with. Your task is to count the number of correct answers in their citizen’s booklet. In order to do it, we provide you with the “SOLUTION MANUAL” which contains the list of correct answers to the problems.

The Citizen’s Booklet

The citizen’s booklet consists of simple math problem of the following type. It has boxes like this:

| 1.79 | 3.70 | 2.99 |
| 8.34 | 7.19 | 5.55 |
| 9.01 | 4.45 | 6.32 |

The citizen has to find a pair of numbers in the box which add up to 10. Having found the pair, they are required to encircle the corresponding numbers and put a tick into the box corresponding to “Got it”. In this example note that only 4.45 and 5.55 add up to 10. The booklet will contain 20 such boxes. The identity number of the citizen is written on top of the booklet.
Your Task

The SOLUTION MANUAL given to you contains the correct answers to the problems. In each box the correct answers are underlined and bold. Your role is to find the number of correct answers in the citizen’s booklet with the help of this SOLUTION MANUAL.

Description of the Game

Now please take a look at the figure attached to the instructions. You will be paid 400 Mohars for correcting the citizen’s booklet in addition to the participation fee of 200 Mohars. If the citizen solves at least ten box correctly then she “passes” the test and is entitled to a payment of 400 Mohars in addition to the participation fee of 200 Mohars. In case she is able to solve only less than ten boxes then she is eligible for a payment of 200 Mohars only and the game ends. However even if the citizen passes the test, you may ask from the citizen a bribe. Notice that if the citizen “passes” then she is eligible for a payment of 400 Mohars more than if she had not “passed”. So you can ask for a maximum bribe of 400 Mohars. You may also choose not to ask for a bribe. If you do ask for a bribe the citizen may accept it or reject it. If she accepts then your earning equals the bribe amount plus 600 Mohars. The citizen’s earning is 600 Mohars less the amount of bribe she paid. If she rejects then she earns only 200 Mohars and your earning is 600 Mohars. Note that you can ask bribe in multiples of 20 Mohars only.

Citizen scores 13 in the matrix task. She is entitled to a payment of 600 Mohars. You demand a bribe of 200 for yourself. Citizen accepts it. Your earning is 600+200=800 Mohars. Her earning is 600-200=400 Mohars.

Citizen scores 13 in the matrix task. She is entitled to a payment of 600 Mohars. You demand a bribe of 350 for yourself. Citizen rejects it. Your earning is 600+0=600 Mohars. Her earning is 200+0=200 Mohars.

You can observe how much you earn for each of your decision from the figure below.

Game Begins

Now please grade the answer sheet.

You will find a response sheet on the desk for you to fill up now.

Please mark whether Citizen has solved at least 10 boxes correctly or not.

Please write down how much, if at all, you want to demand as bribe.

Please put the paper in the envelope and seal it. We will collect envelopes from you one by one and hand it in to the corresponding Citizen.

If you have any questions regarding these instructions, please raise your hand. We will answer your questions in private.

In the meantime please fill out the survey questionnaire.

The response of the citizens have arrived. Check the envelope to see whether she has accepted your demand. Now calculate the total number of mohars that you have earned. In another envelope placed underneath the desk, you will find some money. You can take the money that you have earned. Remember that the exchange rate is 1 Mohar= Rs. 0.50.

You may leave the room now.
(Figure 1(a) was included)
2 Instructions (Citizens in BG)

2.1 Introduction

You are now taking part in an economic decision making study. During the experiment you can earn money by receiving a fictitious currency called “Mohar”. We will pay you 200 Mohars for participating but you can earn additional money depending on the decisions that you and the others make. All Mohars that you earn in the experiment will be exchanged into Rupees at the end of the experiment. The exchange rate is:

\[
1\text{ Mohar} = Rs\ 0.50
\]

Please follow the instructions carefully. If you have any questions please let us know by raising your hand. Your question will be answered by us in private.

In this experiment you will need to solve few math problems without the help of any electronic device. Please note that communication between participants is strictly prohibited during the experiment.

We will give each one of you an identity number. Please do not lose your identity number. This entire experiment is anonymous. We will sometimes form pairs of two participants. The matching of two participants has been randomly determined in advance. You will never be informed of the identity of the participant with whom you have been matched.

2.2 The Game

We have divided the total number of participants in this experiment session randomly into two equal groups: citizens and public officials. All the participants in this room are citizens. All the public officials are located in another room.

In this part each one of you is a “citizen” and you are paired with a “public official” who is in the other room. Nobody will be informed of the identity of the public official with whom he/she is paired. You will receive a booklet which you are required to fill out. The answer booklet contains twenty boxes. In each of the box at least two numbers add up to 10. Your task is to find the two numbers.

Answer Booklet

The answer booklet consists of simple math problems of the following type. It has boxes like this:

\[
\begin{array}{ccc}
1.79 & 3.70 & 2.99 \\
8.34 & 7.19 & 5.55 \\
9.01 & 4.45 & 6.32 \\
\end{array}
\]

You have to find a pair of numbers in the box which add up to 10. Having found the pair, you are required to encircle the two corresponding numbers. In this example note that only 4.45 and 5.55 add up to 10. The booklet will contain 20 such boxes. You will have 10 minutes to complete the boxes. Please write your id on top of the page.
Description of the Game

The public official will be paid 400 Mohars for correcting the citizen’s booklet in addition to a fixed payment of 200 Mohars i.e. his total earnings are 600 Mohars. If you solve at least ten boxes correctly then you “pass” the test and you are entitled to a payment of 600 Mohars i.e. 400 Mohars in addition to the fixed participation fee of 200 Mohars. In case you are able to solve less than ten boxes then you are eligible for the participation fee of 200 Mohars only and the game ends. However if you pass then a public official may ask from you a “bribe” before letting you have the entitlement. Notice that if you “pass” then you are eligible for a payment of 400 Mohars more than if you do not “pass”. So the public official can ask for a maximum bribe of 400 Mohars. He may also choose not to ask for a bribe. If he does ask for a bribe, you may accept it or reject it. If you do accept it then your earning is 600-bribe Mohars and the public official’s earning is 600+bribe Mohars. If you reject it then you earn only 200 Mohars and the public official’s earning is 600 Mohars.

You score 13 in the matrix task. You are entitled to a payment of 600 Mohars. Public official demands a bribe of 200 for himself. You accept it. Public Official’s earning is 600+200=800 Mohars. Your earning is 600-200=400 Mohars.

You score 13 in the matrix task. You are entitled to a payment of 600 Mohars. Public official demands for a bribe of 350 for himself. You reject it. Public Official’s earning is 600+0=600 Mohars. Your earning is 200+0=200 Mohars.

You can observe how much you earn for each of your decision from the figure below.

Game Begins

Now please start solving the box booklet. You have fifteen minutes and your time starts now.

Please stop writing now and hand over the answer sheets.

Now please fill out the exit survey.

Now that the response sheet from the public official is back. Please mark whether you accept or reject to pay the bribe amount demanded on the response sheet.

Please hand in your response sheets one by one and receive your payments now.

(Figure 1(a) is included here)

3 Instructions (Participant B in UG)

3.1 Introduction

You are now taking part in an economic decision making study. During the experiment your payoffs will be stated in terms of a fictitious currency called “Mohar”. You will receive 200 Mohars for participating but you can earn additional money depending on the decisions you and the others make. You will be able to exchange all Mohars into Rupees at the end of the experiment. The exchange rate at which you can convert Mohars into Rupees is:

\[ 1\ Mohar \ = \ Rs \ 0.50 \]

Please follow the instructions carefully. If you have any questions please let us know by raising
your hand. Your question will be answered by us in private. Please note that communication between participants is strictly prohibited during the experiment.

We will give each one of you an identity number. Please do not lose your identity number. This entire experiment is anonymous. We will sometimes form pairs of two participants. The matching of two participants has been randomly determined in advance. You will never be informed of the identity of the participant with whom you have been matched.

3.2 The Game

We have divided the total number of participants in this experiment session randomly into two equal groups: Group A and Group B. All the participants in this room are participants of Group B. All the participants of Group A are located in another room.

In this part each one of you is a “Participant B” and you are paired with a Participant A who is in the other room. You will receive a booklet filled out by the Participant A with whom you are paired with. Nobody will be informed of the identity of the Participant A he/she is paired with. Your task is to count the number of correct answers in Participant A’s booklet. In order to do so, we provide you with a “SOLUTION MANUAL” which contains a list of correct answers to the problems.

Participant A’s Booklet

Participant A’s booklet consists of simple math problem of the following type. It has boxes like this:

```
  1.79  3.70  2.99
  8.34  7.19  5.55
  9.01  4.45  6.32
```

Participant A has to find a pair of numbers in the box which add up to 10. Having found the pair, they are required to encircle the corresponding numbers and put a tick into the box corresponding to “Got it”. In this example note that only 4.45 and 5.55 add up to 10. The booklet will contain 20 such boxes and they have 15 minutes to do the task. The identity number of Participant A is written on the top of the booklet.

Your Task

The Solution Manual given to you contains the correct answers to the problems. In each box the correct answers are underlined and bold. Your role is to find the number of correct answers in the Participant A’s booklet with the help of this Solution Manual.

Description of the Game

You will be paid 400 Mohars for correcting Participant A’s answer booklet in addition to the fixed participation payment of 200 Mohars i.e. you will receive a total of 600 Mohars. If Participant A solves at least ten boxes correctly then she “passes” the test and becomes eligible to take part in next part of the game. If Participant A is unable to solve at least ten boxes correctly then she “fails” the test and will not be able to take part in the next part of the game and the game ends.

Suppose that Participant A solves at least ten boxes correctly and is able to take part in the next part. In this part, you may propose to split 400 Mohars between her and yourself. For example, if you propose to keep for yourself $x$, then Participant A is allocated $400 - x$. However note that...
Participant A may Accept or Reject the proposal suggested by you. If Participant A accepts your proposed division then both of you will earn as per your proposal. If participant A rejects then she gets 200 Mohars i.e. the participation payment only and you earn 600 Mohars.

Participant A scores 13 in the matrix task. She qualifies for the next part. You propose a split of 200 for yourself and 200 for her. Participant A accepts it. Your earning is 600+200=800 Mohars. Her earning is 200+200=400 Mohars.

Participant A scores 13 in the matrix task. She qualifies for the next part. You propose a split of 350 for yourself and 50 for her. Participant A rejects it. Your earning is 600+0=600 Mohars. Her earning is 200+0=200 Mohars.

Game Begins

Now please grade the booklet.
- You will find a response sheet on the desk for you to fill up now.
- Please mark whether Participant A has solved at least 10 box correctly or not.
- Please write down how much, if at all, you want to share with Participant A.
- Please put the paper in the envelope and seal it. We will collect envelopes from you one by one and hand it in to the corresponding Participant A.
- If you have any questions regarding these instructions, please raise your hand. We will answer your questions in private.
- You will be paid once we get the response from Participant As.
- In the mean time please fill out the survey questionnaire.

The response of the Participant A has arrived. Check the envelope to see whether she has accepted your demand. Now calculate the total number of mohars you have earned. In another envelope placed below your desk you will find some money. You can take the money that you have earned from the envelope. Remember that the exchange rate is 1 Mohar = Rs. 0.50.

(Figure 1(b) was included)

4 Instructions (Participant A in UG)

4.1 Introduction

You are now taking part in an economic decision making study. During the experiment your payoffs will be stated in terms of a fictitious currency called “Mohar”. You will receive 200 Mohars for participating but you can earn additional money depending on the decisions you and the others make. You will be able to exchange all Mohars into Rupees at the end of the experiment. The exchange rate at which you can convert Mohars into Rupees is:

$$1 \text{ Mohar} = Rs \ 0.50$$

If you have any questions please let us know by raising your hand. Your question will be answered by us in private. Please note that communication between participants is strictly prohibited during the experiment.

We will give each one of you an identity number. Please do not lose your identity number. This entire experiment is anonymous. We will sometimes form pairs of two participants. The matching
of two participants has been randomly determined in advance. You will never be informed of the identity of the participant with whom you have been matched.

4.2 The Game

We have divided the total number of participants in this experiment session randomly into two equal groups: Group As and Group Bs. All the participants in this room are participants of Group A. All the participants of Group B are located in another room.

Each one of you is paired with a Group B participant who is in the other room. Nobody will be informed of the identity of the Participant B he/she is paired with. Each one of you will receive a booklet which you are required to fill out. The answer booklet contains twenty boxes. In each box at least two numbers add up to 10. Your task is to find the two numbers.

Answer Booklet

The answer booklet consists of simple math problem of the following type. It has boxes like this:

<table>
<thead>
<tr>
<th>1.79</th>
<th>3.70</th>
<th>2.99</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.34</td>
<td>7.19</td>
<td>5.55</td>
</tr>
<tr>
<td>9.01</td>
<td>4.45</td>
<td>6.32</td>
</tr>
</tbody>
</table>

You have to find a pair of numbers in the box which add up to 10. Having found the pair, you are required to encircle the corresponding numbers. In this example note that only 4.45 and 5.55 add up to 10. The booklet will contain 20 such boxes. You will have 15 minutes to solve the 20 boxes. Please write your id on top of the page.

Description of the Game

You are asked to complete the above task. Participant B will be paid 600 Mohars for correcting your answer booklet i.e 400 Mohars in addition to the fixed participation payment of 200 Mohars. If you solve at least ten boxes correctly then you “pass” the test, earn 200 Mohars and can then participate in next part of the game. If you are unable to solve at least ten box correctly then you “fail” the test and will not be able to take part in the next part of the game and the game ends.

Suppose now that you have solved at least ten boxes correctly and thus qualify to participate in the next part. In this part Participant B may propose to split 400 Mohars between himself and you. For example, if Participant B proposes to keep for himself x, then you are allocated 400 – x. But you may Accept or Reject the proposal as suggested by Participant B. If you accept his proposed division then both of you will earn as per the proposal. If you reject then you get 200 Mohars and she earns 600 Mohars. Note this will take place only if you “pass” the test.

You score 13 in the matrix task. You qualify for the next part. Participant B proposes a split of 200 for himself and 200 for you. You accept it. His earning is 600+200=800 Mohars. Your earning is 200+200=400 Mohars.

You score 13 in the matrix task. You qualify for the next part. Participant B proposes a split of 350 for himself and 50 for you. You reject it. His earning is 600+0=600 Mohars. Your earning is 200+0=200 Mohars.

Game begins

Now please start solving the box booklet. You have fifteen minutes and your time starts now.
Please fill out the exit survey now.
Now that the response sheet from Participant B is back. Please mention whether you Accept or Reject the split as proposed by him on the response sheet.
Please hand in your response sheets one by one and receive your payments now.
(Figure 1(b) was included)

5 Instructions-Norms

5.1 Introduction (Common for BG and UG treatments)

You are now taking part in an economic decision making study.

We will give each one of you an identity number. Please do not lose your identity number. This entire study is anonymous. Please do not discuss with your neighbors at any point during the study. Please raise your hands once you have read the questions.

General Instructions

Please write your participant ID in the space provided above.

On the following pages, you will read descriptions of a series of situations. These descriptions correspond to situations in which a person must make a decision. This description will include several possible choices available to, lets say, Individual A.

After you read the description of a situation, you will be asked to evaluate the different possible choices available to Individual A and to decide, for each of the possible actions, whether taking that action would be “socially appropriate” and “consistent with moral or proper social behavior” or “socially inappropriate” and “inconsistent with moral or proper social behavior.” By socially appropriate, we mean behavior that most people agree is the “correct” or “ethical” thing to do. Another way to think about what we mean is that if Individual A were to select a socially inappropriate choice, then someone else might be angry at Individual A for doing so. Social appropriateness rating is on a scale of -3 to +3 where -3 is “very socially inappropriate” and +3 is very socially appropriate.

In each of your responses, we would like you to answer as truthfully as possible, based on your opinions of what constitutes socially appropriate or socially inappropriate behavior.

To give you an idea of how the experiment will proceed, we will go through an example and show you how you will indicate your responses. On the next page you will see an example of a situation.

Example

Individual A is at a local coffee shop near campus. While there, Individual A notices that someone has left a wallet at one of the tables. Individual A must decide what to do. Individual A has four possible choices: take the wallet, ask others nearby if the wallet belongs to them, leave the wallet where it is, or give the wallet to the shop manager. Individual A can choose only one of these four options.

The table below presents a list of the possible choices available to Individual A. For each of the choices, please indicate your rating for the social appropriateness of the action on a scale of -3 to +3. Indicate your response in the table below.
If this were one of the situations for this study, you would consider each of the possible choices above and, for that choice, indicate the extent to which you believe taking that action would be “socially appropriate” and “consistent with moral or proper social behavior” or “socially inappropriate” and “inconsistent with moral or proper social behavior”. Recall that by socially appropriate we mean behavior that most people agree is the “correct” or “ethical” thing to do.

For example, suppose you thought that taking the wallet was very socially inappropriate, asking others nearby if the wallet belongs to them was somewhat socially appropriate, leaving the wallet where it is was somewhat socially inappropriate, and giving the wallet to the shop manager was very socially appropriate. Then you would indicate your responses as follows:

<table>
<thead>
<tr>
<th>Individual A’s choice</th>
<th>Your rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Take the wallet</td>
<td>-3</td>
</tr>
<tr>
<td>Ask others nearby if the wallet belongs to them</td>
<td>+1</td>
</tr>
<tr>
<td>Leave the wallet where it is</td>
<td>-1</td>
</tr>
<tr>
<td>Give the wallet to the shop manager</td>
<td>+3</td>
</tr>
</tbody>
</table>

Are there any questions about this example situation or about how to indicate your responses?

On the following pages, the situations deal with decisions that “Individual A” might have to make. For each situation, you will receive a sheet, with a table on which to indicate your responses.

For each situation, the experimenter will read a description of the situation. You will then indicate whether each possible choice available to Individual A is socially appropriate or socially inappropriate.

At the end of the session today, we will select one of the two situations by a coin toss (Head - Situation I and Tail - Situation II). We will then ask you to randomly choose one category from an envelope containing all the categories. Thus, we will select both a situation and category at random. For each situation and category, we will calculate the most frequently occurring response from all the responses in the room today. We will pay you Rs. 110 for your participation today. However if you give the same response as that most frequently given by other people in the room, then you will receive an additional Rs. 110 i.e. you will earn a total of Rs. 220. This amount will be paid to you, in cash, at the end of the experiment.

For instance, suppose that we randomly select the example situation above and the possible choice “Leave the wallet where it is”. Now if your response had been “somewhat socially inappropriate,” i.e rating -1 and if this was the response selected by most other people in today’s session, then you would receive Rs. 110, in addition to the Rs. 110 participation fee - your total earning would be Rs. 220. Otherwise you would receive only participation fee which is Rs. 110.

If you have any questions from this point on, please raise your hand and wait for the experimenter to come to you.

Please wait to turn the page until the experimenter asks you to do so. If you have any questions, please raise your hand and wait for the experimenter. All earnings in the hypothetical situation given below is stated in terms of a fictitious currency called Mohar. Each mohar in the situation can be exchanged into Rupees at the rate of 1 Mohar = Re. 0.50
5.2 The Situation (BG)

Please look at the figure given below. Citizens and public officials play a game where they are seated in two separate rooms but each citizen is randomly matched with exactly one public official. A citizen is given 20 problems to solve in 15 minutes. The public official grades the answer sheet of the citizen with whom he is matched. If the citizen solves at least 10 problems correctly, she “passes” the test but if she scores less than 10 she “fails” the test. The citizen is entitled to a prize of 400 in addition to a base amount of 200 if she solves at least 10 problems and “passes” the test but she earns only 200 if she fails. However even if the citizen solves 10 problems or more correctly the supervisor demands a bribe in order to let the citizen pass and earn the prize. In other words whether to let the citizen pass is entirely his discretion. He may demand a bribe amount of \{0,20,40,...400\}. He can also choose not to take a bribe i.e. demand 0. The citizen upon receiving a demand for bribe may accept or reject the bribe offer.

If she accepts the bribe offer then she gets 200 (base participation fee) + 400 (prize money) - bribe amount. If she rejects the bribe demand then she does not gets the prize but only gets 200 (base participation fee). If the citizen accepts the bribe demand then the public official gets 600 + bribe amount demanded. If she rejects to pay the bribe then the public official gets only 600. Let us go through the figure again to calculate the earnings in each scenario.

Citizen scores 13 in the matrix task. Citizen is entitled to a payment of 600 Mohars. Public official demands for a bribe of 200 for himself. Citizen accepts it. Public Official’s earning is 600+200=800 Mohars. Citizen’s earning is 600-200=400 Mohars.

Citizen scores 13 in the matrix task. Citizen is entitled to a payment of 600 Mohars. Public official demands for a bribe of 350 for himself. Citizen rejects it. Public Official’s earning is 600+0=600 Mohars. Citizen’s earning is 200+0=200 Mohars.

Rate the action of the public official and the citizen on a scale of -3 to +3 as stated above. However the citizen may accept to pay a bribe or she may reject it.

Remember you are not being asked to report your personal appropriateness rating but social appropriateness rating and you will be paid if your rating matches with the rating of most other participants.

Rate the action of Public official on a scale of -3 and +3 in the response sheet given to you. Rate the action of Citizen on a scale of -3 to +3 in the response sheet given to you.

Remember you will be rewarded if your rating matches with the rating of most other people in the room today.

Figure 1(a) shown below gives a visual description of the payoffs.

(Figure 1(a) was included)
Response Sheet

Identity Number. _________________________________

Situation

Rate Public Official’s decision.

<table>
<thead>
<tr>
<th>Amount Sought by the public official as bribe</th>
<th>Amount of the prize left with the citizen</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 i.e. Public official does not ask for a bribe</td>
<td>400</td>
<td></td>
</tr>
<tr>
<td>10-50</td>
<td>390-350</td>
<td></td>
</tr>
<tr>
<td>60-200</td>
<td>340-300</td>
<td></td>
</tr>
<tr>
<td>110-150</td>
<td>290-250</td>
<td></td>
</tr>
<tr>
<td>160-200</td>
<td>240-200</td>
<td></td>
</tr>
<tr>
<td>210-250</td>
<td>190-150</td>
<td></td>
</tr>
<tr>
<td>260-300</td>
<td>140-200</td>
<td></td>
</tr>
<tr>
<td>310-350</td>
<td>90-50</td>
<td></td>
</tr>
<tr>
<td>360-390</td>
<td>40-10</td>
<td></td>
</tr>
<tr>
<td>400 i.e. Public official demands the entire amount as bribe</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Total amount that can be extracted as bribe is 400.

Now rate the decision of the Citizen.
<table>
<thead>
<tr>
<th>Amount Sought by the public official as bribe</th>
<th>Citizen’s response</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 i.e. Public official does not demand a bribe</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>10-50</td>
<td>Accept</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reject</td>
<td></td>
</tr>
<tr>
<td>60-200</td>
<td>Accept</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reject</td>
<td></td>
</tr>
<tr>
<td>110-150</td>
<td>Accept</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reject</td>
<td></td>
</tr>
<tr>
<td>160-200</td>
<td>Accept</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reject</td>
<td></td>
</tr>
<tr>
<td>210-250</td>
<td>Accept</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reject</td>
<td></td>
</tr>
<tr>
<td>260-300</td>
<td>Accept</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reject</td>
<td></td>
</tr>
<tr>
<td>310-350</td>
<td>Accept</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reject</td>
<td></td>
</tr>
<tr>
<td>360-390</td>
<td>Accept</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reject</td>
<td></td>
</tr>
<tr>
<td>400 i.e. Public official demands the entire amount as bribe</td>
<td>Accept</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reject</td>
<td></td>
</tr>
</tbody>
</table>
5.3 The Situation (UG)

Please look at the figure given below. Participants A and Participants B play a game where they are seated in two separate rooms but each Participant A is randomly matched to exactly one Participant B. Participant A is given 20 problems to solve in 15 minutes. Participant B is supposed to grade the answer sheet of the Participant A he was matched with. If Participant A solves at least 10 problems correctly then she qualifies to take part in the next part of the game. Otherwise she earns only 200 and leaves. If she qualifies for the next part of the game then she is eligible for a transfer from Participant B. Each participant B has 400 between himself and Participant A. Participant B then splits 400 between himself and Participant A. He can share any amount including 0 and 400 i.e. he can share nothing with Participant A or he could give away the entire amount to her. P-A in turn could accept or reject the proposed division by Participant B. If she rejects the offer then she gets only 200 and P-B gets 600. If she accepts the offer then P-A gets 200+400 - amount that P-B keeps with himself and P-B gets 600+amount he keeps with himself. Let us go through the figure to further clarify.

Participant A scores 13 in the matrix task. Participant A qualifies for the next part. Participant B proposes a split of 200 for himself and 200 for her. Participant A accepts it. Participant B’s earning is 600+200=800 Mohars. Participant A’s earning is 200+200=400 Mohars.

Participant A scores 13 in the matrix task. Participant A qualifies for the next part. Participant B proposes a split of 350 for himself and 50 for her. Participant A rejects it. His earning is 600+0=600 Mohars. Participant B’s earning is 200+0=200 Mohars.

Rate the action of Participant B on a scale of -3 and +3 in the response sheet given to you.
Rate the action of Participant A on a scale of -3 to +3 in the response sheet given to you. Remember you will be rewarded if your rating matches with the rating of most other participants. Figure 1(b) shown below gives a visual description of the payoffs. (Figure 1(b) was included)
**Response Sheet**

Identity Number. _______________________________________

**Situation**

Rate Participant B’s Decision.

<table>
<thead>
<tr>
<th>Total amount to be divided is 400</th>
<th>The part Participant B proposes to keep with himself</th>
<th>Amount offered to Participant A</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 i.e. Participant B does not keep anything with himself</td>
<td></td>
<td>400</td>
<td></td>
</tr>
<tr>
<td>10-50</td>
<td></td>
<td>390-350</td>
<td></td>
</tr>
<tr>
<td>60-200</td>
<td></td>
<td>340-300</td>
<td></td>
</tr>
<tr>
<td>110-150</td>
<td></td>
<td>290-250</td>
<td></td>
</tr>
<tr>
<td>160-110</td>
<td></td>
<td>240-110</td>
<td></td>
</tr>
<tr>
<td>210-250</td>
<td></td>
<td>190-150</td>
<td></td>
</tr>
<tr>
<td>260-300</td>
<td></td>
<td>140-200</td>
<td></td>
</tr>
<tr>
<td>310-350</td>
<td></td>
<td>90-50</td>
<td></td>
</tr>
<tr>
<td>360-390</td>
<td></td>
<td>40-10</td>
<td></td>
</tr>
<tr>
<td>400 i.e. Participant B keeps everything for himself</td>
<td></td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Now rate the action of Participant A.
Total amount to be divided is 400

<table>
<thead>
<tr>
<th>The part participant B proposes to keep with himself</th>
<th>Participant A’s response</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 i.e. Participant B does not keep anything with himself</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10-50</td>
<td>Accept</td>
<td>-</td>
</tr>
<tr>
<td>60-200</td>
<td>Reject</td>
<td>-</td>
</tr>
<tr>
<td>110-150</td>
<td>Accept</td>
<td>-</td>
</tr>
<tr>
<td>160-110</td>
<td>Reject</td>
<td>-</td>
</tr>
<tr>
<td>210-250</td>
<td>Accept</td>
<td>-</td>
</tr>
<tr>
<td>260-300</td>
<td>Reject</td>
<td>-</td>
</tr>
<tr>
<td>310-350</td>
<td>Accept</td>
<td>-</td>
</tr>
<tr>
<td>360-390</td>
<td>Reject</td>
<td>-</td>
</tr>
<tr>
<td>400 i.e. Participant B keeps everything for himself</td>
<td>Accept</td>
<td>-</td>
</tr>
</tbody>
</table>
Chapter Three
Corruption, Norm Violation and Decay in Social Capital

Ritwik Banerjee*
Aarhus University, Denmark

Abstract

Corruption has enormous economic and social consequences. While the direct economic consequences of corruption have been well studied, its social consequences remain relatively less understood. This paper aims to causally link corruption to lower trust in society using data from an experiment. In one treatment, subjects play a harassment bribery game while in the control, they play a strategically identical but differently framed ultimatum game (UG). Each treatment is followed by a trust game. We also elicit the social norm governing the two strategically identical but differently framed situations. Finally, we use this novel setting to study the association between behavior in a trust game and response to the trust question in the World Value Survey. Our results show that a) subjects in the harassment bribery game treatment trust less than those in the strategically identical but differently framed ultimatum game treatment; b) the negative spillover effects of corruption on trust increases with an increase in the bribe and with a decrease in the social appropriateness of the demand; c) lower trust in the bribery game treatment is explained by lower expected return on trust – this is also predicted by a simple theoretical framework where Bayesian agents update their beliefs about trustworthiness depending on whether they meet an honest or a dishonest person; d) surprisingly, for both the treatments, social norm violation engenders the decay in trust; e) the WVS-trust question captures expectations about others’ trustworthiness, however the response to the WVS-trust is stable while behavior in the trust game is susceptible to short term fluctuations.

Keywords: Corruption, Social Trust, Trust Games

JEL Classification: C91 C92 D03

*Email: rbanerjee@econ.au.dk. Address: Department of Economics and Business, Fuglesangs Allé 4 8210 Århus V, Denmark. I thank Alexander Koch, Nabanita Datta Gupta and seminar participants at Aarhus University for their comments and suggestions. I also thank Sujoy Chakrabarty and Bibek Banerjee for making it possible for me to run the experiment at Institute of Management and Technology and Manas Pathak for his research assistance. The experiment was funded with grants from INCLUDE and Department of Economics and Business, Aarhus University.
1 Introduction

Social capital, which comprises of commonly held values such as trust, trustworthiness and cooperative norms, is increasingly seen today as an important component of a successful economic environment. Given that social capital helps circumvent the necessity for expensive complete contracts and thereby decreases the costs of enforcing contracts (North, 1990; Williamson, 1985), it is not surprising that it has been found to have a positive instrumental role in a wide range of economic activities: from economic growth (Knack and Keefer, 1997) to financial development (Guiso et al., 2004, 2008) and trade and investment (Guiso et al., 2009).

Studies show that this vital ingredient of economic activity is negatively associated with corruption in a cross country panel framework. Figure 1 documents this association in a dynamic panel of countries with trust data from World Value Survey (WVS) and corruption (perception) data from International Country Risk Guide (ICRG), over four WVS waves. It illustrates the stylized fact that not only is corruption and trust negatively related cross-sectionally, but the movement of most countries have followed a trajectory from high trust - low corruption to low trust - high corruption during the period, as is indicated by the arrows which point towards the South East for most countries.

This association has been studied primarily by political scientists and to some extent by economists, however the precise causal link and the mechanisms driving the association remains less known. Some have taken the view that low levels of trust in a society may engender and nurture corruption as people fail to develop cooperative ethos. A lack of trust may also diminish the sense of doing something wrong or “immoral”, leading to a perception of high corruption in the society (Xin and Ruden, 2004), which in turn may lead to greater prevalence of corruption (Bardhan, 1997; Innes and Mitra, 2013). Some authors have taken the cross country route (La Porta et al., 1997; Bjornskov, 2011) while others have used more qualitative data (Seligson, 2002; Davis et al., 1997; Moreno, 2002) to take this view on the causal link.

Corruption has also been viewed as a cause for the erosion of social capital (Anderson and Tverdova, 2003; Chang and Chu, 2006; Della Porta, 2000), suggesting that corruption undermines trust in institutions. This view has been supported with evidence from the impact of political scandals on trust (Bowler and Karp, 2004), and by relating confidence in institutions entrusted to control corruption to interpersonal trust (Rothstein and Stolle, 2002)\(^1\). Others still, have interpreted the relation as one of mutual causality (see for instance Uslaner (2002), Morris and Klesner (2010)) where higher corruption reduces lower trust which in turn triggers higher corruption.

Despite the fact that the association between trust and corruption has been widely documented in a number of settings, the causal interpretation between the two rests on a weak foreground. The potential simultaneity in the association between corruption and trust leads to a strong endogeneity

\(^1\)The literature exploring the relation between trust and corruption often focuses on either political/institutional trust or interpersonal trust. While conceptually different in terms of the objects of trust, both originate from the same behavioral primitive.
bias. Besides, inclusion of country level institutional variables in the empirical specifications, which are collinear with the trust levels in society, leads to imprecise estimation and can potentially increase bias in the presence of omitted variables. Though some studies have tried to correct for the endogeneity bias in the estimation of the relation between corruption and trust by using an instrumental variable approach, the exogeneity of the instruments chosen are not above board\(^2\).

The matter is further complicated by the lack of data which are time varying and comparable across countries.

Figure 1: Within and between country gradient between trust and corruption

Note: Arrows indicate the movement of countries in the Trust-Corruption space from Wave 2 to Wave 5 in the World Value Survey. The arrows point to the South East for more countries suggesting that over the period, most countries have moved towards higher corruption and that this movement has also been associated with lower trust. The dotted line shows the cross sectional gradient. Data source: Corruption - International Country Risk Guide (ICRG); Trust - World Value Survey. Arrows in some data points are missing as these countries have only one data point each since WVS has brought more and more countries within its fold over time.

In this paper, we contend that the question about the causal relationship between corruption and trust may also be addressed from a behavioral standpoint because both corruption (or more

\(^2\)For example, political constraints, monarchy etc., as choice of instrument in Bjornskov (2011), and percentage of Christian population, in Bjornskov (2003), can be argued to be related to both corruption and trust.
generally unethical behavior) and trust originate from behavioral primitives and both have been extensively studied through lab based experimental methods (for a review of experimental corruption games see Abbink (2006) and Serra and Wantchekon (2012)). Thus, we take an experimental approach to analyze if people exhibit less trust in a standard trust game after having played a corruption game than after having played a strategically identical but differently framed ultimatum game. The difference in behavior between the two frames have been analyzed in Banerjee (2014). In this paper, we show how these two frames trigger different response on trust and identify the channel through which the causality between corruption and trust plays out.

In our between subject design, “citizens” and “public officials” play a real effort harassment bribery game. A citizen performs a task and earns a prize if successful. A public official, however, may demand a bribe in order to let the citizen have her prize which the citizen may accept or reject. In the strategically identical but differently framed treatment, Participant A (analogous to the Citizen) - upon successfully completing the task - earns the right to go to the second stage of the game. At the second stage, Participant B (analogous to the Public Official) plays an ultimatum game, with the same stake size as the prize of the bribery frame, and decides how much to share with Participant A, which the latter can accept or reject. In our experiment we randomly assign people to either the bribery frame or the strategically identical but differently framed ultimatum game frame. Having primed the subjects through the two frames, we measure the trust behavior of the subjects in a trust game. In this way, not only do we cleanly identify the causal impact of corruption on trust, if any, but also answer whether lower trust in people is associated with greater unethical behavior. Furthermore, the two environments, despite being strategically identical, is possibly governed by different social norms. To examine this and to understand the effect of social norms governing corruption on trust behavior, we elicit social appropriateness measure using a coordination tool developed by Krupka and Weber (2012). The design of our experiment thus allows us to precisely identify the mechanisms - preferences vs expectations - through which corruption affects trust.

First, our results confirm that the two frames trigger different behavioral responses - in particular, the bribery frame successfully imposes the intended frame of immorality. As hypothesized, it is indeed the case that the two frames are governed by different social norms and this partly explains the difference in actual behavior. Second, the Citizens trust less than the Participant As and the baseline subjects in the trust game but we find no difference in trustworthiness between them. Third, trust levels are inversely related to the bribe amount. Fourth, we find that the expectations of the Citizens about the trustworthiness of the matched partner, is lower when compared to that of the Participant As and baseline subjects. The negative shock to expectation is driven by the violation

3Harassment bribery is a form of bribery where a public official asks for a bribe from a citizen who is entitled to a service that the official is obligated to provide. Petty bribery of this nature is very common in developing countries where citizens, despite being entitled to government services (e.g. passport, driver’s license), have to pay a bribe in order to obtain them or avoid inordinate procedural delays. Harassment bribery has been studied through experimental games in the past using by and Abbink et al. (2012).
of social appropriateness norms and explains part of the difference in trust behavior. Interestingly, there is no independent effect of the corruption frame on trust (i.e. there is no priming or mindset effect) but there is a mindset effect of corruption on expectations about trustworthiness. Hence, our findings suggest that norm violation and corruption mindset effect affect expectations about trustworthiness, which in turn leads to lower trust. Finally, we find a weak negative association between corrupt and trusting behavior among Public Officials.

We further use this design to resolve another important issue - namely, the interpretation of the answer to the question in WVS, which is taken as measuring trust in the literature: “Generally speaking, would you say that most people can be trusted or you need to be very careful in dealing with people?” To resolve the conflicting finding that in some studies the answer to this question (WVS-trust henceforth) has been found to be correlated with behavior in experimental trust games (Fehr et al., 2003; Bellemare and Kroger, 2007) while not in others (Glaeser et al., 2000; Lazzarini et al., 2004), Sapienza et al. (2013) studied the correlation between the two by factoring in the beliefs of the subjects about others’ behavior. They found that while the sender’s behavior in a trust game is driven both by beliefs and preferences, WVS-trust captures the belief based component. While we find this to be true in a cultural context characterized by low trust (India) and thus different from the one in which Sapienza et al. (2013) conduct their study, there is more to it. In our experiment, the correlation between WVS-trust and expectation about trustworthiness holds for Participant As and baseline subjects, but not for the Citizens. Taking into consideration the fact that the expectation of the Citizens about the trustworthiness of the matched partners is lower than that of the Participant As and baseline subjects, our results suggest that belief based expectations are vulnerable to locally determined short term fluctuations, while the answer to the WVS-trust is more stable. Our interpretation of this result is that while the reference frame of beliefs, elicited through trust games, is more immediate, subjects contextualize the WVS-trust question in a longer term pan-societal context and therefore the two measures may not always be used interchangeably.

Our contributions to the literature are the following. We provide a clean identification of the causal link that corruption leads to lower trust using an experimental approach. In doing so, we pin down the precise behavioral mechanism through which corruption affects trust and show that the decline in the social capital in a corrupt environment has much to do with the violation of social appropriateness norms. With this, we also identify one potential channel through which behavioral spillovers work in wider social contexts in general and in experimental games in particular - namely that of norm violation.

The behavioral spillover effect, also sometimes called a mindset effect, may be defined as an effect which is observed only when a game is played together with other games but not when the game is played in isolation. Such effects, which are common in the experimental literature, have been found to enhance cooperation (Cason et al., 2012; Brandts and Cooper, 2006; Albert et al., 2007; Cason and Gangadharan, 2013), induce rationality (Cherry et al., 2003; Cherry and
Shogren, 2007) and even change actions when subjects hear about the actions of others in their group (Huck et al., 2011). Duffy and Ochs (2009) even find that a cooperative norm emerges in a fixed matching prisoner's dilemma game but not in a random matching version and such behavioral spillovers influence even new subjects with a Pareto improving coordination equilibrium sustaining in a larger group (Weber, 2006). Though some studies have inferred that higher cognitive load induced by a greater outcome entropy, uncertainty and lesser path dependence induces higher positive behavior spillovers\(^4\) (Bednar et al., 2012; Cason et al., 2012), we know surprisingly little about the mechanisms behind negative behavioral spillovers. More importantly, the explanations offered in these studies shed little light in terms of how and why behavioral spillovers play out in real life. To the best of our knowledge, ours is the first study to show that social norm violation, through an effect on beliefs about others behavior, plays an important role in generating negative spillovers.

Besides, we also contribute to the framing literature and show that the framing effect may, in part, be explained by the social appropriateness norms of the two frames. Within this set-up, we re-examine what the interpretation of the answer to the WVS-trust question is in a cultural context which is different from the ones in past studies. This exercise of validating the behavioral correlates of the answer to WVS-trust question in a different cultural context itself is important since WVS aims to investigate “...Political and Sociocultural Change” around the world. Hence, it is important that questions used in WVS in general and the WVS-trust question in particular, is interpreted by respondents in the same way across the world, as it should, in order to allow for comparability. Our study relates WVS-trust to expectations of trustworthiness but with the caveat that this interpretation may break down in certain situations.

The rest of the paper is organized as follows: Section 2 lays out the experimental design and briefly sketches the important survey questions. Section 3 presents the broad results and the mechanisms behind the results and Section 4 discusses the results and offers the concluding remarks.

2 Experimental Design

A simple way to examine the effect of corrupt transaction on trust is to compare the trust behavior of victims in a corruption game with that of subjects who do not participate in the corruption game (i.e. baseline treatment). However, such an exercise may end up overestimating or underestimating the true effects of corruption on trust due to intertwined behavioral confounds. To preclude such confounds, we develop a counterfactual of the corruption game treatment, which is strategically identical to the corruption game but is differently framed. Any treatment effect which is found in the trust behavior following the corruption frame and its strategically identical but neutrally

\(^4\)Though differently names, the explanations offered in these studies are essentially similar. Greater variability in outcome or “outcome entropy” leads to lesser path dependence and thus higher uncertainty.
framed version can then be attributed solely to the effect of the corruption frame. However, for completeness we also conduct the baseline trust game and compare the results of the other two treatments with the baseline trust results in Section 3.

With this cue, we design a two part experiment where in the first part (similar to Banerjee (2014)) subjects are primed either through a corruption game, more specifically a harassment bribery game (BG henceforth), or a strategically identical but neutrally framed ultimatum game (UG henceforth). The two frames differ in terms of the language used - loaded (e.g. Citizen, Public Official, bribe etc.) vs. neutral (e.g. Participant A, Participant B, transfer etc.), and also in terms of the reference frame or the sense of expectation among the subjects. After having randomly treated subjects either through the BG or the UG, we let them play a modified version of the trust game (Berg et al., 1995) in the second part and then observe the trust behavior.

The design of the BG and UG treatment relies on the prior that any amount demanded off subjects in a bribery game may be considered unfair in the bribery game, but the same amount when retained in an ultimatum game may be considered fair. In other words, an act of demanding bribe may be considered immoral while that of retaining an amount in an ultimatum game may be considered moral. Thus, the two frames are expected to trigger different emotional responses despite being strategically identical. We exploit this and analyze their impact on trust behavior in the second part. In addition, we also compare the impact of trust in the two frames with that of baseline level of trust.

This strategy of studying behavior after triggering a certain behavioral response has been adopted in past studies; for instance, Burnham et al. (2000) study the effect of framing the matched partner as a friend or foe, Drouvelis et al. (2010) prime cooperation in social dilemma games, Ariely et al. (2003) prime individuals to “arbitrary anchors“ e.g. social security number and finds a difference in the willingness to pay behavior and more recently, Buser and Dreber (2013) examine the effect of competition on cooperation.

2.1 First Part - Bribery or Ultimatum Game

2.1.1 Bribery Game Treatment

Figure 2(a) lays out the bribery game. A citizen (C) has ten minutes to complete a real effort task, following which she gets a prize winning code. If she successfully completes the task and gets the prize winning code then she is entitled to a prize of $400 Mohars (M400). Otherwise she earns only the participation fee of M200.

---

5 We introduce a computer based task where citizens have to count the number of occurrences of the letter “A” in five different panels, each with a random sequence of letters A, B, C, D, E. The panels are so designed that they appear one after the other and a subject is not be able to proceed to the next one unless she correctly counts the number of As in the current one. They have ten minutes within which they are required to find the prize winning code, which is mentioned at the end of the fifth panel. The task is designed in such a way that a Public Official does not need to grade the answers of the Citizen but is still able to figure out whether she has been able to complete the task or not by looking at the prize winning code in her response sheet.
However, even if the citizen does get the prize winning code, a public official may demand a bribe of amount \( b \) in order to let the citizen have her entitled prize. Notice that the bribe of amount \( b \) is an extract from the value of the citizen’s entitlement, which in this case is \( M \) 400. Any \( b > 400 \) is likely to be rejected by a reasonable citizen. So, we let \( b \in \{0, 100, 200, 300, 400\} \). Upon receiving a bribe demand for \( b \), a citizen may accept it or reject it. If she accepts the demand for a bribe, she earns \( M(200 + 400 - b) \) whereas if she rejects it, she earns the participation fee of \( M \) 200 only.

On the other hand a Public Official gets a participation fee of \( M \) 200 and a salary of \( M \) 400 for the task of approving the prize for the citizens. Thus, if he demands a bribe \( b \) and his demand is accepted, then he earns \( M(200 + 400 + b) \). If his demand is rejected however, he earns \( M(200 + 400) \) i.e. \( M \) 600 only. Since the citizen always earns less than the public official, an act of bribery cannot be explained by alternative explanations such as inequity aversion. This is a one shot game which reflects the standard harassment bribe situation where a briber and a bribee meet only once.

The real effort task for the citizen is calibrated in a way such that she is more likely to successfully complete the task and get the prize winning code. At the same time it is crucial for her to perform the task in order to induce in her a sense of entitlement and hence a sense of harassment if a bribe is demanded.

### 2.1.2 Ultimatum Game Treatment

Figure 2(b) lays out the ultimatum game treatment. In this treatment Participant A (analogous to citizens in BG) has ten minutes to complete the real effort task, following which she gets a prize winning code. If Participant A (P-A henceforth) successfully completes the task then she qualifies for the next round of the game in which Participant B (P-B henceforth) decides to keep \( Mx \) for himself and makes an offer \( M(400 - x) \) to P-A. Thus in this case P-A’s completion of the task only determines whether she proceeds to the next part of the game or not. If she does proceed to the next stage, then she can decide whether to accept or reject the offer in which cases she receives \( M(200 + 400 - x) \) and \( M \) 200, respectively. If P-B’s offer is accepted then P-B gets \( M(200 + 400 + x) \), if it is rejected, then he gets only \( M(200 + 400) \), i.e. \( M \) 600.

This alternative frame, though strategically identical to BG, is designed to trigger a different behavioral response in a P-A than that of a citizen in BG. This is achieved not only by changing the sense of entitlement in the two frames, but also by the naming of the subjects as “Player 1” and “Player 2” instead of “public official” and “citizen” and by the use of words such as “transfer” instead of “bribe”. Thus the expectations of the subjects and their psychological reference points are different in BG and UG. Also note that we implement a strategy method to elicit C/P-A’s accept/reject decision.

\[ \text{In both the treatments of our experiment all the subjects were able to complete the task within the stipulated time and get the prize winning code.} \]
2.2 Second Part - Trust Game

Following the BG/UG treatment, subjects play a trust game (also called an investment game). Our design of the trust game closely follows the game proposed by Berg et al. (1995) where a sender is matched to a receiver. The receiver is matched in a manner such that he is not the same matched person in BG/UG. In our game the sender is endowed with $M=400$. He can then decide how much to send, $t \in \{0, 50, 100, \ldots, 400\}$, to a receiver. Any amount sent is then multiplied by 3. The receiver then decides how much to return, $w \in \{0, 50, 100, \ldots, 3t\}$, to the sender. As a result, the payoff of a sender equals $M(400 - t + w)$ whereas that of the receiver equals $M(3t - w)$. The predominant interpretation in the literature is that $t$ is a measure of trust, while $w$ is a measure of trustworthiness.

Subjects make their decisions both as a sender and a receiver in a strategy elicitation method. Behavioral spillovers have been found to be stronger when games are played with the same subjects than when they are played with different subjects (Cason et al., 2012). Our stranger matching protocol ensures that we do not pick up a trivial effect where a sender, after being a victim of a bribe demand, sends a lower amount (i.e. shows less trust) to the corrupt person she is matched with.

In an exhaustive metastudy of the existing literature comparing strategy elicitation method versus direct method, Brandts and Charness (2011) finds that in no case where treatment effect was found in strategy elicitation method, it was not found in direct method. Given this finding and the fact that strategy elicitation method helps us obtain a rich data set, which otherwise may not have been available, we decided to employ this method. This allowed us to perform a robust sensitivity analysis. We believe that the treatment effect which we find will have
and under role uncertainty. In the first decision, subjects play as a sender and they decide how much of their initial endowment of M400 they are willing to send to the receiver. They choose any amount between M0 and M400 in multiples of fifty. In the second decision, they indicate how much they expect to be returned by their matched receiver for each possible amount sent. Thus, we obtain their response both on and off the equilibrium path, enabling us to perform a richer sensitivity analysis. To motivate the subjects to report their beliefs accurately, we incentivize this question by paying an additional M100 if the response matches the actual return of the receiver for a randomly chosen amount sent. The subjects then take a third decision, where they play the role of a receiver. In this decision they indicate the amount they are willing to return to the sender for each possible amount which they may have received. The earnings from this part are determined by randomly picking either the role of a sender or that of a receiver.

2.3 Survey Data

At the end, the subjects write an exit survey which is designed to acquire demographic data and measure various personality traits. Our main survey measure of trust came from the standard trust question of the World Value Survey: “Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?” This question was part of the survey which Sapienza et al. (2013) had administered a week before the actual experiment was conducted. In our case, it was important that this question was asked after the experiment was conducted in order to allow for the possibility that experimental behavior could influence its response. Table 11 in the Appendix gives the summary statistics of the responses to this question. In our sample, 32% of the students said that most people can be trusted. The corresponding figure in the nationally representative data from all the waves of the World Value Survey in India is 35%.

The survey included two additional measures of trust. The first one, also taken from WVS, was: “Do you think most people would try to take advantage of you if they got the chance, or would they try to be fair?” The second question was, “Suppose you left your wallet with Rs. 500 in the Metro. On a scale of 1 to 10, how much do you think are the chances that you will get it back?” The response to these questions have not been included in the analysis below but the pairwise correlations and summary statistics have been reported in Table 11 in the Appendix.

Besides the trust questions, we also obtain risk measures by asking survey questions which have been used to measure risk in large scale behavioral surveys such as German Socio-Economic Panel (SOEP). For example, one question directly asks subjects to make an assessment about their willingness to take risk in general on a scale of 1 to 10: “How willing are you to take risks, in general?” Though the survey also includes other questions about risk attitudes in specific dimensions such as financial matters, car driving, health and career, our focus in the analysis below relies on the general measure of risk since Dohmen et al. (2011) find that the general measure is been accentuated had we used a direct method.
the best predictor of risky behavior elicited through an experiment. The descriptive statistics of
the variables used in the analysis are presented in Table 7 in the Appendix.

2.4 Experiment Procedure

The experiment was conducted in Delhi in January 2014, with student subjects recruited from
Institute of Management Technology. A total of 218 students were recruited and each subject
participated in only one role of one treatment. Sessions were randomly assigned UG and BG
treatments and subjects were randomly assigned roles. The subjects were divided into two equal
groups - each group was allocated a different room. The instructions for the respective roles were
read out and several examples were worked out and the earnings of each role were mentioned.

The Citizens/Participant A completed the computer based task which then led them to the prize
winning code. They mentioned the code in the response sheet which was then handed out to the
matched Public Official/Participant B. The PO/P-B then made their bribe demand decision which
was subsequently returned to C/P-A for their acceptance or rejection decision. The earning from
this part was computed and mentioned on the response sheets, which was returned to the C/P-A
subsequently. It was crucial that they knew how much less they were being paid than what they
were entitled to as it could potentially reinforce the sense of unfairness.

Following The BG/UG treatment, the subjects played the trust game. They made their decisions
both as sender and receiver but their earnings from this part were determined by randomly picking
one of the roles. The final earnings of the subjects were determined by randomly picking one of
the two parts. If the first part was chosen, then the subjects were paid according to the roles they
played. If the second part was chosen then one group was randomly given the sender’s earnings
and the other group the receiver’s earnings.

The sessions lasted for one and a half hour. The instructions were read out in English. All
payoffs for this experiment were stated in terms of a fictitious currency called Mohar\(^9\). At the end
of the experiment participants exchanged their earnings for Rupees at the rate of \(100\text{ Mohars}=\text{Rs.50}\). The earnings ranged from Rs. 100 to Rs. 700 with an average of Rs. 252 (~ 10 USD in PPP terms).

2.5 Social Norm

A different experiment is conducted to gather data on the social norm governing the two situations
BG and UG using a tool developed by Krupka and Weber (2012). The subjects report their
subjective assessment of how appropriate or inappropriate most people find the actions in a
particular situation to be. They are rewarded if their appropriateness ratings match those of
most other people (mode) in the room. The reward, attached to the mode of the distribution of
ratings, helps elicit a second order belief about the societal (defined as the others in the classroom)

\(^{9}\)A unit of exchange used in medieval India, an etymological history of which can be traced back to the persian
word mohr, meaning seal
viewpoint. Following Krupka and Weber (2012), we interpret this as the social norm governing the situation.

Subjects see either the BG or the UG treatment. Then they rate each action for each role as one of the following - “very socially inappropriate”, “socially inappropriate”, “somewhat socially inappropriate”, “value neutral”, “somewhat socially appropriate”, “socially appropriate” and “very socially appropriate”. The responses were later converted into numerical scores of -3, -2, -1, 0, +1, +2 and +3 respectively.

Each subject is paid a participation fee of Rs. 110. In the BG situation subjects report their social appropriateness rating for the actions of the Public Official and that of the Citizen. One bribe amount is randomly chosen. The modal response of the appropriateness rating for the randomly selected amount is noted. If a subject’s rating for the randomly selected category for either the public official or the citizen is the same as the modal response, then she is paid double the participation fee, i.e. Rs. 220. Those who are given the UG situation follow exactly the same protocol as above except that subjects give their social appropriateness ratings for the actions of P-A and P-B.

The second experiment was conducted in December 2012 in connection with a different study, though the situations which the subjects evaluated were identical to the BG and UG frame described above. A session described either the BG environment or the UG. Each session comprised of around 20 subjects and four such sessions were conducted. A total of 40 subjects participated in the norm elicitation of BG and another 40 in that of UG. The design we implemented is an entirely between subject one in order to avoid possibilities of experimenter demand effect and confounds arising from subjective perceptions of the situations. Thus, none of the subjects who participated in the UG and BG treatments, took part in the social norm treatment.

3 Results

3.1 Trust Behavior

Since we were interested in the phenomenon of harassment bribery, we calibrated the task in a manner such that the majority of the subjects crossed the threshold, and as it turned out all the subjects completed the task within the stipulated time. In all we had 45 and 46 pairs of subjects in UG and BG, respectively. In the following discussion (but not in the actual experiment) we use the word *ultim* to denote the UG treatment equivalent of the bribe i.e. the amount that P-B proposes to keep for himself. Also, since our focus in this paper is on the behavioral impact of being at the receiving end of corruption on trust, we will restrict our attention to the trust behavior of Citizens in BG, Participant As in UG and the subjects of the baseline trust game, unless otherwise stated.

We find significant differences in the way people perceive the BG and UG frames. It is reflected in the distributions of bribe/*ultim* amount in Figure 3a and also in the mean difference tests in
Table 1. A Kolmogorov-Smirnov test rejects the null hypothesis that the distributions of bribe and *ultim* are equal ($p$-value = 0.03). Not only do we find that 19% of the subjects do not demand a bribe in the BG treatment as opposed to none in UG ($\chi^2$ test, $p$-value=0.00), the mean bribe demanded is $M_{182.6}$ in BG, but the mean *ultim* retained is $M_{268.9}$ in UG ($t$-test, $p$-value<0.001). For the restricted sample, i.e. given that a bribe/*ultim* was demanded, we reject the hypothesis that bribe and *ultim* amount are equal ($t$-test, $p$-value=0.02). The difference in behavior of the PO in BG with that of the P-B in UG clearly indicates that the BG frame successfully induces an immoral environment. This is further reflected from the acceptance/rejection decision of the C/P-A which we see in Figure 3a. The percentage of Citizens who reject a bribe is always greater than that of Participant As for all possible bribe amount indicating that a bribe demand is considered more unfair than a demand for an equivalent amount of *ultim*. Table 1 shows that the difference is statistically significant for the amounts 200 and 300 ($\chi^2$ test, $p$-value<0.001 for both the amounts). The results clearly show that the two frames, though strategically identical, have triggered different behavioral responses not only among PO/P-B but also among C/P-A.

That the subjects perceive the two frames differently is also indicated by the left panel of Figure 3b which plots the difference in social norm governing the two frames. A zero bribe demand is considered very socially appropriate but a zero *ultim* is considered socially inappropriate. Furthermore, the social appropriateness of the bribe demand decreases with an increase in the bribe amount, but that of *ultim* is maximum at equal split. The full distribution of the ratings, including the mean and mode, for each bribe/*ultim* category, is reported in Table 8. We shall later examine if this data can be used to predict trust behavior in the trust game.

Figure 4 lays out the distribution of the levels of trust between BG and UG and compares the mean trust levels (amount shared as a sender in the trust game) of the Citizens with that of the Participant As. A chi-square test rejects the null hypothesis that both distributions are equal ($p$-value = 0.04)\(^\text{10}\). The full sample mean amount sent by the Citizens in the role of sender is $M_{205.4}$ while that of the Participant As is $M_{261.1}$. Table 1 shows that the difference is statistically significant ($t$-test, $p$-value=0.04). The difference between the amount sent remains statistically significant even when conditioned on those who decide to send a positive amount ($t$-test, $p$-value=0.08). These results indicate that subjects who play the role of those at the receiving end of the bribery game - the Citizens, tend to trust less than those who play a strategically identical but differently framed role, the Participant As, in an ultimatum game.

Interestingly, it is the trust behavior in the BG, and not in the UG, which differs from the native distribution of trust as measured by the baseline trust treatment. This is indicated by the fact that both distribution and mean trust differs between BG and baseline (KS, $p$-value=0.10 and $t$-test, $p$-value=0.06). However, neither the distribution nor the means are different across the UG and baseline trust behavior (KS, $p$-value=0.80 and $t$-test, $p$-value=0.88). Our apprehension in Section 2 that the baseline trust may not be an appropriate counterfactual with which one can

\(^{10}\)A Kolmogorov-Smirnov test however fails to reject the null of equality of distributions in this case.
Figure 3: Bribery and Ultimatum Game

(a) Bribe/ultim demand and accept/reject

The figure on the left compares the bribe/ultim distribution in BG and UG. The one on the top right compares the mean bribe demanded with the mean ultim for the full sample while that on the bottom right compares the acceptance rates of the Citizens/Participant As for the bribe/ultim demand.

(b) Social Norm and Actual Behavior

The figure on the left compares the social norm of the bribe and ultim demand for each amount and the shaded region marks the difference. The figure on the right plots the number of subjects who demanded a particular amount of bribe/ultim in the BG/UG treatment. Clearly, the pattern of difference in social norm closely follows the difference in actual behavior.
compare the effect of corruption on trust turned out to be misplaced. However, the subsequent discussion mainly focuses on the comparison between the trust behavior of P-A in UG and that of C in BG.

Figure 4: Mean Trust in the two treatments.

The figure on the top lays out the distributions of the amount shared by a Citizen and Participant A in the trust game. The bottom left and bottom right figures compare the mean amount shared by C in BG, P-A in UG and baseline subjects for full sample (including those who did not share anything) and restricted sample (excluding those who did not share anything) respectively.

The difference in the trusting behavior between the BG and the UG treatment persists after controlling for the bribe/ultim amount demanded off a citizen and mean expected return on trust (ERoT), as shown in col (1) in Table 3. An increase in the amount of bribe reduces trust but trust and ultim are statistically uncorrelated (despite a negative correlation coefficient). This is also clear from Figure 5 and the pattern is broadly consistent with the stylized fact we observe in Figure 1. Instead of bribe/ultim demanded off a citizen, Col (2) - (5) controls for the social norm ratings of the bribe/ultim demand, besides other variables such as the average expected return on trust, demographics and preferences. Interestingly, the treatment effect disappears in these specifications with social norm as a covariate, but the social norm coefficient remains positive and statistically significant. The coefficients suggest that for each unit increase in mean social appropriateness rating, the average trust level increases by $M_{18}$ to $M_{19.5}$, depending on the specification. Mean ERoT has a positive statistically and economically significant coefficient, suggesting that an increase in
expected return increases trust levels. Except the negative coefficient for female - which indicates that females trust less than males - none of the demographic variables explain the trust levels. Col (5) of Table 3 reports the ordered probit\textsuperscript{11} estimates of the amount shared - the direction of the results are consistent with the ordinary least square estimates.

In Table 4 we briefly look at trust behavior of public officials. We find that trust is negatively associated with a demand for a bribe as indicated by col (1) and (3), though the strength of the association is weak. In col (4) - (5), we replace bribe demand with the social appropriateness rating corresponding to the bribe amount and find that the coefficient of social appropriateness measure is economically and statistically significant implying that POs who took a more socially appropriate action in the bribery game also trusted more in the trust game.

Figure 5: Amount demanded as bribe or ultim and trust levels

![Graph showing trust levels versus bribe demand for citizen and participant A.]

The panel on the right plots citizen’s trust with the bribe demanded off her while that on the left plots the relationship between Participant A’s trust and the ultim demanded off her. While the correlation in the left panel is -0.29 (p-value=0.05) that on the right is -0.12 (p-value=0.20). However, the difference in slope between in the two treatments are not statistically different.

\textsuperscript{11}Since trust or the amount sent by the sender in our case - $I$, is an ordinal and discrete variable, one may argue that the actual propensity to trust $I^*$ is latent and thus unobservable. The mapping between $I$ and $I^*$ is given by $I_i = x_i^\prime \beta + \varepsilon_i$ and $I_i = t$ if $s_{t-1} < I^* \leq s_t$, $t = 0, ..., T$. $\beta$ - the parameter of interest can then be obtained by an ordered probit estimate.
Table 1: Mean Differences in BG/UG and Trust Game

<table>
<thead>
<tr>
<th>Variables</th>
<th>BG</th>
<th>UG</th>
<th>Difference</th>
<th>p-valueb</th>
</tr>
</thead>
<tbody>
<tr>
<td>% who asked for bribe/ultim</td>
<td>81</td>
<td>100</td>
<td>19</td>
<td>0.00***</td>
</tr>
<tr>
<td>Bribe/ultim amount (Full Sample)</td>
<td>182.6</td>
<td>268.8</td>
<td>71.1</td>
<td>0.00**</td>
</tr>
<tr>
<td>Bribe/ultim amount (Restricted Sample)</td>
<td>245.9</td>
<td>268.8</td>
<td>22.9</td>
<td>0.02**</td>
</tr>
<tr>
<td>% who accepted when amount=100</td>
<td>93.5</td>
<td>97.8</td>
<td>4.3</td>
<td>0.31</td>
</tr>
<tr>
<td>% who accepted when amount=200</td>
<td>80.4</td>
<td>97.8</td>
<td>17.4</td>
<td>0.00***</td>
</tr>
<tr>
<td>% who accepted when amount=300</td>
<td>73.3</td>
<td>41.3</td>
<td>32.0</td>
<td>0.00***</td>
</tr>
<tr>
<td>% who accepted when amount=400</td>
<td>17.8</td>
<td>10.8</td>
<td>7</td>
<td>0.34</td>
</tr>
<tr>
<td>% who decided to trust</td>
<td>89.1</td>
<td>95.6</td>
<td>6.5</td>
<td>0.25</td>
</tr>
<tr>
<td>Trust Amount (Full Sample)</td>
<td>205.4</td>
<td>261.1</td>
<td>55.7</td>
<td>0.04**</td>
</tr>
<tr>
<td>Trust Amount (Restricted Sample)</td>
<td>230.5</td>
<td>273.3</td>
<td>42.8</td>
<td>0.08*</td>
</tr>
</tbody>
</table>

aThe restricted sample consists of only those who demanded a bribe/ultim or those who chose to trust.
b*** denotes significance at the 1% level, ** at the 5% level and * at the 10% level. For comparing proportions in the table, p-values from \( \chi^2 \) test of equality of proportions are reported. For comparison of sample mean, p-values from \( t \)-test are reported.

### 3.2 Trustworthiness behavior

To analyze the trustworthiness behavior, we convert the amount returned as a receiver to a Return Ratio (RR) where \( RR = \frac{\text{Amount Returned}}{\text{Amount Received}} \). Col (9) - (11) in Table 3 report the ordinary least square regression results of trustworthiness of the C/P-A - measured in terms of average Return Ratio. We find no evidence that there is a difference in the trustworthiness of Citizens and Participant As. Also, neither the bribe amount or the social appropriateness measure predict trustworthiness. This is evident from the proximity of the BG and UG Return Ratio curves which are plotted in 6(b) and is also borne out of the fact that for none of the trust categories, the difference in return ratio is statistically significant. The only variable which seems to capture trustworthiness is mean ERoT, indicating that a subject’s expectation about her matched partner’s trustworthiness seems to predict her own trustworthiness behavior. It is difficult to say whether subjects form their expectations about others’ trustworthiness from their own trustworthiness behavior or the other way round, but in a homogenous subject pool like ours, both mechanisms can potentially be at work (Sapienza et al., 2013). We also do observe the monotone concave shape of the Return Ratio curve observed in past studies (Bellemare and Kroger, 2007).

Interestingly, it is important to note from further disaggregation in Table 9 in the Appendix that, ERoT predicts trustworthiness for different trust categories. However, for higher amounts, the slope of ERoT is smaller for the bribery game treatment than the ultimatum game treatment. This indicates that the same level of ERoT has a higher effect on trustworthiness in UG than in BG, for higher trust levels. For lower amounts the direction of the interaction coefficient is same as that for higher amounts but the
behavior from the expectations they have about others trustworthiness behavior, the expectation shock suffered by the citizens seems to have lowered its response to trustworthiness. As in the above, it is fair to say that no causal link can be established between the two concepts from our set up\textsuperscript{13}. For an example of a study which studies the interplay between preference and beliefs, see Blanco et al. (Forthcoming).

Finally, the trustworthiness of PO, measured by mean Return Ratio of all trust categories, does not have any association with the bribe amount demanded, as indicated by results given in col (6) - (8) in Table 4.

3.3 Mechanisms

3.3.1 Theoretical framework

We present a formal set up through which the mechanism driving the spillover from corruption to trust may be examined. “Citizens” interact with “Public Officials” in order to obtain a certain public service. Public officials may be “honest” and “good” or “corrupt” and “bad”. Whether a citizen meets an honest public official or a corrupt one is entirely fortuitous. Following their interactions with the public officials, citizens form beliefs, $\mu$, about how likely public officials are to be trustworthy and then make a decision whether or not to trust a public official.

Assume that a person, who is honest, is more likely to be trustworthy, say with probability $h$ than a person who is corrupt, say $l$, such that $0 \leq l < h \leq 1$. Also assume that a Bayesian citizen gets an imperfect signal about trustworthiness of the set of matched partners, by observing whether she meets an honest official or a corrupt one in the first stage. Then she uses this signal to update his posterior belief that a matched person in the second stage is trustworthy. Note that unlike in typical applications where a Bayesian agent updates her belief about exactly the same individual she meets, we follow Acemoglu and Wolitzky (2014) and assume that a citizen learns something about how subjects, in the pool of matched public officials she is dealing with, are by observing examples of their behavior. Bayesian updating then tends to make citizens put more weight on the behavior they immediately observe.

Let $\mu_0$ be the prior belief that a matched partner is trustworthy. Suppose that a citizen meets an honest official in the first stage. His updated posterior about a matched partner’s trustworthiness in the second stage, $\mu$ is given by

$$
\mu(H) = \frac{\mu_0 h}{\mu_0 h + (1 - \mu_0)l}
$$

estimates are imprecisely estimated. With a larger sample, perhaps the coefficients corresponding to lower amounts too will be statistically significant.

\textsuperscript{13}We did try to estimate the relation simultaneously taking advantage of the fact that risk preference affects ERoT but not trustworthiness and thus makes the system of equations identifiable, however no significant results were found.
This follows in view of the fact that the official is either honest and trustworthy (probability $\mu_0 h$) or an honest official coming from that subset of officials who are not trustworthy (probability $(1 - \mu_0) l$). Similarly, given that a citizen meets a corrupt person in the first stage, his updated posterior is given by

$$
\mu(L) = \frac{\mu_0(1 - h)}{\mu_0(1 - h) + (1 - \mu_0)(1 - l)}
$$

(2)

It is easy to see that the $\mu(H) > \mu_0$ and $\mu(L) < \mu_0$ i.e. the posterior belief of a citizen about meeting a trustworthy person in the future is higher if she meets an honest public official while it is lower if she meets a corrupt public official. This leads to our first prediction.

**Prediction 1.** If a citizen meets a corrupt (honest) public official, then her belief about the trustworthiness of the person she will meet in the second stage decreases (increases).

The updated beliefs play an important role in the subsequent trust decision, which we model as that in a simple trust game. Assume that the initial endowment in the trust game is given by $\bar{x}$, sharing decision by $t$ and there are only two types of receivers with whom the sharing decisions are made - one, less trustworthy i.e. those who keep a high amount with themselves (say $\bar{w}$) and two, more trustworthy i.e. those who keep a small amount with themselves (say $w$). Preferences of those making the trust decisions are given by $U(., \lambda, \theta)$ where $\lambda$ and $\theta$ are the risk (implicit in the curvature of the utility function) and pro-sociality paramater.

Expected Utility of an individual who shares $s$ as a sender in the trust game is given by -

$$
EU = \mu U(\bar{x} - t + (3t - \bar{w}), \lambda, \theta) + (1 - \mu) U(\bar{x} - t + (3t - w), \lambda, \theta)
$$

(3)

Assuming an interior solution\textsuperscript{14}, the first order condition is given by

$$
2\mu U'(\bar{x} + 2t^* - \bar{w}, \lambda, \theta) + 2(1 - \mu) U'(\bar{x} + 2t^* - w, \lambda, \theta) = 0
$$

$$
t^* = t^*(\mu, \lambda, \theta)
$$

(4)

Plugging equation (1) in (4) gives the optimal trust decision if one meets an honest person before, $t^*(H) = t^*(\mu(H), \lambda, \theta)$, while plugging equation (2) in (4) gives the optimal trust decision if one meets a dishonest person before, $t^*(L) = t^*(\mu(L), \lambda, \theta)$.

**Prediction 2.** If a citizen meets a corrupt (honest) public official, then the trust shown by the citizen decreases (increases).

\textsuperscript{14}In this framework $t^* > 0$ iff $EU(t^*) \geq U(\bar{x})$ i.e. an individual decides to share iff his expected return from sharing is higher than that from retaining the entire amount with himself.


Totally differentiating the first order condition with respect to \( t \) and \( \mu \), we can show that \( \frac{dt^*}{d\mu} > 0 \). Since \( \mu(H) > \mu_0 \) and \( \mu(L) < \mu_0 \) implies \( \mu(H) > \mu(L) \), it follows that \( t^*(\mu(H), \lambda, \theta) > t^*(\mu(L), \lambda, \theta) \).

Our interest lies in estimating how trust responds to whether a citizen meets an honest official or a corrupt one. We therefore estimate equation (4) from the framework described above. The linear approximation to the functional form in equation (4) is

\[
t = \beta_{\mu} \mu + \beta_{\lambda} \lambda + \beta_{\theta} \theta + \epsilon
\]

(5)

The specification in (5) shows that trust decision will depend on the belief of the citizen about how trustworthy the trustee is, the risk preference and pro-sociality. However, notice that \( \mu \) is determined at the previous stage and thus equation (5) should be jointly estimated with equation (6)

\[
\mu = \gamma 1(H) + \nu
\]

(6)

where \( 1(H) \) is a dummy variable indicating whether the citizen meets an honest official or not.

The data from the experiment allows us to jointly estimate (5) and (6), as we shall demonstrate in the next subsection.

### 3.3.2 Empirical Validity

In order to pin down the precise channel through which corruption affects trust, we analyze the data collected on expectations (as proxy for beliefs described in Section 4.1) and preferences. For this purpose we define Expected Return on Trust as \( ERoT = \frac{(Expected\ Return\ for\ a\ trust\ level - Trustlevel)}{Trustlevel} \). Figure 6(a) plots the average expected return on trust for each possible trust levels using the full strategy vector data. Two observations stand out - first, the expected return is significantly less for the Cs in BG than the P-As in UG, especially for lower levels of trust; second, \( ERoT \) is negative for largely negative, specially for lower levels of trust. The first observation tells us that citizens have suffered a negative shock in their expectations about trustworthiness of their matched partners, when compared to Participant A or baseline subjects. Second, subjects by and large display low expectations (largely negative) about the trustworthiness when full strategy vector is considered though their expectations corresponding to their actual response is greater than zero and thus consistent (not reported).
Figure 6: Expected Return on Trust

6(a)

Note: The panel plots the average Expected Return on Trust for full strategy vector data of subjects in baseline and those playing Citizen and Participant A in BG and UG, respectively.

6(b)

Note: The above figure plots the average return ratio (i.e. trustworthiness), for all possible bribe amounts, of subjects in baseline and those playing Citizen and Participant A in BG and UG, respectively.

The OLS estimates of col (2) - (4) and the ordered probit estimates of col (5) show that the bribery frame by itself does not lead to the treatment effect, and thus the treatment effect is most likely not generated through priming or mindset effects. However, the positive, statistically and economically significant coefficients corresponding to the average social norm shows that if a Citizen/Participant A faces a bribe/ultim demand which is considered socially inappropriate then she trusts less. Now, combine this with the finding that average social appropriateness measure of bribe demand in the BG, -0.38, is significantly lower than that of the ultim demand in UG, 0.04 (MW test, p-value=0.04). This implies that Citizens have faced more socially inappropriate demands on average than Participant As and this in turn has generated lower trust.

Besides social norm, ERoT, averaged over all trust categories, turns out to be positive and
statistically significant when it is used as a covariate in the regression of trust (col (2) - (5) in Table 3) and this implies that the expectation of trustworthiness of the matched partner predicts trust. However, it is important to note that these specifications yield inconsistent estimates since the mean ERoT is endogenously determined. Not only is this apparent conceptually from equation (5) and (6) in Section 4.1, but also empirically from Fig 6 where ERoT is seen to systematically vary with treatment. Thus treatment dummy, and also possibly social appropriateness measure, may affect the expected trustworthiness in the first stage while trust decisions may also be influenced by demographic characteristics in the second stage, besides mean ERoT and preferences. Hence, the structural equations in terms of the experimental data are given by

\[
TRUST = \beta_0 + \beta_1 E(\text{mean ERoT}|BG, NORMMEAN) + \beta_2 X + \beta_\lambda + \beta_\theta + \epsilon \tag{7}
\]

\[
\text{mean ERoT} = \gamma_0 + \gamma_1 BG + \gamma_2 NORMMEAN + \nu \tag{8}
\]

where \(TRUST\) is the trust decision of the sender, \(X\) is the set of exogenous characteristics of the subject, \(BG\) is the treatment dummy and \(NORMMEAN\) is the exogenously elicited appropriateness measure of the bribe/ultim demand faced by the subject. The joint estimation of equations (7) and (8) assumes that the appropriateness of the bribe/ultim demand and the treatment dummy affects trust through the expected return on trust but not independently i.e. the exclusion restriction holds. Our parameters of interest are \(\beta\) which is consistently and unbiasedly estimated using a 3SLS procedure.

Col (6) - (8) in Table 3 reports the second stage results. Col (6) and (7) show that the coefficient of ERoT is positive and significant at 10% level and that in col (8) it is marginally insignificant with a \(p\)-value=0.10. Interestingly, the first stage results show that the treatment dummy has a negative and significant effect and mean social appropriateness measure a positive and significant effect on mean ERoT. The significant negative effect of the treatment dummy reflects that there is a mindset effect working insofar as the expectations of trustworthiness is concerned but there is no independent effect of the treatment on trust (not reported) i.e. the very mindset of being in a potentially corrupt environment reduces expectations about a prosocial and fair behavior. The positive significant effect of social norm on expectations of trustworthiness, on the other hand, indicates that subjects respond positively to a fair outcome, not necessarily an outcome beneficial for themselves. Note from Figure 3b that while a bribe amount and its fairness has a negative and monotonic relation, less ultim is not always considered fair - the peak of social appropriateness in UG is in fact at equal split. Table 2 compares the difference in trust and mean ERoT for

\[15\text{Formally we test for endogeniety using a Durbin-Wu-Hausman test and this leads to a rejection of the null hypothesis that the mean ERoT is exogenous (\(p\)-value=0.07).}
Table 2: Trust, Mean ERoT and high and low social appropriateness

<table>
<thead>
<tr>
<th>UG, BG Amount</th>
<th>Trust Difference</th>
<th>p-value</th>
<th>Difference in Mean ERoT</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>200,000</td>
<td>4</td>
<td>0.94</td>
<td>0.07</td>
<td>0.60</td>
</tr>
<tr>
<td>400,400</td>
<td>10</td>
<td>0.90</td>
<td>-0.38</td>
<td>0.13</td>
</tr>
</tbody>
</table>

*a* Most socially appropriate choice

*b* Most socially inappropriate choice

greatest social appropriateness and lowest social appropriateness in UG and BG. The greatest social appropriateness in UG and BG correspond to an ultim and bribe demand for UG and BG, respectively and least social appropriateness in the two frames correspond to ultim/bribe amount of 400 each. Two other observations stand out: one, there is no statistical difference in mean ERoT between subjects who received the bribe/ultim demand which is considered most fair and least fair and two, there is no statistical difference in trust between subjects who received the bribe/ultim demand which is considered most fair and least fair. This indicates that it is social norm of bribe/ultim and not the amount of bribe/ultim per se which significantly affects mean ERoT and the subsequent trust decisions. This leads us to the interesting conclusion that the underlying model which predicts trust decisions is actually one which is determined by whether an individual has faced norm violation or not\(^\text{16}\). So why is it that the trust decisions differed in the two frames? The reason is that the socially inappropriate demand was exercised more frequently in BG than in UG.

Therefore, results from the joint estimation of ERoT and trust confirm the following mechanism at work: the corruption frame decreases expectations of trustworthiness and the social appropriateness of bribe/ultim demand increases the same. The expectations of trustworthiness in turn then determine the actual trust decisions. Thus the first stage results affirm Prediction 1 while the second stage results confirm Prediction 2 as obtained from the theoretical framework discussed earlier.

We also collected data on altruistic and risk preferences, which we use to analyze if preferences can help explain the trust pattern which we observe. We follow Dohmen et al. (2011), who find that the approach of asking people a survey question about their “willingness to take risks in general ... generates a useful all-round measure”, to obtain the risk profile (proxy for $\gamma$ in Section 3.3.1) of the subjects. For altruistic preference, we ask the subjects how much they will be willing to share with a charity if they win a lottery of Rs. 1000 (proxy for $\theta$ in Section 3.3.2). The general risk measure turns out to be a significant predictor of trust as is clear from Col (4), (5) and (8) (and elsewhere\(^\text{16}\).

\(^{16}\) We estimated the trust decision model separately for BG and UG with norm as the covariate, a chow test however rejects the null of equality of the coefficients corresponding to norm for the two subsamples. It is possible that, as Table 2 indicates, at the corner (highest and least social appropriateness) norm violation predicts trust but not at the interiors.
in the analysis) in Table 3 but altruism does not\(^{17}\). Neither risk nor altruism is significant when interacted through the treatment variable (not reported).

To summarize, the discussion above indicates that the variation in trust can be explained by the treatment and social appropriateness norm which affect trust through the mean ERoT. In other words, the bribery game treatment and the social (in)appropriateness of the bribe/ultim lowers the expectation about trustworthiness and this in turn reduces the trust levels.

### 3.4 WVS-trust Question and trust behavior in the experiment

In this section we are interested in examining how the WVS-trust question correlates with the trust behavior in the experiment and then in comparing the correlation across the two treatments. Hence for the rest of the analysis we restrict our attention to response from Citizens, Participant As and the baseline respondents in the trust game experiment.

Table 5 reports the probit regression of the WVS-trust on trust and trust worthiness. Col (1) shows that the trust decision in the experiment is not related to WVS-trust - not even when the relation is examined separately for C, P-A and baseline (not reported). Col (2) - (9) report the coefficients for trustworthiness (i.e. return ratio) for each trust category for all possible amounts sent after controlling for the expected return on trust. If ERoT is not included as a covariate in the regression, then the coefficients of the return ratios are positive and statistically significant only for higher amounts sent (not reported). However, once ERoT is included, we notice that WVS-trust is actually predicted by ERoT and not the trustworthiness - a 10% increase in ERoT increases WVS-trust by 2.2 to 3.8 percentage points (ppts). This indicates that subjects’ responses to WVS-trust are not independently associated with trustworthiness. However, WVS-trust is associated with the expectations that subjects form about the matched receivers’ responses and these expectations in turn play a role in determining their trustworthiness. In other words, WVS-trust captures the expectation of how trustworthy the matched partner is, when the stakes are high.

We now take a closer look at how WVS-trust captures the belief component of the senders belonging to different groups. To get a preliminary sense of the relation, we report the probit regression estimates of WVS-trust on ERoT, treatment dummy and their interaction in Table 10 in the Appendix. ERoT is a significant predictor of WVS-trust - a 10% increase in ERoT increases the probability that WVS-trust will take value 1 by 2.8 to 10.8 ppts. However, the negative coefficients for interaction terms, specially corresponding to higher amounts, reflect that the response of ERoT to WVS-trust is significantly higher for UG than for BG. In other words, the same level of ERoT leads to a greater increase in WVS-trust in the UG treatment.

To understand this further, in Table 6, we report the probit regression results of WVS-trust on Expected Return on Trust for all possible amounts sent and the risk preference. Panel (A) reports

\(^{17}\)We expected altruism to play a role in trust behavior, particularly we expected to see a decrease in altruism in Citizens. It turned out to be true but statistically insignificant. Perhaps altruism could have been measured in a better way.
the results for Citizens, Participant As and baseline subjects, panel (B) for P-A and baseline, panel (C) for P-A only and panel (D) for C only. Interestingly, the coefficient of ERoT corresponding to most trust levels is statistically significant in Panel (A), (B) and (C) but not in the case of Panel (D). Besides, the coefficients are economically significant too - a 10% increase in ERoT leads to a 2.4 ppts to 6.8ppts increase in the probability that WVS-trust will take value 1 in Panel B. The corresponding figures in Panel (C) are 3.1 ppts to 8.8 ppts in our sample\textsuperscript{18}. As the hypothetical amount sent increases, the strength of the association between expectations about others’ trustworthiness and response to WVS-trust increases. Also, it increasingly better explains the variation in WVS as is indicated by the increase in Pseudo-$R^2\textsuperscript{19}$ as the hypothetical amount sent increases.

This indicates that not only do the expectations of the Citizens affect the WVS-trust less than that of the P-As, the expectations of citizens do not at all predict the response to WVS-trust. Whereas ERoT does predict WVS-trust question for the P-As and the baseline respondents. This, in conjunction with our observation from Figure 6 that ERoT is lower for Citizens than for P-As and baseline, indicates that WVS-trust is stable in the short run and is not affected by the corruption frame. On the other hand, expectations or beliefs about how others will behave, are subject to short run fluctuations. This suggests that the response to WVS-trust comes from a global perception about expectations of trustworthiness in general whereas the expectations measured in experimental games are locally determined.

\section{Discussion and Conclusion}

In real life people often interact with each other strategically in many different settings. These interactions subsequently affect our priors which in turn influence our subsequent decisions. However, the majority of experimental research focuses on behavior in experimental games in isolation. To mimic some of the real life situations of interrelated behavioral effects, past studies have designed experiments that generate “behavioral spillover effects” (Knez and Camerer, 2000; Ahn et al., 2001; Cason et al., 2012). Such spillovers can be positive when coordination or cooperation improves (Cason et al., 2012) or negative when it reduces cooperation (Buser and Dreber, 2013). These effects are different from the psychological effects of priming and pure mindset effects. While priming arises from very subtle interventions, which works even by simply reminding people of some priming elements, behavioral spillovers arise when there are real consequences in terms of monetary payoffs.

In this paper we provide evidence of both negative spillover - effect of corrupt transaction on trust, and no spillover - no effect of ultimatum game on trust. Subjects are randomized into

\textsuperscript{18}Following Kessler and Vesterlund (Forthcoming), we are more confident about the direction of the marginal effect than its precise size which may change with a larger and/or a different sample.

\textsuperscript{19}It is well known that Pseudo-$R^2$ cannot be interpreted as a $R^2$, but the former still has a meaning when compared to another Pseudo-$R^2$ of the same type and predicting the same outcome in the same data. In such a situation, a comparison of Pseudo-$R^2$ indicates which model better predicts the outcome.
either a corruption game or a strategically identical but differently framed ultimatum game. Both treatments are followed by a trust game. The negative spillover effect of corruption on trust provides a causal link that corruption leads to lower trust. Furthermore, we provide an insight into what drives the negative spillover effect. Evidence suggests that a demand for bribe violates the social appropriateness norm and norm violation leads to a negative shock in the belief that the matched partner is trustworthy. Since actual trust behavior is shaped in part by the belief about how trustworthy a matched partner is, the negative shock in the latter immediately triggers a decrease in the level of trust. Why do corrupt behavior cause a negative shock in the belief about others’ trustworthiness? We show that the violation of social appropriateness norms in the bribery game leads the Citizens to lower their prior about the prevalence of norm violators in the trust game. With a readjusted prior about trustworthiness, the optimal response in the trust game is one with lower trust levels. Interestingly, while negative spillover effect can be attributed to norm violation, Gächter et al. (2013) find that observed peer effect in their gift exchange game is explained by social preferences rather than social norm. Also, it is important to note that the spillover effect of the kind we find in our experiment is perhaps short-lived. However, we conjecture that repeated interactions in a corrupt environment may lead to a behavioral stationary state of low trust. Of course our experiment, being a one shot game, has no way of verifying this conjecture.

We also analyze another related issue in the trust literature. We comment on the interpretation of the question on trust which is widely used in surveys like WVS and GSS. Like Sapienza et al. (2013), we find that the WVS-trust measures the belief about others’ trustworthiness. This re-examination is conducted in India - a society with a relatively low level of generalized trust. We thus examine the behavioral underpinnings of the WVS-trust question in a society whose characteristics are markedly different from societies where such analysis has been conducted before. Hence, this exercise by itself is important to validate the claim that WVS in general and the WVS-trust question in particular is interpreted in a similar manner across societies which are fundamentally differently organized.

Does this then imply that the WVS-trust, with its convenience of being a survey question, may be widely employed to measure people’s belief about how trustworthy others are? Our results suggests that we need to be careful in this regard. We find that while the belief about others’ trustworthiness is malleable or subject to short term fluctuations, WVS-trust is more stable in the shorter run. The implication of this finding is the following. After the important work of Dohmen et al. (2011), who find a robust correspondence between elicited risk in experiments and a survey question on attitude to risk in general, a number of studies, including ours, have come to rely on the survey question about risk in general to measure risk preference. The convenience of replacing a treatment by a question is hard to overstate. If similar questionnaires were to be found to robustly measure other behavioral correlates, then we could replace costly treatments with such questionnaires. Our study shows that in the context of measuring beliefs about others’ trustworthiness, we cannot make a blanket use of the WVS-trust question since it cannot capture
the short term changes in locally formed beliefs. The reference group, which dictates behavior in
trust games, especially in a homogenous subject pool like in our case, is typically different from the
reference group people have in mind when they answer the WVS-trust question. As a result, trust
games and the WVS-trust question remain relevant in their respective domains of inquiry, namely
behavioral analysis in the short run and generalized levels of trust in the long run, respectively.

References


Table 3: Citizen’s Trust and Trustworthiness

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
<th>(11)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>trust</td>
<td>trustworthiness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bg</td>
<td>-61.65**</td>
<td>-32.79</td>
<td>-28.35</td>
<td>-32.31</td>
<td>-0.30</td>
<td>0.02</td>
<td>0.03</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(26.56)</td>
<td>(25.33)</td>
<td>(25.76)</td>
<td>(25.19)</td>
<td>(0.24)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>buamount</td>
<td>-0.29**</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.12)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>normmean</td>
<td>19.49**</td>
<td>19.24**</td>
<td>18.12**</td>
<td>0.18**</td>
<td>-0.01</td>
<td>-0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(7.96)</td>
<td>(7.88)</td>
<td>(7.81)</td>
<td>(0.07)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>meanerot</td>
<td>99.38***</td>
<td>99.33***</td>
<td>100.34***</td>
<td>89.68**</td>
<td>0.18**</td>
<td>-0.01</td>
<td>-0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(35.84)</td>
<td>(35.58)</td>
<td>(35.44)</td>
<td>(0.34)</td>
<td>(0.07)</td>
<td>(0.07)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>age</td>
<td>-2.99</td>
<td>-4.42</td>
<td>-0.04</td>
<td>-3.02</td>
<td>-4.49</td>
<td>-0.00</td>
<td>-0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(7.02)</td>
<td>(6.90)</td>
<td>(0.06)</td>
<td>(6.99)</td>
<td>(7.00)</td>
<td>(0.06)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>female</td>
<td>-59.08**</td>
<td>-52.34*</td>
<td>-0.51*</td>
<td>-59.31**</td>
<td>-51.55*</td>
<td>-0.01</td>
<td>-0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(29.35)</td>
<td>(28.61)</td>
<td>(0.27)</td>
<td>(29.82)</td>
<td>(31.30)</td>
<td>(0.27)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>catscore</td>
<td>-4.47</td>
<td>-4.07</td>
<td>-0.05</td>
<td>-4.46</td>
<td>-4.10</td>
<td>-0.01</td>
<td>-0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.54)</td>
<td>(3.44)</td>
<td>(0.04)</td>
<td>(3.51)</td>
<td>(3.50)</td>
<td>(0.04)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fam_income</td>
<td>-4.44</td>
<td>-1.40</td>
<td>0.01</td>
<td>-3.93</td>
<td>-3.58</td>
<td>0.03*</td>
<td>0.03*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(16.21)</td>
<td>(15.78)</td>
<td>(0.15)</td>
<td>(19.29)</td>
<td>(19.41)</td>
<td>(0.15)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>altruism</td>
<td>0.03</td>
<td>0.00</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.00)</td>
<td>(0.06)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>risk_general</td>
<td>16.62**</td>
<td>0.17***</td>
<td>17.64**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(6.72)</td>
<td>(0.06)</td>
<td>(7.03)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>334.51***</td>
<td>255.32***</td>
<td>782.06**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(36.58)</td>
<td>(17.56)</td>
<td>(357.15)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First Stage</td>
<td>bg</td>
<td>-0.13*</td>
<td>-0.12*</td>
<td>-0.11*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.07)</td>
<td>(0.07)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>normmean</td>
<td>0.03*</td>
<td>0.03*</td>
<td>0.03*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>91</td>
<td>91</td>
<td>91</td>
<td>91</td>
<td>91</td>
<td>91</td>
<td>91</td>
<td>91</td>
<td>91</td>
<td>91</td>
<td></td>
</tr>
<tr>
<td>R-squared</td>
<td>0.180</td>
<td>0.180</td>
<td>0.236</td>
<td>0.297</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pseudo R-squared</td>
<td>0.0846</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log Likelihood</td>
<td>-172.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

† The coefficient is marginally insignificant with p-value=0.10
†† Note: Numbers in the parentheses are Standard errors. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable in col (1) - (7) is trust, that in col (9) - (11) is trustworthiness measured as mean Return Ratio. Col (5) reports the ordered probit estimation. Col (6) - (8) jointly estimate mean ERoT and trust.
Table 4: Public Official’s Trust and Trustworthiness

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>bribe</td>
<td>-0.28*</td>
<td>-0.26</td>
<td>-0.30*</td>
<td>-0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.15)</td>
<td>(0.16)</td>
<td>(0.15)</td>
<td>(0.00)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>normmean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>age</td>
<td>0.24</td>
<td>0.04</td>
<td>-0.46</td>
<td>-0.62</td>
<td>0.02**</td>
<td>0.02**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(11.83)</td>
<td>(11.19)</td>
<td>(11.81)</td>
<td>(11.17)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>female</td>
<td>-17.92</td>
<td>-21.01</td>
<td>-19.44</td>
<td>-22.30</td>
<td>-0.03</td>
<td>-0.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(43.78)</td>
<td>(41.48)</td>
<td>(43.56)</td>
<td>(41.27)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>catscore</td>
<td>-0.40*</td>
<td>-0.36*</td>
<td>-0.39*</td>
<td>-0.35*</td>
<td>-0.00***</td>
<td>-0.00***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.21)</td>
<td>(0.20)</td>
<td>(0.21)</td>
<td>(0.20)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>fam_income</td>
<td>17.92</td>
<td>10.29</td>
<td>18.32</td>
<td>10.55</td>
<td>-0.01</td>
<td>-0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(28.66)</td>
<td>(27.28)</td>
<td>(28.48)</td>
<td>(27.11)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>altruism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>risk_general</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>305.65***</td>
<td>278.84</td>
<td>232.78</td>
<td>249.98</td>
<td>199.15</td>
<td>0.43***</td>
<td>-0.02</td>
<td>-0.11</td>
</tr>
<tr>
<td></td>
<td>(32.10)</td>
<td>(317.78)</td>
<td>(310.25)</td>
<td>(312.12)</td>
<td>(305.51)</td>
<td>(0.04)</td>
<td>(0.30)</td>
<td>(0.30)</td>
</tr>
<tr>
<td>Observations</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.076</td>
<td>0.159</td>
<td>0.287</td>
<td>0.169</td>
<td>0.295</td>
<td>0.010</td>
<td>0.345</td>
<td>0.422</td>
</tr>
</tbody>
</table>

† Note: Numbers in the parentheses represent the Standard errors. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable in col (1) - (5) is trust, that in col (6) - (8) is trustworthiness measured in terms of Return Ratio.
Table 5: Marginal Effects from Probit Regression of WVS-Trust on trust and trustworthiness

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>trust</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>trustworthiness</td>
<td>0.00</td>
<td>-0.07</td>
<td>-0.21</td>
<td>-0.32</td>
<td>-0.03</td>
<td>0.15</td>
<td>0.06</td>
<td>-0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.16)</td>
<td>(0.23)</td>
<td>(0.29)</td>
<td>(0.31)</td>
<td>(0.38)</td>
<td>(0.39)</td>
<td>(0.39)</td>
<td>(0.38)</td>
<td></td>
</tr>
<tr>
<td>ERoT</td>
<td>0.04</td>
<td>0.11</td>
<td>0.27***</td>
<td>0.35***</td>
<td>0.22*</td>
<td>0.30**</td>
<td>0.33**</td>
<td>0.38***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.08)</td>
<td>(0.10)</td>
<td>(0.12)</td>
<td>(0.13)</td>
<td>(0.14)</td>
<td>(0.13)</td>
<td>(0.13)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>127</td>
<td>127</td>
<td>127</td>
<td>127</td>
<td>127</td>
<td>127</td>
<td>127</td>
<td>127</td>
<td>127</td>
</tr>
<tr>
<td>Pseudo R-sq</td>
<td>0.00807</td>
<td>0.00246</td>
<td>0.0127</td>
<td>0.0415</td>
<td>0.0492</td>
<td>0.0215</td>
<td>0.0435</td>
<td>0.0500</td>
<td>0.0539</td>
</tr>
</tbody>
</table>

† Note: Numbers in the parentheses represent the Standard errors. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is WVS-trust. The independent variable in col (1) is trust while that in col (2) - (9) is trustworthiness measured in terms of Return Ratio for each possible amount sent by the sender. All regressions have been run with a constant term.
Table 6: Marginal Effects from Probit Regression of WVS-Trust on ERoT and Risk

<table>
<thead>
<tr>
<th>Panel (A)</th>
<th>if sent 50</th>
<th>if sent 100</th>
<th>if sent 150</th>
<th>if sent 200</th>
<th>if sent 250</th>
<th>if sent 300</th>
<th>if sent 350</th>
<th>if sent 400</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERoT</td>
<td>0.05</td>
<td>0.09</td>
<td>0.21**</td>
<td>0.24**</td>
<td>0.20*</td>
<td>0.29***</td>
<td>0.33***</td>
<td>0.34***</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.06)</td>
<td>(0.08)</td>
<td>(0.10)</td>
<td>(0.11)</td>
<td>(0.11)</td>
<td>(0.11)</td>
<td>(0.11)</td>
</tr>
<tr>
<td>risk_general</td>
<td>0.08***</td>
<td>0.08***</td>
<td>0.08***</td>
<td>0.07***</td>
<td>0.08***</td>
<td>0.07***</td>
<td>0.08***</td>
<td>0.07***</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Observations</td>
<td>127</td>
<td>127</td>
<td>127</td>
<td>127</td>
<td>127</td>
<td>127</td>
<td>127</td>
<td>127</td>
</tr>
<tr>
<td>Pseudo R-sq</td>
<td>0.0868</td>
<td>0.0930</td>
<td>0.119</td>
<td>0.117</td>
<td>0.102</td>
<td>0.122</td>
<td>0.137</td>
<td>0.137</td>
</tr>
</tbody>
</table>

| Panel (B) | ERoT      | 0.06       | 0.11        | 0.24**      | 0.25**      | 0.27**      | 0.49***     | 0.57***     | 0.68***     |
|           |           | (0.07)     | (0.08)      | (0.10)      | (0.11)      | (0.13)      | (0.15)      | (0.15)      | (0.12)      |
| risk_general | 0.08***    | 0.07***    | 0.07***     | 0.07***     | 0.07***     | 0.07***     | 0.07***     | 0.07***     |
|           |           | (0.02)     | (0.02)      | (0.02)      | (0.02)      | (0.02)      | (0.02)      | (0.02)      | (0.02)      |
| Observations | 81         | 81         | 81          | 81          | 81          | 81          | 81          | 81          |
| Pseudo R-sq | 0.094      | 0.104      | 0.134       | 0.128       | 0.125       | 0.165       | 0.184       | 0.242       |

| Panel (C ) | ERoT      | 0.06       | 0.12        | 0.31**      | 0.29        | 0.33        | 0.88***     | 0.65***     | 0.69***     |
|            |           | (0.10)     | (0.11)      | (0.15)      | (0.19)      | (0.22)      | (0.28)      | (0.23)      | (0.16)      |
| risk_general | 0.08**     | 0.08**     | 0.08***     | 0.07**      | 0.06**      | 0.05**      | 0.06**      | 0.06**      |
|            |           | (0.03)     | (0.03)      | (0.03)      | (0.03)      | (0.03)      | (0.03)      | (0.03)      | (0.03)      |
| Observations | 45         | 45         | 45          | 45          | 45          | 45          | 45          | 45          |
| Pseudo R-sq | 0.0713     | 0.0848     | 0.125       | 0.102       | 0.105       | 0.216       | 0.187       | 0.265       |

| Panel (D) | ERoT      | -0.01      | 0.02        | 0.14        | 0.22        | 0.01        | 0.03        | 0.10        | -0.14       |
|           |           | (0.11)     | (0.10)      | (0.16)      | (0.21)      | (0.19)      | (0.17)      | (0.15)      | (0.17)      |
| risk_general | 0.09***    | 0.09***    | 0.09***     | 0.09***     | 0.09***     | 0.09***     | 0.09***     | 0.09***     |
|           |           | (0.03)     | (0.03)      | (0.03)      | (0.03)      | (0.03)      | (0.03)      | (0.03)      | (0.03)      |
| Observations | 46         | 46         | 46          | 46          | 46          | 46          | 46          | 46          |
| Pseudo R-sq | 0.110      | 0.111      | 0.123       | 0.128       | 0.110       | 0.111       | 0.119       | 0.122       |

† Note: Numbers in the parentheses represent the Standard errors. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is WVS-trust. The independent variables are Expected Return on Trust (ERoT) and risk preference. Panel A includes Citizens(C), Participant As(P-A) and baseline subjects (BS), Panel B includes P-A and BS, Panel C includes P-A only and Panel D includes C only. The number of observations differs in the panels due to inclusion of different set of subjects in each panel.
Appendix 1

Table 7: Summary Statistics

<table>
<thead>
<tr>
<th>BG/UG Variables</th>
<th>Description</th>
<th>Mean</th>
<th>Std. Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>bg</td>
<td>=1 if Treatment is BG</td>
<td>0.51</td>
<td>0.50</td>
</tr>
<tr>
<td>buamount</td>
<td>bribe/ultim demanded from C/P-A</td>
<td>232.97</td>
<td>109.60</td>
</tr>
<tr>
<td>bribe</td>
<td>bribe demanded by PO in BG of a</td>
<td>225.27</td>
<td>110.14</td>
</tr>
<tr>
<td></td>
<td>bribe/ultim amount</td>
<td></td>
<td></td>
</tr>
<tr>
<td>normmean</td>
<td>Social appropriateness measure</td>
<td>-0.17</td>
<td>1.57</td>
</tr>
<tr>
<td>Observations</td>
<td></td>
<td>91</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Trust Game &amp; Demographic Variables</th>
<th>Full Sample</th>
<th>C+P-A+Baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td>trust</td>
<td>245.41</td>
<td>232.97</td>
</tr>
<tr>
<td>trustworthiness</td>
<td>0.39</td>
<td>0.39</td>
</tr>
<tr>
<td>meanerot</td>
<td>-0.02</td>
<td>-0.02</td>
</tr>
<tr>
<td>female</td>
<td>0.26</td>
<td>0.25</td>
</tr>
<tr>
<td>age</td>
<td>24.37</td>
<td>24.35</td>
</tr>
<tr>
<td>catscore</td>
<td>100.32</td>
<td>95.06</td>
</tr>
<tr>
<td>fam_income</td>
<td>3.56</td>
<td>3.57</td>
</tr>
<tr>
<td>altruism</td>
<td>294.27</td>
<td>283.52</td>
</tr>
<tr>
<td>risk_general</td>
<td>5.76</td>
<td>5.90</td>
</tr>
</tbody>
</table>

Cronbach’s alpha for Return Ratio $^{(b)}$ 0.86 0.86
Cronbach’s alpha for ERoT $^{(b)}$ 0.92 0.90
Observations 218 91

$^{(a)}$Categorical variable where category 1: <Rs. 20,000, 2: between Rs. 20,001 and Rs. 50,000, 3: between Rs. 50,001 and Rs. 1,00,000 and 4: >Rs. 1,00,001.

$^{(b)}$A measure of internal consistency, i.e. how closely related the set of items are as a group.
Table 8: Frequency distribution of norm ratings for each bribe/ultim

<table>
<thead>
<tr>
<th>Bribe/Ultim</th>
<th>-3</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>Mean UG</th>
<th>-3</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>Mean BG</th>
<th>p-value&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>2.5</td>
<td></td>
<td>10</td>
<td>10</td>
<td>5</td>
<td></td>
<td>0</td>
<td>30</td>
<td>-0.43</td>
<td>7.5</td>
</tr>
<tr>
<td>100</td>
<td>12.5</td>
<td>32.5</td>
<td>10</td>
<td>5</td>
<td>10</td>
<td>30</td>
<td>0</td>
<td>-0.43</td>
<td>0</td>
<td>10</td>
<td>25</td>
<td>10</td>
<td>20</td>
<td>35</td>
<td>0</td>
<td>0.45</td>
<td>0.03</td>
</tr>
<tr>
<td>200</td>
<td>0</td>
<td>5</td>
<td>2.5</td>
<td>25</td>
<td>10</td>
<td>30</td>
<td>27.5</td>
<td>1.4</td>
<td>0</td>
<td>30</td>
<td>17.5</td>
<td>27.5</td>
<td>12.5</td>
<td>7.5</td>
<td>5</td>
<td>-0.35</td>
<td>0.00</td>
</tr>
<tr>
<td>300</td>
<td>5</td>
<td>25</td>
<td>17</td>
<td>0</td>
<td>30</td>
<td>17.5</td>
<td>5</td>
<td>-0.03</td>
<td>27.5</td>
<td>32.5</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1.88</td>
<td>0.00</td>
</tr>
<tr>
<td>400</td>
<td>97.5</td>
<td>2.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-2.98</td>
<td>92.5</td>
<td>2.5</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-2.68</td>
<td>0.29</td>
</tr>
</tbody>
</table>

<sup>a</sup>Mann Whitney Rank Sum test reported p-value of the mean difference. Shaded cells for each category in each treatment denote the modal ratings.

Table 9: Trustworthiness and ERoT

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>if sent 50</th>
<th>if sent 100</th>
<th>if sent 150</th>
<th>if sent 200</th>
<th>if sent 250</th>
<th>if sent 300</th>
<th>if sent 350</th>
<th>if sent 400</th>
</tr>
</thead>
<tbody>
<tr>
<td>bg</td>
<td>0.06</td>
<td>0.08&lt;sup&gt;*&lt;/sup&gt;</td>
<td>0.09&lt;sup&gt;**&lt;/sup&gt;</td>
<td>0.05&lt;sup&gt;*&lt;/sup&gt;</td>
<td>0.02</td>
<td>0.03</td>
<td>0.05&lt;sup&gt;*&lt;/sup&gt;</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.05)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.02)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>ERoT</td>
<td>0.18&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.23&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.26&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.24&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.18&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.21&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.26&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.23&lt;sup&gt;***&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.04)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>bg*ERoT</td>
<td>0.02</td>
<td>-0.07</td>
<td>-0.07</td>
<td>-0.08</td>
<td>-0.02</td>
<td>-0.09</td>
<td>-0.19&lt;sup&gt;***&lt;/sup&gt;</td>
<td>-0.11&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.07)</td>
<td>(0.07)</td>
<td>(0.09)</td>
<td>(0.07)</td>
<td>(0.07)</td>
<td>(0.06)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.26&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.32&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.37&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.38&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.41&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.39&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.39&lt;sup&gt;***&lt;/sup&gt;</td>
<td>0.40&lt;sup&gt;***&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.03)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Observations</td>
<td>91</td>
<td>91</td>
<td>91</td>
<td>91</td>
<td>91</td>
<td>91</td>
<td>91</td>
<td>91</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.173</td>
<td>0.287</td>
<td>0.327</td>
<td>0.241</td>
<td>0.22</td>
<td>0.245</td>
<td>0.316</td>
<td>0.298</td>
</tr>
</tbody>
</table>

<sup>†</sup>Note: Numbers in the parentheses represent the Standard errors. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is trustworthiness. The independent variables are Expected Return on Trust (ERoT), treatment dummy and an interaction.
Table 10: Marginal Effects from Probit Regression of WVS-trust on ERoT with treatment interaction terms

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>bg</td>
<td>-0.06</td>
<td>-0.05</td>
<td>-0.03</td>
<td>-0.03</td>
<td>-0.03</td>
<td>0.29*</td>
<td>0.15</td>
<td>0.27**</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.09)</td>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.17)</td>
<td>(0.15)</td>
<td>(0.13)</td>
</tr>
<tr>
<td>ERoT</td>
<td>0.02</td>
<td>0.10</td>
<td>0.28**</td>
<td>0.31*</td>
<td>0.37**</td>
<td>1.08***</td>
<td>0.72***</td>
<td>0.81***</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.10)</td>
<td>(0.14)</td>
<td>(0.17)</td>
<td>(0.18)</td>
<td>(0.40)</td>
<td>(0.28)</td>
<td>(0.23)</td>
</tr>
<tr>
<td>bg*ERoT</td>
<td>0.00</td>
<td>-0.04</td>
<td>-0.09</td>
<td>-0.04</td>
<td>-0.36</td>
<td>-1.04**</td>
<td>-0.64**</td>
<td>-0.88***</td>
</tr>
<tr>
<td></td>
<td>(0.13)</td>
<td>(0.14)</td>
<td>(0.21)</td>
<td>(0.27)</td>
<td>(0.26)</td>
<td>(0.43)</td>
<td>(0.32)</td>
<td>(0.27)</td>
</tr>
<tr>
<td>Observations</td>
<td>91</td>
<td>91</td>
<td>91</td>
<td>91</td>
<td>91</td>
<td>91</td>
<td>91</td>
<td>91</td>
</tr>
<tr>
<td>Pseudo R-sq</td>
<td>0.00520</td>
<td>0.0134</td>
<td>0.0385</td>
<td>0.0379</td>
<td>0.0310</td>
<td>0.0849</td>
<td>0.0745</td>
<td>0.102</td>
</tr>
</tbody>
</table>

† Note: Numbers in the parentheses represent the Standard errors. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is WVS-trust. The independent variables are Expected Return on Trust (ERoT).
Table 11: Alternative measures of trust and the mutual correlations

(a) Definitions of different measures of trust

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description/Question</th>
<th>Values</th>
<th>Percentage/Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>trust_wvs</td>
<td>Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?</td>
<td>Most people can be trusted</td>
<td>68%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Need to be very careful</td>
<td>32%</td>
</tr>
<tr>
<td>most_adv</td>
<td>Do you think most people would try to take advantage of you if they got the chance, or would they try to be fair?</td>
<td>Most of the time they would try to be fair</td>
<td>63%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Most of the time they would try to take advantage</td>
<td>37%</td>
</tr>
<tr>
<td>wallet</td>
<td>Suppose you left your wallet in the Metro. On a scale of 1 to 10, how much do you think are the chances that you will get it back?</td>
<td>1 to 10 (Mean Reported)</td>
<td>3.04</td>
</tr>
<tr>
<td>trust</td>
<td>Amount sent as a sender in the trust game</td>
<td>0 to 400 in intervals of 50</td>
<td>245</td>
</tr>
</tbody>
</table>

(b) Correlations

<table>
<thead>
<tr>
<th></th>
<th>trust_wvs</th>
<th>most_adv</th>
<th>wallet</th>
<th>send</th>
</tr>
</thead>
<tbody>
<tr>
<td>trust_wvs</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>most_adv</td>
<td>0.21</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>wallet</td>
<td>0.1</td>
<td>0.1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>send</td>
<td>0.07</td>
<td>0.04</td>
<td>0.12</td>
<td>1</td>
</tr>
</tbody>
</table>
Appendix 2: Instructions

1 Instructions for First Experiment (Bribery, Ultimatum and Trust Games)

Welcome.

You are now taking part in an economic experiment. 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. It is prohibited to communicate with the other participants during the experiment. Should 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 identity or the choices that others know. We will give each one of you an identity number which will facilitate your payments.

Overview

The experiment consists of 2 parts and a survey. Please start by reading the instructions for the first part carefully. You will receive the instructions for the second part after the first part is finished. And so on.

Earnings

During the experiment you can earn money by receiving a fictitious currency called “mohor”. All mohors that you earn in the experiment will be exchanged into Rupees at the end of the experiment, The exchange rate is: 1 Mohor = Rs 0.50 (M denotes Mohor henceforth). We will pay you 200 mohors for participating but you can earn additional money depending on the decisions you and the others make. The experiment consists of two parts but remember you will be randomly paid for only one of the two parts which will be determined by a toss. Since you donot know which one you will be paid for, make your decisions for both the parts carefully.

1.1 Part 1 : Subject given either 1.1.1 or 1.1.2

1.1.1 Bribery Game

You may be paired with another participant in this experiment. The matching of two participants will be randomly done. You will not be informed of the identity of the participant with whom you have been matched.

Each participant today will receive a base participation fee of 200 M. We have divided the total number of participants in this experiment session randomly into two equal groups: Citizens (C) and Public Officials (PO).

A citizen performs a task in ten minutes. Her task is to count the number of occurrences of the letter 'A' from a random sequence of letters. She has to perform this task for five different sequences of letters. If she is unable to complete the task in ten minutes then she leaves the game.
with her participation fee of 200 M and the public official gets 600 M. If she successfully does complete the task she is entitled to a payment of 400 M (in addition to participation fee of 200 M). However before the experimenter hands over the entitlement to the citizen, she needs the approval of the Public Official who receives a salary of 400 M for his job of approval (in addition to participation fee of 200 M). A Public Official, however, can ask for a bribe before approving the entitlement for the corresponding citizen. He may ask for a bribe of 100, 200, 300, 400 M. He may choose not to ask for a bribe as well i.e. ask for 0 bribe. The information for demand for bribe then is forwarded to the Citizen who can then decide whether to Accept or Reject the demand for bribe. The final earnings will depend on the choices each one makes i.e. choices about demand for bribe if one is a PO, choices about Acceptance or Rejection if one is a C.

Take a look at the figure below to further clarify the rules of the game and the earnings.

Please go through the two examples given below.

Example. Citizen completes the task. She is entitled to a payment of 600 M. Public official demands a bribe of 100 M for himself. Citizen accepts it. Public Official’s earning is 200+400+100=700 M. Citizen’s earning is 200+400-100=500 M.

Example. Citizen completes the task. She is entitled to a payment of 600 M. Public official demands for a bribe of 400 M for himself. Citizen rejects it. Public Official’s earning is 600+0=600 M. Citizen’s earning is 200+0=200 M.

Instruction for Citizens

In this room all of you are citizens. Note that you are matched anonymously with a participant sitting in the other room. He is your corresponding public official who is in charge of approving your entitlement if you do earn it.

In order to earn the entitlement of 400 M, you have to perform a simple task in 10 minutes following which you will earn the key to your entitlement. What you will see is five sequences of some random letters. Your task is to count the exact number of ‘A’s for each of the sequences. Only when you have correctly counted the number of As for a sequence will you be able to go to the next sequence. When you count successfully for all the sequence you will receive the key for your entitlement and you can use the key to make your choices subsequently.

In order to complete the task, click here to count the number of As for the five sequences of random letters.

1. Input your Identity number. Enter the correct the number of ’A’s for each of the five sequences.

2. If you have completed counting the number of As, please write down the Key to your entitlement on a piece of paper and proceed. If you have not been able to complete the task in ten minutes then you can collect your participation fee and leave.

Right click here and open the link in a new tab to input your choices

3. Please indicate on the response sheet and on the website whether you accept or reject the demand for bribe for all possible bribe amount. We will match your response with the actual bribe...
amount demanded and determine your earnings.

4. We will now collect your responses and match them with the response of your corresponding public official.

5. You will now receive the response sheet, which mentions the bribe demanded, your acceptance/rejection decision and the final earning from this part.

**Instruction for Public officials**

In this room all of you are Public officials. You are responsible for the approving the entitlement for the citizens. Before approving you can ask for a bribe from the citizen. You may also choose not to ask for a bribe.

1. Now make your decision whether to ask for a bribe before you approve the entitlement for the citizen and if yes how much to ask for. Click here to input your choice.

2. We now collect your decision and match them with Citizen’s acceptance/rejection decision.

### 1.1.2 Ultimatum Game

**Common for Participant A and Participant B**

You may be paired with another participant in this experiment. The matching of two participants will be randomly done. You will not be informed of the identity of the participant with whom you have been matched.

Each participant today will receive a base participation fee of 200 M. We have divided the total number of participants in this experiment session randomly into two equal groups: Participant A (P-A) and Participant B (P-B).

Participant A performs a task in ten minutes. Her task is to count the number of occurrences of the letter "A" from a random sequence of letters. She has to perform this task for five different sequences of letters. If she cannot perform the task in ten minutes then she leaves the game with her participation fee 200 M and P-B gets 600 M. If she successfully completes the task, then she is entitled to playing the next stage of the game i.e. she earns the right to second stage. P-B’s earning is 400M in addition to the participation fee of 200 M (i.e. 600 M) for her role, which is the following. In the second stage P-B decides to divide 400 M between himself and P-A. For example if P-B keeps x with himself then P-A may get 400 – x M. The amount which P-A can transfer can only be in multiple of 100s i.e. either 0, 100, 200, 300, 400 in which cases he gets 400, 300, 200, 100, 0 respectively for himself. P-A can then decide whether or not to Accept or Reject the amount which is offered. If P-A accepts the offer then P-B gets (600 + x) M and P-A gets (200 +x) M. If she rejects the earnings then P-A and P-B get only 600 M and 200 M respectively.

Take a look at the figure below to further clarify the rules of the game and the earnings.

Please go through the two examples given below.

Example. P-A completes the task and proceeds to the next round. P-B divides 400 M into 100 M for himself and 300 M for himself. P-A accepts it. P-B’s earning is 200+400+100=700 M. P-A’s
earning is 200+400-100=500 M.

Example. P-A completes the task and proceeds to the next round. P-B divides 400 M into 400 M for himself and 0 for P-A. P-A rejects it. P-B's earning is 600+0=600 M. P-A's earning is 200+0=200 M.

**Instruction for Participant A**

In this room all of you are Participant As. Note that you are matched anonymously with a participant sitting in the other room, he is your corresponding Participant B whom you will play in second round.

In order to proceed to second round, you will have to perform a simple task in 10 minutes following which you will earn the key to the second round. What you will see is five sequences of some random letters. Your task is to count the exact number of 'A’s for each of the sequences. Only when you have correctly counted the number of As for a sequence will you be able to go to the next sequence. When you count successfully for all the sequence you will receive the key and you can use the key to proceed.

1. Right click here and open the link in a new tab. Enter the correct the number of 'A’s for each of the five sequences.

2. If you have completed counting the number of As, please write down the Key to your entitlement on a piece of paper and proceed. If you have not been able to complete the task in ten minutes then you can collect your participation fee and leave.

3. Please right click here and open a new tab and indicate on the response sheet whether you accept or reject the demand for bribe for all possible bribe amount. We will match your response with the actual bribe amount demanded and determine your earnings.

4. We will now collect your responses and match them with the response of your corresponding P-B.

5. You will now receive the response sheet, which mentions the division proposed by P-B, your acceptance/rejection decision and the final earnings from this part.

**Instruction for Participant B**

In this room all of you are Participant B.

1. Now make your decision about how to divide 400 M between yourself and Participant A i.e. whether to share anything with him, if yes how much. Please right click here to open a new link and input your decision. Your sharing amount should be in multiples of 100.

2. We now collect your decision.

3. You are now being informed about whether your decision has been accepted and your and the P-A’s earnings.
1.2 Part 2: Trust Game: Common for Citizen and Public Official in BG and Participant A and Participant B in UG

In this part a participation fee of 200 M to all participants.

You will now play the sending task. In this task, participants are divided into two groups: Senders and Receivers. Both the Senders and the Receivers are given 400 M. First Sender makes a decision. The Sender can choose to send 0, 100, 200, 300 or 400 M to the Receiver. Any amount sent will be tripled. The Sender keeps any amount of money not send to the Receiver.

The Receiver can send back any amount up to the total amount received (that is, the amount the Sender sent multiplied by 3).

Earnings The Sender’s earnings in part 2 are := 400 M - any amount sent to the Receiver + any amount sent back to the Sender

The Receiver’s earnings in part 2 are: = any amount received from the Sender multiplied by three — any amount sent back to the Sender.

You will be asked to make a decision both as a Sender and as a Receiver. One of your roles will be randomly picked. You will be matched with another randomly matched participant in the other role (note that your matched partner here will be different from the matched partner in Part 1). Your decision and the decision of the other player determine your earnings.

Example

You are in the role of a Sender. You have chosen to send 200 M to the Receiver. Hence, the Receiver could send back between 0 and 600 M (= 3 × 200 M) to you. The Receiver has chosen to send back 300 M to you.

• Your earnings in Part 2 are therefore: 400 M - 200 M (the amount you sent) + 300 M (the amount received back) = 500 M.

• The earnings of the Receiver in Part 2 are therefore: 600 M (amount sent to the Receiver) - 300 M (amount sent back by the Receiver) = 300 M

You are in the role of a Receiver. The Sender has chosen to send 400 M to you. Hence, you could send back between 0 and 1200 M ( = 3 × 400 M) to the Sender. You have chosen to send back 100 M to the Sender.

• Your earnings in part 2 are therefore: 1100 M = 1200 M (the amount you received) - 100 M (the amount you sent back).

• The earnings of the Sender in part 2 are therefore : 100 M = 400 - 400 + 100= 100M

Practice exercise.

• You are in the role of a Sender. You have chosen to send 100 M to the Receiver. The Receiver has chosen to send back 0 M to you.

Your earnings in part 4 are: The earnings of the Receiver in part 4 are:

• You are in the role of a Receiver. The Sender has chosen to send 300 M to you. You have chosen to send back 100 M to the Sender.
Your earnings in part 2 are:
The earnings of the Sender in part 2 are:
To make your decisions, right click here and open a new tab.

**Decision as Sender**

1. Mark how much you will like to send the receiver. (in multiples of 50 M)

2. How much you expect the receiver to return you back if you send him 100 M (i.e. Receiver receives 300 M). If your prediction matches the actual decision of the matched receiver then you will win a bonus of 100 M. While making your decision choose between 0,100, 200 or 300 only. (Please mention in multiples of 50 M)

3. How much you expect the receiver to return you back if you send him 200 M (i.e. Receiver receives 600 M). If your prediction matches the actual decision of the matched receiver then you will win a bonus of 100 M. While making your decision choose between 0,100, 200, 300, 400, 500 or 600 only. (Please mention in multiples of 50 M)

4. How much you expect the receiver to return you back if you send him 300 M (i.e. Receiver receives 900 M). If your prediction matches the actual decision of the matched receiver then you will win a bonus of 100 M. While making your decision choose between 0,100, 200, 300, 400, 500, 600, 700, 800 and 900 only. (Please mention in multiples of 50 M)

5. How much you expect the receiver to return you back if you send him 400 M (i.e. Receiver receives 1200 M). If your prediction matches the actual decision of the matched receiver then you will win a bonus of 100 M. While making your decision choose between 0,100, 200, 300, 400, 500, 600, 700, 800, 900 1000, 1100 or 1200 only. (Please mention in multiples of 50 M)

**Decision as Receiver**

1. Mark how much you will like to send back to the sender if you received 150 M (i.e. the sender sent you 50 M). Choose between 0 and 150 but in multiples of 50 only

2. Mark how much you will like to send back to the sender if you received 300 M (i.e. the sender sent you 100 M). Choose between 0 and 300 but in multiples of 50 only

3. Mark how much you will like to send back to the sender if you received 450 M (i.e. the sender sent you 100 M). Choose between 0 and 450 but in multiples of 50 only

4. Mark how much you will like to send back to the sender if you received 600 M (i.e. the sender sent you 200 M). Choose between 0 and 600 but in multiples of 50 only

5. Mark how much you will like to send back to the sender if you received 750 M (i.e. the sender sent you 100 M). Choose between 0 and 750 but in multiples of 50 only

6. Mark how much you will like to send back to the sender if you received 900 M (i.e. the sender sent you 300 M). Choose between 0 and 900 but in multiples of 50 only

7. Mark how much you will like to send back to the sender if you received 1150 M (i.e. the sender sent you 100 M). Choose between 0 and 1150 but in multiples of 50 only

8. Mark how much you will like to send back to the sender if you received 1200 M (i.e. the sender sent you 400 M). Choose between 0 and 1200 but in multiples of 50 only.
Now we will randomly determining your role and determine your earnings. Coin toss for Role determination. Heads := you will be paid for Part 1 and Tails:= you will be paid for Part 2. We will toss once more if you get Part 2 in the first toss in order to determine the roles.

We have come to the end of the experiment. Please fill out the exit survey.

1.3 Exit Survey

1. Identity Number

2. Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people? (Code one answer):

   (a) Most people can be trusted
   (b) Need to be very careful

3. Do you think most people would try to take advantage of you if they got the chance, or would they try to be fair?

   (a) Most of the time they would try to be fair
   (b) Most of the time they would try to take advantage

4. Would you say that most of the time people try to be helpful, or that they are mostly just looking out for themselves?

   (a) Most of the time people are helpful
   (b) Most of the time they are just looking out for themselves

5. Suppose you left your wallet with Rs. 500 in the Metro. On a scale of 1 to 10, how much do you think are the chances that you will get it back?

6. How much you trust people from various groups. Could you tell me for each whether you trust people from this group completely, somewhat, not very much or not at all?

   (a) Your family
   (b) Your neighborhood
   (c) People you know personally
   (d) People you meet for the first time
   (e) People of another religion
   (f) People of another caste
7. How widespread do you think bribe taking and corruption is in this country?
   (a) Almost no public officials are engaged in it
   (b) A few public officials are engaged in it
   (c) Most public officials are engaged in it
   (d) Almost all public officials are engaged in it

8. Please mention for each of the following actions whether you think it can always be justified, never be justified, or something in between. 1: Never justified 10: Justifiable
   (a) Claiming government benefits to which you are not entitled
   (b) Avoiding a fare on public transport
   (c) Cheating on taxes if you have a chance
   (d) Someone accepting a bribe in the course of their duties

9. 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 tick a box on the scale, where the value 0 means: ‘not at all willing to take risks’ and the value 10 means: ‘very willing to take risks’.

10. How do you see yourself: are you a person who is fully prepared to take risks when it comes to car driving/motorcycle riding etc. or do you try to avoid taking risks? Please tick a box on the scale, where the value 0 means: ‘not at all willing to take risks’ and the value 10 means: ‘very willing to take risks’.

11. How do you see yourself: are you a person who is fully prepared to take risks when it comes to financial matters or do you try to avoid taking risks? Please tick a box on the scale, where the value 0 means: ‘not at all willing to take risks’ and the value 10 means: ‘very willing to take risks’.

12. How do you see yourself: are you a person who is fully prepared to take risks when it comes to sports or leisure activities or do you try to avoid taking risks? Please tick a box on the scale, where the value 0 means: ‘not at all willing to take risks’ and the value 10 means: ‘very willing to take risks’.

13. How do you see yourself: are you a person who is fully prepared to take risks when it comes to career or do you try to avoid taking risks? Please tick a box on the scale, where the value 0 means: ‘not at all willing to take risks’ and the value 10 means: ‘very willing to take risks’.
14. How do you see yourself: are you a person who is fully prepared to take risks when it comes to health matters or do you try to avoid taking risks? Please tick a box on the scale, where the value 0 means: ‘not at all willing to take risks’ and the value 10 means: ‘very willing to take risks’.

15. A short survey about yourself

(a) Sex
   i. Male
   ii. Female

(b) Age

(c) Are you the only child of your parents?
   i. Yes
   ii. No

(d) Ancestral State

(e) Mother Tongue

(f) Political Identity

(g) Total Gross Family Income

(h) Caste Affiliation

(i) Religious Affiliation

(j) Religiosity - Not Religious (1) ... Strongly Religious (4)

(k) Your secondary level exam score was (in %)

(l) Your higher secondary level exam score was (in %)

(m) Your CAT/GMAT score (in percentile)

(n) Were the instructions clear? Please write a line or two.

(o) What strategies did you use? Please write a line or two.

(p) How did you decide on the strategies that you chose? Please write a line or two.

(q) What do you think this experiment was about? Please write a line or two.
2 Instructions for second experiment (Social Norms)

2.1 Bribery Game

2.1.1 Introduction

You are now taking part in an economic decision making study.

We will give each one of you an identity number. Please do not lose your identity number. This entire study is anonymous. Please do not discuss with your neighbors at any point during the study.

We will use identity number for payment. Please raise your hands once you have read the questions.

Please write your participant ID in the space provided above.

On the following pages, you will read descriptions of a series of situations. These descriptions correspond to situations in which a person must make a decision. This description will include several possible choices available to, let’s say, Individual A.

After you read the description of the decision, you will be asked to evaluate the different possible choices available to Individual A and to decide, for each of the possible actions, whether taking that action would be “socially appropriate” and “consistent with moral or proper social behavior” or “socially inappropriate” and “inconsistent with moral or proper social behavior.” By socially appropriate, we mean behavior that most people agree is the “correct” or “ethical” thing to do. Another way to think about what we mean is that if Individual A were to select a socially inappropriate choice, then someone else might be angry at Individual A for doing so. Social appropriateness rating is on a scale of -3 to +3 where -3 is “very socially inappropriate” and +3 is very socially appropriate.

In each of your responses, we would like you to answer as truthfully as possible, based on your opinions of what constitutes socially appropriate or socially inappropriate behavior.

To give you an idea of how the experiment will proceed, we will go through an example and show you how you will indicate your responses. On the next page you will see an example of a situation.

Example: Situation

Individual A is at a local coffee shop near campus. While there, Individual A notices that someone has left a wallet at one of the tables. Individual A must decide what to do. Individual A has four possible choices: take the wallet, ask others nearby if the wallet belongs to them, leave the wallet where it is, or give the wallet to the shop manager. Individual A can choose only one of these four options.

The table below presents a list of the possible choices available to Individual A. For each of the choices, please indicate your rating for the social appropriateness of the action on a scale of -3 to +3. Indicate your response, in the table below.
If this were one of the situations for this study, you would consider each of the possible choices above and, for that choice, indicate the extent to which you believe taking that action would be “socially appropriate” and “consistent with moral or proper social behavior” or “socially inappropriate” and “inconsistent with moral or proper social behavior”. Recall that by socially appropriate we mean behavior that most people agree is the “correct” or “ethical” thing to do.

For example, suppose you thought that taking the wallet was very socially inappropriate, asking others nearby if the wallet belongs to them was somewhat socially appropriate, leaving the wallet where it is was somewhat socially inappropriate, and giving the wallet to the shop manager was very socially appropriate. Then you would indicate your responses as follows:

<table>
<thead>
<tr>
<th>Individual A’s choice</th>
<th>Your Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Take the wallet</td>
<td>-3</td>
</tr>
<tr>
<td>Ask others nearby if the wallet belongs to them</td>
<td>+1</td>
</tr>
<tr>
<td>Leave the wallet where it is</td>
<td>-1</td>
</tr>
<tr>
<td>Give the wallet to the shop manager</td>
<td>+3</td>
</tr>
</tbody>
</table>

Are there any questions about this example situation or about how to indicate your responses?

On the following pages, the situations deal with decisions that “Individual A” might have to make. For each situation, you will receive a sheet, with a table on which to indicate your responses.

For each situation, the experimenter will read a description of the situation. You will then indicate whether each possible choice available to Individual A is socially appropriate or socially inappropriate.

At the end of the session today, we will select one of the two situations by a coin toss (Head - Situation I and Tail - Situation II). We will then ask you to randomly choose one category from an envelope containing all the categories. Thus, we will select both a situation and category at random. For each situation and category, we will compute the most frequently occurring response from all the responses in the room today. We will pay you Rs. 110 for your participation today. However if you give the same response as that most frequently given by other people in the room, then you will receive an additional Rs. 110. This amount will be paid to you, in cash, at the conclusion of the experiment.

For instance, if we were to select the example situation above and the possible choice “Leave the wallet where it is,” and if your response had been “somewhat socially inappropriate,” i.e rating -1, then you would receive Rs. 110, in addition to the Rs. 110 participation fee, if this was the response selected by most other people in today’s session. Otherwise you would receive only participation
fee which is Rs. 110.

If you have any questions from this point on, please raise your hand and wait for the experimenter to come to you.

Please wait to turn the page until the experimenter asks you to do so. If you have any questions, please raise your hand and wait for the experimenter.

2.1.2 The Situation

Take a look at Figure 1. Citizens and public officials play a game where they are seated in two separate rooms but each citizen is randomly matched to exactly one public official. A citizen is given 20 problems to solve in 15 minutes. The corresponding matched supervisor grades the answer sheet of the citizen he is matched with. If the citizen solves at least 10 problems correctly he “passes” the test but if she scores less than 10 she “fails” the test. The citizen is entitled to a payment of 400 in addition to a base amount of 200 if she solves at least 10 problems and “passes” but she earns only 200 if she fails. However even if the citizen solves 10 problems or more correctly the supervisor demands a bribe in order to let the citizen pass and earn more. In other words whether to let the citizen pass is entirely his discretion. He may demand a bribe amount of \{10, 20, 30...400\}. He can also choose not to take a bribe. Rate the action of the official and citizen on a scale of -3 to +3 as stated above.

However the citizen may accept to pay a bribe or she may reject it.

Remember you are not being asked report your personal appropriateness rating but social appropriateness rating and you will be paid if your rating matches with the rating of most other people.

Rate the action of Public official on a scale of -3 and +3 in the response sheet given to you.

Rate the action of Public Citizen on a scale of -3 to +3 as stated above.

Remember you will be rewarded if your rating matches with the rating of most other people in the room today.

Figure 1 below gives a visual description of the payoffs.

Figure 1 referred to Figure 2(a) in the paper.

2.1.3 Response Sheet

Identity Number. _________________________________

Situation

Rate Public Official’s decision.
Total amount that can be extracted as bribe is 400.

<table>
<thead>
<tr>
<th>Amount Sought by the public official as bribe</th>
<th>Amount of the entitlement left with the citizen</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 i.e. Public official does not ask for a bribe</td>
<td>400</td>
<td></td>
</tr>
<tr>
<td>10-50</td>
<td>390-350</td>
<td></td>
</tr>
<tr>
<td>60-100</td>
<td>340-300</td>
<td></td>
</tr>
<tr>
<td>110-150</td>
<td>290-250</td>
<td></td>
</tr>
<tr>
<td>160-200</td>
<td>240-200</td>
<td></td>
</tr>
<tr>
<td>210-250</td>
<td>190-150</td>
<td></td>
</tr>
<tr>
<td>260-300</td>
<td>140-100</td>
<td></td>
</tr>
<tr>
<td>310-350</td>
<td>90-50</td>
<td></td>
</tr>
<tr>
<td>360-390</td>
<td>40-10</td>
<td></td>
</tr>
<tr>
<td>400 i.e. Public official demands the entire amount as bribe</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Rate the decision of the Citizen
<table>
<thead>
<tr>
<th>Total surplus that can be extracted as bribe is 400</th>
<th>Citizen’s response</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount Sought by the public official as bribe</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 i.e. Participant B does not keep anything with himself</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>10-50</td>
<td>Accept</td>
<td></td>
</tr>
<tr>
<td>60-100</td>
<td>Accept</td>
<td></td>
</tr>
<tr>
<td>110-150</td>
<td>Accept</td>
<td></td>
</tr>
<tr>
<td>160-200</td>
<td>Accept</td>
<td></td>
</tr>
<tr>
<td>210-250</td>
<td>Accept</td>
<td></td>
</tr>
<tr>
<td>260-300</td>
<td>Accept</td>
<td></td>
</tr>
<tr>
<td>310-350</td>
<td>Accept</td>
<td></td>
</tr>
<tr>
<td>360-390</td>
<td>Accept</td>
<td></td>
</tr>
<tr>
<td>400 i.e. Participant B keeps everything for himself</td>
<td>Accept</td>
<td></td>
</tr>
</tbody>
</table>

### 2.2 Ultimatum Game

#### 2.2.1 Introduction

You are now taking part in an economic decision making study. We will give each one of you an identity number. Please do not lose your identity number. This entire study is anonymous. Please do not discuss with your neighbors at any point during the study.

We will use identity number for payment. Please raise your hands once you have read the questions.

Please write your participant ID in the space provided above.

On the following pages, you will read descriptions of a series of situations. These descriptions correspond to situations in which a person must make a decision. This description will include several possible choices available to, let's say, Individual A.

After you read the description of the decision, you will be asked to evaluate the different possible choices available to Individual A and to decide, for each of the possible actions, whether taking that action would be “socially appropriate” and “consistent with moral or proper social
behavior” or “socially inappropriate” and “inconsistent with moral or proper social behavior.” By socially appropriate, we mean behavior that most people agree is the “correct” or “ethical” thing to do. Another way to think about what we mean is that if Individual A were to select a socially inappropriate choice, then someone else might be angry at Individual A for doing so. Social appropriateness rating is on a scale of -3 to +3 where -3 is “very socially inappropriate” and +3 is very socially appropriate.

In each of your responses, we would like you to answer as truthfully as possible, based on your opinions of what constitutes socially appropriate or socially inappropriate behavior.

To give you an idea of how the experiment will proceed, we will go through an example and show you how you will indicate your responses. On the next page you will see an example of a situation.

**Example: Situation**

Individual A is at a local coffee shop near campus. While there, Individual A notices that someone has left a wallet at one of the tables. Individual A must decide what to do. Individual A has four possible choices: take the wallet, ask others nearby if the wallet belongs to them, leave the wallet where it is, or give the wallet to the shop manager. Individual A can choose only one of these four options.

The table below presents a list of the possible choices available to Individual A. For each of the choices, please indicate your rating for the social appropriateness of the action on a scale of -3 to +3. Indicate your response, in the table below.

<table>
<thead>
<tr>
<th>Individual A’s choice</th>
<th>Your rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Take the wallet</td>
<td></td>
</tr>
<tr>
<td>Ask others nearby if the wallet belongs to them</td>
<td></td>
</tr>
<tr>
<td>Leave the wallet where it is</td>
<td></td>
</tr>
<tr>
<td>Give the wallet to the shop manager</td>
<td></td>
</tr>
</tbody>
</table>

If this were one of the situations for this study, you would consider each of the possible choices above and, for that choice, indicate the extent to which you believe taking that action would be “socially appropriate” and “consistent with moral or proper social behavior” or “socially inappropriate” and “inconsistent with moral or proper social behavior”. Recall that by socially appropriate we mean behavior that most people agree is the “correct” or “ethical” thing to do.

For example, suppose you thought that taking the wallet was **very socially inappropriate**, asking others nearby if the wallet belongs to them was somewhat **socially appropriate**, leaving the wallet where it is was **somewhat socially inappropriate**, and giving the wallet to the shop manager was **very socially appropriate**. Then you would indicate your responses as follows:
<table>
<thead>
<tr>
<th>Individual A’s choice</th>
<th>Your Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Take the wallet</td>
<td>-3</td>
</tr>
<tr>
<td>Ask others nearby if the wallet belongs to them</td>
<td>+1</td>
</tr>
<tr>
<td>Leave the wallet where it is</td>
<td>-1</td>
</tr>
<tr>
<td>Give the wallet to the shop manager</td>
<td>+3</td>
</tr>
</tbody>
</table>

Are there any questions about this example situation or about how to indicate your responses?

On the following pages, the situations deal with decisions that “Individual A” might have to make. For each situation, you will receive a sheet, with a table on which to indicate your responses.

For each situation, the experimenter will read a description of the situation. You will then indicate whether each possible choice available to Individual A is socially appropriate or socially inappropriate.

At the end of the session today, we will select one situation and one category randomly. Thus, we will select both a situation and category at random. For each situation and category, we will have recorded the most frequently occurring response of most people in the room today. If you give the same response as that most frequently given by other people then you will receive a reward of Rs. 110 in addition to the participation fee of Rs. 110. This amount will be paid to you, in cash, at the end of the experiment.

For instance, if we were to select the example situation above and the possible choice “Leave the wallet where it is,” and if your response had been “somewhat socially inappropriate,” i.e rating -1, then you would receive Rs. 110, in addition to the Rs. 110 participation fee, if this was the response selected by most other people in today’s session. Otherwise you would receive only participation fee which is Rs. 110.

If you have any questions from this point on, please raise your hand and wait for the experimenter to come to you. Please wait to turn the page until the experimenter asks you to do so. If you have any questions, please raise your hand and wait for the experimenter.

### 2.2.2 The Situation

Take a look at Figure 2. Participants A and Participants B play a game where they are seated in two separate rooms but each Participant A is randomly matched to exactly one Participant B. Participant A was given 20 problems to solve in 10 minutes. Participant B was supposed to grade the answer sheet of the worker he was matched with. If Participant A solves at least 10 problems correctly then she takes part in a the next part of the game. Otherwise she earns only 110 and leaves. If she proceeds to the next part of the game then she is eligible for a transfer from Participant B i.e. each participant B has 400 between himself and Participant A. Participant B the splits 400 between himself and Participant A. He could share any amount including 0 and 400 i.e. he could share nothing with P-A or he could give away the entire amount to her. P-A in turn could accept or reject the proposed division by P-B. If he rejects the offer then he gets only 110
and P-B gets 500. If he accepts the offer then P-A gets 110+amount P-B shares with him and P-B gets 500+share he keeps with himself. Let us go through the figure to further clarify.

Rate the action of Participant B on a scale of -3 and +3 in the response sheet given to you.

Rate the action of Participant A on a scale of -3 to +3 as stated above.

Remember you will be rewarded if your rating matches with the rating of most other people.

Figure 2 below gives a visual description of the payoffs.

Figure 2 referred to Figure 2(b) in the paper.

### 2.2.3 Response Sheet

Identity Number. _________________________________

Rate Participant B’s Decision.

<table>
<thead>
<tr>
<th>Total amount to be divided is 400</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>The part he proposes to keep with himself.</td>
<td>Amount of the share left with Participant A</td>
</tr>
<tr>
<td>0 i.e. Participant B does not keep anything with himself</td>
<td>400</td>
</tr>
<tr>
<td>10-50</td>
<td>390-350</td>
</tr>
<tr>
<td>60-100</td>
<td>340-300</td>
</tr>
<tr>
<td>110-150</td>
<td>290-250</td>
</tr>
<tr>
<td>160-110</td>
<td>240-110</td>
</tr>
<tr>
<td>210-250</td>
<td>190-150</td>
</tr>
<tr>
<td>260-300</td>
<td>140-100</td>
</tr>
<tr>
<td>310-350</td>
<td>90-50</td>
</tr>
<tr>
<td>360-390</td>
<td>40-10</td>
</tr>
<tr>
<td>400 i.e. Participant B keeps everything for himself</td>
<td>0</td>
</tr>
</tbody>
</table>

Rate the action of Participant A
<table>
<thead>
<tr>
<th>Total amount to be divided is 400</th>
<th>Participant A’s response</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>The part participant B proposes to keep with himself.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 i.e. Participant B does not keep anything with himself</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>10-50</td>
<td>Accept</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reject</td>
<td></td>
</tr>
<tr>
<td>60-100</td>
<td>Accept</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reject</td>
<td></td>
</tr>
<tr>
<td>110-150</td>
<td>Accept</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reject</td>
<td></td>
</tr>
<tr>
<td>160-110</td>
<td>Accept</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reject</td>
<td></td>
</tr>
<tr>
<td>210-250</td>
<td>Accept</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reject</td>
<td></td>
</tr>
<tr>
<td>260-300</td>
<td>Accept</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reject</td>
<td></td>
</tr>
<tr>
<td>310-350</td>
<td>Accept</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reject</td>
<td></td>
</tr>
<tr>
<td>360-390</td>
<td>Accept</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reject</td>
<td></td>
</tr>
<tr>
<td>400 i.e. Participant B keeps everything for himself</td>
<td>Accept</td>
<td></td>
</tr>
</tbody>
</table>
Chapter Four
Does bureaucracy attract more corrupt people? An experimental evidence from India

Ritwik Banerjee∗1, Tushi Baul2, and Tanya Rosenblat2

1Department of Economics and Business, Aarhus University
2Department of Economics, Iowa State University

Abstract

Do corrupt people self select themselves in professions where the scope of corruption is high? We conduct a corruption experiment with private sector job aspirants and students aspiring to join the highest echelons of Indian bureaucracy. The game we implement is novel and models embezzlement of resources in which “supervisors” evaluate the performance of “workers” and then pay them. We find that aspirant bureaucrats indulge in corruption more than private sector aspirants. The public sector aspirants over report the amount required to pay the workers more than their private sector counterparts. However, we do not find any statistically significant difference in the likelihood of being corrupt, between the two sectors. We also find no evidence that pro-sociality or competitiveness are significant sorting factors for the aspirants of either sector. The findings suggest that, contrary to the widely held view that “system corrupts people”, bureaucracy in India attracts more corrupt people, possibly due to the corrupt rent seeking potential in it.

JEL Classification: C91 D73 O12 K42
Keywords: Corruption, Experiments, Bureaucracy, Public Sector

∗rbanerjee@econ.au.dk; Address: Department of Economics and Business, Fuglesangs Allé 4 8210 Århus V, Denmark. We thank Alexander Koch, Torben M Andersen, Lata Gangadharan, Klaus Abbink and other seminar participants at ESA Cologne 2012 and Aarhus University for their comments and suggestions. We also thank Sujoy Chakraborty and Rajat Kathuria without whose help we could not have conducted the experiment. Banerjee thanks Ashokankur Datta, Sayan Dasgupta and Anindya Sengupta for hosting him in Delhi during the period ofthe study.
1 Introduction

“Just as it is not possible, not to taste honey or poison put on the surface of the tongue, so it is not possible for the Government servants dealing with money not to taste it in however small quantities.”

- Kautilya in Arthashastra (circa 300BC).

This insightful surmise, from the ancient Indian political economy treatise - Arthashastra, authored more than 2300 years ago by Kautilya, finds more resonance in much of the developing world today than ever before. However, in the developing countries with a large, inefficient and corrupt government, the more important question is whether it is the surface of the tongue which is on the lookout for honey. In other words, are more corrupt people self selecting into professions where the possibilities of corruption are high?

The narrative which governs public sector corruption is the following - systemic corruption is embedded in the public office and thus even honest people turn corrupt once they become a part of it. There are several reasons why that may be the case. An honest person may join the government and recalibrate his ex ante prior of monitoring and inspection downwards. It may as well be that rampant corruption among peers may lower his moral cost (Innes and Mitra, 2013; Bardhan, 1997). However if corruption is an equilibrium strategy for most people in a sector for a sufficiently long time, then that sector may begin to attract corrupt people.

The rich literature on labor market sorting which has evolved in the recent past indicates that the dimensions of sorting are indeed numerous. Abowd et al. (1999) show that more productive firms employ more productive employees while Krueger and Schkade (2007) conclude gregarious workers flock towards jobs with higher social interaction. Women with the same abilities as men have been found to self select themselves out of competitive environments (Niederle and Vesterlund, 2007). One of the few papers that looks at the interaction between ethicality, changes in the economic incentives and sorting is Ferraz and Finan (2011). They find that greater public monitoring of federal funds in Brazil attracts better politicians. However, corruption as a potential dimension of sorting has not been examined.

Two studies are closely related to ours. Alatas et al. (2009) experimentally examine the corruption propensities of Indonesian bureaucrats and students and find that the former are less likely to engage in corruption than the latter. The lack of selection effect in their data is attributed to the experience effect, i.e. the effect coming from different real life experiences among the subject pools. However, the selection effect is difficult to identify in their setting for several reasons. The bureaucrats are typically observed only after they have joined the public service. Thus, at the time of the experiment, it is very plausible that they have internalized the norms of the profession they are already in. Furthermore, both Alatas et al. (2009) and a more recent study by Hanna and Wang
rely on one survey based question about whether the subject wants to join the public or the private sector, and this may not be a robust reflection of true occupational preference. Hanna and Wang (2013) also study cheating behavior among students as opposed to corruption. While the two are conceptually related, the moral cost of indulging in cheating is much less than that of making corrupt decisions and therefore behavior in the two may not be identical.

In our study, we overcome some of these limitations by measuring and contrasting corruption propensities between students who aspire to become bureaucrats at the highest level in the government and those who aspire to join the private corporate sector. We implement a uniquely designed corruption game which mimics embezzlement: “supervisors” decide how much to overreport, if any, the amount they need to pay “workers” and how much to underpay them, if at all, for their task. We find that corruption is fairly high among both groups - when averaged over the two groups, 66% subjects opt for corrupt strategies. The results further show that the public sector aspirants are systematically more corrupt than their private sector counterparts. However, we find no statistical difference in the percentage of those who engage in corruption in the two sectors. Thus, our findings suggest that while corruption is a dimension of sorting into the public sector labor market, private sector aspirants are equally likely to be corrupt.

Our contribution to the literature are two fold. One, we clearly identify selection effect vis-a-vis corruption with aspirant bureaucrats as opposed to in-service bureaucrats - our subject pool in that sense is unique. This allows us to compare the corruption propensities of aspirant bureaucrats with those of aspirants for a job in the private corporate sector. Two, contrary to existing corruption games in the literature which model petty corruption (see Serra and Wantchekon (2012) for a review), we introduce a new corruption game which attempts to model embezzlement rather than bribery.

The rest of the paper is divided into the following sections. Section 2 describes the experiment with the identification strategy in Section 2.1, the design of the experiment in Section 2.2 and the experimental procedure in Section 2.3. The results are given in Section 3 while Section 4 offers the concluding remarks.

2 The Experiment

Shleifer and Vishny (1993) make a crucial distinction between two types of corruption - petty corruption, which does not entail theft of government resources, and embezzlement, which does. It is not difficult to imagine why an embezzlement of public resources causes more loss to welfare than...
petty corruption. Typically, the beneficiaries of these public entitlements are poorer sections of the population and any siphoning off of the public resource causes a far higher welfare depletion. Also, there has been some justification of petty bribery as “speed money” and thus efficiency improving, however no such rationale exists for embezzlement and the scale of misappropriation under the latter is huge.

In India, embezzlement of public resources has taken gargantuan proportions. MNREGA, a flagship program in India, which guarantees 100 days of compulsory work a year to 900 million people at the cost of 3.6% of annual government spending, has been found to have a massive leakage of funds, often aided by fictitious spells of workers and “ghost households” (Niehaus and Sukhtankar, 2013). Similarly, several studies on food, health and education suggest enormous embezzlement of government resources.

Considering the above, the literature on experimental corruption games has largely focused on petty bribery. A thorough understanding of corruption thus entails understanding and observing the incentive structure faced by the most important cog in the wheel connecting the government and the citizens - the bureaucrat. Hence, our corruption game mimics a situation where resources from the government are claimed by a bureaucrat in order to deliver entitlements to the citizenry. But first we discuss the strategy for identifying self-selection.

2.1 Identification Strategy

We compare the degree and extent of corruption across subject pools - one in which students aspire to join the bureaucracy in the government and another, where they aspire to join the private corporate sector. We are aware of the pitfalls of subject pool comparison as highlighted by Levitt and List (2007), but it remains a widely used instrument through which subject pool effects have been studied: e.g. in the context of financial market trading (Alevy et al., 2007), sports-card dealers (List, 2003), nurses (Barr et al., 2004), CEOs (Fehr and List, 2004) and even public affairs officials (Potters and Winden, 1996). Subject pool comparisons have been used in the experimental corruption literature as well: Cameron et al. (2009) study corruption behavior across cultures and

---

3Some authors, beginning with Leff (1964) and Huntington (1968) and later Lui (1985) and others, argue that in developing countries where institutions are inefficient, petty corruption can be efficiency improving and thus growth enhancing. However this view has been rejected several times subsequently (see Shleifer and Vishny (1993) and Mauro (1995)).

4Valsecchi (2013) finds that of all the corruption related cases prosecuted or coordinated by the General Attorney Office in Indonesia, 80% were those involving embezzlement.

5Niehaus and Sukhtankar (2012) find that none of the statutory wage increase is passed on to workers in an MNREGA program implying a marginal leakage of 100%.

6According to a Planning Commission report, 58% of subsidized targeted food grains through Public Distribution System in India are diverted. Olken (2006) estimates that at least 18% of the subsidized rice is diverted from the Operasi Pasar Khusus (OPK) program in Indonesia.

7Reinikka and Svensson (2004) notes that 87% of school grants, meant to cover non-wage expenditure of primary schools in Uganda, were usurped by local officials in connivance with politicians. Interestingly, the official figure of Uganda’s total public expenditure in the mid 1990s on education was 20%.
Alatas et al. (2009) compare corruptibility between student subjects and working bureaucrats. The difference between subject pool comparison in past corruption studies and ours is that we measure corruptibility in a non-standard subject pool, i.e. from representatives of the population of interest. Notice that in order to identify selection, our interest lies not in public sector bureaucrats or private sector managers but in aspirant bureaucrats and aspirant managers. Our subjects come precisely from these two subject pools. In that sense, our experiment may be categorized as an artefactual field experiment according to the taxonomy developed by Harrison and List (2004). What follows next is a description of the subject pools.

2.1.1 Public Sector Aspirants

In order to join the Indian government administrative service at the highest level, aspiring candidates have to take a three level national examination called the Union Public Service Commission (UPSC) Examination. Each year roughly 400,000-500,000 students appear for the Preliminary level of the UPSC exam. Around 2% of them are invited for the second level - the main exam. Finally 0.4% of the students manage to appear for the interview out of which 700-800 students are finally selected - this translates into a success rate of 0.25%-0.3%. Admittedly a difficult exam, UPSC however allows students to appear for the exam several times.

Once a candidate clears this exam, she can choose from a number of services like Indian Foreign Service (IFS), Indian Administrative Service (IAS), Indian Police Service (IPS), Indian Postal Service, Indian Revenue Service (IRS) etc. with a starting salary of approximately Rs. 720,000 ($28,800 in PPP terms). In other words, people qualifying through the UPSC exam comprise of the entire spectrum of personnel on whom the government machinery rests. In particular, Indian Administrative Service, which admits within its folds only 80-100 students every year, is the “executive bulwark” through which constitutional propriety is enforced, policies implemented and entitlements delivered. However, it is this service which has earned the flak for being inefficient and thoroughly corrupt in the recent years.

Due to several reasons, over the years Delhi in general and Jawaharlal Nehru University (JNU) in particular, has become an important hub for the aspirants of the UPSC exam. JNU is a premiere research university in India and it attracts students from all across India and offers a wide range of choice of majors at bachelors and graduate levels, mainly in humanities and social sciences. In

---

Note that the expected salary faced by a public sector aspirant is much lower than that of a private sector aspirant since the probability of passing the UPSC exam is much lower than getting a job in the private sector.

It has been hailed as the “Steel frame of India” by Vallabhai Patel - a statesman and the first home minister of India.

Indian bureaucracy has been ranked as the worst in Asia by the Hong Kong based Political and Economic Risk Consultancy (PERC). Debroy and Bhandari (2013) estimate that $18.42 billion are pocketed by public officials through corruption. In a separate study Transparency International estimates that truckers alone pay police officials Rs. 22,200 annually in order to cross state borders. Bertrand et al. (2008) find that the average license getter in Delhi pays Rs. 1080 while the official license fee is Rs. 450. They further note that the differential is not merely transfers, usurped by “the insidious nexus of bureaucrats and agents” but it is distortionary as well.
our study, the subject pool consisting of public sector aspirants are made up by students who are preparing for the UPSC exam based in JNU.

2.1.2 Private Sector Aspirants

The comparison group in our study comprises of private sector aspirants i.e. students who aspire to work for the private corporate sector in India. For this, we have recruited subjects at International Management Institute (IMI) - a premier management training institute, located adjacent to JNU in Delhi, which offers an MBA degree in Marketing, Finance and Human Resource. Students at IMI train themselves with an aim to join India Inc. Over the years IMI has been fairly successful in placing its students in the private corporate sector within India as well as abroad. In 2011, out of around 150 students at IMI, 27% have been placed within the financial sector, 21% with consulting firms, 11% with the FMCG sector and 5% with hospitality, Information Technology and KPO firms each. The average annual salary of the placed students is Rs. 900,000 (≈ $36000 in PPP terms). Subjects in both these groups are in several ways representative of the public and private sector aspirants - our population of interest. While subjects in the former group are randomly selected from the student pool who are going to write the UPSC exam and may as well be future bureaucrats occupying high offices, those in the latter have joined IMI after having qualified in the Common Admission Test (a competitive test similar to GMAT which MBA students in the country are required to take). The placement opportunities available to the IMI graduates discussed above further reinforce the notion that these students will go on to occupy important positions in India Inc. It is also important to note that other than unethicality, occupational sorting between the public and private sector may also be determined by other characteristics such as gender, caste, prosociality etc - however, we do not test for them in this paper.

2.2 Design

In our experimental design, we model a situation where a public official evaluates the performance of a worker and is mandated to pay her accordingly. To pay the worker, the public official receives resources from the government. The situation is an abstraction of many programs and instances such as those discussed above, where entitlements are disbursed to the citizens. We use loaded terms with words like “Overreport” and “Underpay” in our design of the corruption game. Studies in the past have shown that the use of loaded terms does impose a sense of “immorality” on the subjects (Banerjee, 2014) and the results obtained are externally valid (Armantier and Boly, 2013). Subjects, participating in our pen-and-paper, real effort task experiment are divided into two groups - “workers” and “supervisors”. They are seated in two separate rooms and each worker is anonymously matched to a supervisor. Workers are asked to solve twenty matrix problems, such
as in Ariely et al. (2009) in ten minutes. Their answers are graded by a randomly matched supervisor located in an adjacent room. The workers are entitled to a payment of 1 experimental token for each correctly solved matrix, as reported by the supervisor.

A supervisor, after having graded the answersheet, goes to a token counter located outside the room and claims the total number of tokens that he needs to pay the worker and pay himself his salary of 2 tokens. After collecting the tokens, the supervisor comes back to his seat where he finds an envelope with the workers identity marked on it. He puts the worker’s earnings into the envelope and seals it. He keeps the rest of the tokens with himself and leaves the worker’s envelope in a box. These envelopes are matched and sent to the corresponding workers. The supervisor then makes one of the following choices:

1. Overreport at the token counter i.e. claim more tokens than you need and Underpay the worker’s performance i.e. put less tokens in the worker’s envelope than he deserves.

2. Overreport at the token counter i.e. claim more tokens than you need and be Truthful with the worker’s performance i.e. put the exact number of tokens that the worker deserves.

3. Be Truthful at the token counter i.e. claim exact number of tokens you need and Underpay the worker’s performance i.e. put less tokens in the worker’s envelope than he deserves.

4. Overreport at the token counter i.e. claim more tokens than you need and Overpay the worker’s performance i.e. put more tokens in the worker’s envelope than he deserves.

5. Be Truthful at the token counter i.e. claim exact number of tokens you need and be Truthful with the worker’s performance i.e. put the exact number of tokens that the worker deserves.

Note that one of the unique features of our corruption game is that there are two margins of corruption - overreporting and underpaying. Given this setting, the supervisor has a number of strategy combinations. Our interest mainly lies in the following subset - overreport and underpay, overreport and pay honestly, report honestly and underpay and report honestly and pay honestly. There are two less probable albeit interesting cases that may arise - one, a supervisor may overreport and pay the worker more than she deserves and two, he may underreport and pay her less. These strategies, despite the involvement of overreporting and underpayment may not lead to corrupt earnings. The three other strategy combinations entail incurring a deliberate self inflicted loss.

After the corruption game, supervisors participate in two diagnostic games and are randomly paid for one of them. In the first game, each of them is allocated 10 tokens and is asked to pay any number of tokens to one of five listed charity organizations (Camerer, 2003). This aims to measure

\footnote{In each matrix they are asked to find two numbers which add up to ten.}

\footnote{Giving subjects an explicit option to embezzle, we believe, does not create experimenter demand effect - not only does it show that making decisions under such explicit framing does entail a moral cost, there is also no reason to assume that this loaded frame affects the two subject pools asymmetrically.}
prosocial behavior among subjects. In the second game, they are asked to choose a wage scheme, by which they will like themselves to be paid, for a two minute long addition task. They can choose the wage scheme from one of the following - fixed wage, variable wage and tournament wage. The aim here is to capture the preference for competition among subjects and examine if occupational sorting can be explained in part by attitude to competition\textsuperscript{13}. For the exact descriptions, see the instructions in Appendix 2.

\subsection*{2.3 Experiment Procedure}

The experiment was conducted at Jawaharlal Nehru University (JNU) and International Management Institute (IMI) in New Delhi. The public sector group consisted of those students from JNU who were either going to appear for the UPSC exam in 2013 or those who had already appeared in it in 2012 or both. The recruitment of public sector aspirants was very carefully conducted as we wanted to include only those who were genuine UPSC aspirants.

We finally had 82 subjects participating in our experiment from JNU, out of which 41 subjects played the role of supervisors - they were all aspirant bureaucrats. From IMI we had 66 participants with 33 in the role of supervisors - they were private sector aspirants. Thus, we had 41 and 33 matched pairs in public and private sector aspirants, respectively. Three sessions were conducted at JNU, each with at least 24 students. Only one session was conducted at IMI and it had 66 students. The students were asked to come to two different rooms as per their preassigned roles, and they were randomly matched. We were particularly careful to maintain the anonymity of subjects in the experiment because we expected dishonest behavior. Thus, the responses of all subjects were anonymous and were tracked using a random identity number. All earnings were exchanged to Rupee at the rate 1 token=Rs. 50.

\section*{3 Results}

Figure 1 plots the data in an overreport-underpayment panel where the size of the circle becomes larger with corrupt earning. For instance, a circle at (5,10) denotes that the subject underpaid the worker by 5 tokens and overreported at the token counter by 10 tokens leading to a total corrupt earning of 15 tokens. All the elements of the strategy space discussed in Section 2.2 can be seen here. Also, the percentage of subjects in each of the elements are given in Table 1. The solid dots show the points where corrupt earnings are zero. The correlation between number of tokens overreported and underpaid is weakly positive. Interestingly, the solid dots are also present in areas other than at the (0,0) coordinate - we call this phenomenon “benevolence”. The beneficent

\textsuperscript{13}We are aware that the measure may be confounded by attitude to risk. But such a composite measure serves our purpose since the aim here is to obtain a measure which controls for attitude to competition and risk. The task itself is simple and does not depend on heterogeneity in ability.
Figure 1: Overreporting and Underpayment

Note: Red (Green) represents public (private) sector aspirants. The figure is a scatter plot of overreporting and underpayment, weighted by total corrupt earnings. For instance, a circle at (5,10) means that the supervisor underpaid 5 tokens, overreported 10 tokens and thereby has corrupt earnings of 15 tokens.

supers are those who overreport at the token counter and overpay the workers at the same time or *vice versa*\(^{14}\). One of the interesting findings from our data is that while benevolence is not associated with religiosity, it is negatively related to workers’ true performance (Table 3). In other words, supervisors were more likely to overreport and overpay workers if the performance of the workers were low\(^{15}\). We also find a significant positive relation between corrupt earning with the respective self declared degree of religiosity.

Figure 2 in the Appendix 1 shows the distribution of the overreporting, underpayment and corrupt earning for both the private and public sector aspirants. Table 2 presents the full sample and corrupt sample mean differences in corrupt earnings, the *t*-tests of differences in mean and Mann-Whitney ranksum tests of difference in distributions. Average number of tokens overreported is significantly higher among public sector aspirants, at 4.02, than their private sector counterparts, at 0.67 (*t*-test, \(p\)-value<0.001). The average underpayment by private sector aspirants, at 2.61, is less than that of the public sector aspirants which is 3.54, but the difference is not statistically significant (*t*-test, \(p\)-value=0.46). The mean corrupt earning, which is the sum of overreporting and underpayment, is significantly higher for the public sector group than for their private counterpart.

\(^{14}\)There also exists one supervisor from the public sector aspirant group who underreports at the token counter and underpays the worker.

\(^{15}\)We find weak correlation between charitable giving and benevolence. Furthermore, corruptibility is found to have an insignificant but negative relation with charitable giving, with \(\rho = -0.13\) for private sector and \(\rho = -0.08\) for public sector.
The mean difference in overreporting and corrupt earnings (not underpayment) holds even for the sample which chose to be corrupt.

The demographic characteristics of the private and public aspirants are compared in Table 5 in the Appendix 1. Since a student typically appears in the UPSC exam several times before being able to pass it, the public sector aspirants are on an average slightly older than the private sector aspirants (24.5 as compared to 22.8). There was also a higher proportional representation of lower caste in the public sector with only one student in the private sector group belonging to the lower caste, a higher percentage of female subjects and higher family income among the private sector aspirants. Due to the unbalance in the subject characteristics, it is important to examine whether the sectoral choice can be explained by these differences. We estimate a probit model with the choice of professional sector as the dependent variable and corrupt earnings along with demographic and other exogenous variables as independent variables. The vector of independent variables were chosen so as to control for all possible factors that may affect occupational choice. The presence of public servants in the family may inspire children to follow suit, higher family income may nudge students to acquire relatively expensive skill sets which are more valuable in the private sector, students from a lower caste may be attracted to government to benefit from the institutionalized affirmative action policies, a perception of higher corruption in the society at large might affect his choice and finally attitude to competition may determine sector choice. We also use data from the diagnostic treatments - charity and choice of competitive scheme as a proxy for prosociality and attitude to competition, respectively.

The marginal effects are reported in col (1) - (2) of Table 4. The results indicate that there is an effect of corrupt earning on choice of sector after controlling for demographic and other individual variables. Corrupt earning has a positive effect on probability of choosing public sector and is highly significant in both specifications. The size of the estimate suggests that a 10% increase in corrupt earning leads to a 0.2 percentage point increase in probability of being an aspirant public servant. Presence of public servants among immediate family members, lower caste identity, motivation for public service significantly increase the probability of choosing the public sector while family income significantly decreases it. However, unlike in Hanna and Wang (2013), we find no effect of choice of incentive scheme and charity contribution on choice of sector. Col (3) - (4) report the ordinary least square estimates of demographic and preference variables on corrupt earnings - none of the variables predict corrupt earnings.

Table 2 reports the proportion of subjects, across the two subject pools, who indulged in corruption. We find that 66% of the public sector overreported as against only 33% of private sector aspirants ($\chi^2$-test, $p$-value=0.005). Also, 59% of the former underpayed as against only 39% of the latter ($\chi^2$-test, $p=0.10$). However, when it comes to corruption, though 70% of former are found to be more corrupt as against 63% of the latter, the difference is not statistically significant ($\chi^2$-test, $p$-value=0.516). This suggests that while most of the public sector corrupt indulged in corruption
by both overreporting and underpaying, the corrupt private sector aspirants who underpaid were not the ones who overreported and vice-versa. These results when juxtaposed with that of the bribe amount, implies that while public sector aspirants are not more likely to be corrupt, but on an average they are more corrupt.

4 Conclusion

We introduce a corruption game which is distinctly different from the bribery and petty corruption games that are used in the corruption experimental literature. The game, which models the most significant source of leakage of public resources - embezzlement, was used to examine whether selection into the public sector occurs with corruption as a sorting dimension. The experiment was conducted in India with two treatment groups - aspirants of Indian bureaucracy and aspirants of the private corporate sector. We found that both public sector and private sector aspirants were equally likely to be corrupt, however the public sector aspirants on an average are more corrupt than the private sector counterparts. Our results show that just like tall people self select themselves into professions where they could put their height to an advantage (e.g. basketball), people with lower moral costs (or more skillful in corruption) may aspire to join professions where there exists a possibility of corruption. However, the deeper psychological mechanisms driving the selection are still not well understood with the data that we have. This could be an excellent topic for a future follow up study.

From a policy standpoint, our results are important as they suggest that the occupational choice of a person is correlated with his/her intrinsic characteristics, such as the moral cost of indulging in unethical acts, besides standard socioeconomic factors. This makes corruption even harder to fight in an environment which is already widely corruption ridden. Not only is it important that existing corruption is eliminated through systematic interventions in institutions, it is also important that a strong and credible signal is sent to aspirant bureaucrats that corrupt activities will be effectively dealt with.

References


Armantier, Olivier and Amadou Boly, “Comparing Corruption in the Laboratory and in the Field in Burkina Faso and in Canada,” *The Economic Journal*, 2013, pp. n/a–n/a.


Leff, N. H., American Behavioral Scientist, Nov 1964, 8, 8–14.


Table 1: Percentage of subjects in each category

<table>
<thead>
<tr>
<th>Category</th>
<th>Percentage of subjects (Public, Private)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Over-report</td>
<td>7.3, 9.1</td>
</tr>
<tr>
<td>Over-report</td>
<td>12.2, 18.2</td>
</tr>
<tr>
<td>Over-report</td>
<td>46.3, 6.1</td>
</tr>
<tr>
<td>Be truthful</td>
<td>22.0, 33.3</td>
</tr>
<tr>
<td>Be truthful</td>
<td>4.9, 30.3</td>
</tr>
<tr>
<td>Under-report</td>
<td>7.3, 3.0</td>
</tr>
</tbody>
</table>

Note: Percentage of subjects in each category from the public and private sector aspirants. Though (under-report, underpay) option was not explicitly mentioned, it was exercised as an option.

Table 2: Mean difference corruption

<table>
<thead>
<tr>
<th>Category</th>
<th>Private Aspirants</th>
<th>Public Aspirants</th>
<th>Mean difference&lt;sup&gt;a&lt;/sup&gt;</th>
<th>t/χ²-test, p-value</th>
<th>MW Test&lt;sup&gt;b&lt;/sup&gt;, p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overreporting</td>
<td>0.67</td>
<td>4.02</td>
<td>-3.36**</td>
<td>0.000</td>
<td>0.001</td>
</tr>
<tr>
<td>Underpayment</td>
<td>2.61</td>
<td>3.54</td>
<td>-0.93</td>
<td>0.458</td>
<td>0.193</td>
</tr>
<tr>
<td>Corrupt Earning</td>
<td>3.27</td>
<td>7.56</td>
<td>-4.29**</td>
<td>0.006</td>
<td>0.031</td>
</tr>
<tr>
<td>1(Overreporting&gt;0)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.33</td>
<td>0.66</td>
<td>-0.32**</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td>1(Underpayment&gt;0)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.39</td>
<td>0.59</td>
<td>-0.19*</td>
<td>0.100</td>
<td></td>
</tr>
<tr>
<td>1(Corrupt Earning&gt;0)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.63</td>
<td>0.70</td>
<td>-0.07</td>
<td>0.516</td>
<td></td>
</tr>
</tbody>
</table>

Note: For Overreporting, Underpayment, and Corrupt Earning, the mean difference is reported as Mean difference<sup>a</sup>.

<sup>a</sup>*** p<0.01, ** p<0.05, * p<0.1.

<sup>b</sup>Mann Whitney Ranksum test

<sup>c</sup>1(>0) is an identity variable and denotes all subjects who made a non-zero decision. A χ² test reports p-values corresponding to the mean difference.
Table 3: Correlation between Corrupt earning, Benevolence and worker entitlement, Religiosity

<table>
<thead>
<tr>
<th></th>
<th>Corrupt earning</th>
<th>Benevolence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Worker’s True Performance</td>
<td>-0.17</td>
<td>-0.20*</td>
</tr>
<tr>
<td>Religiosity</td>
<td>0.19*</td>
<td>-0.07</td>
</tr>
</tbody>
</table>

Note: Non-parametric correlation coefficients are reported. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Regression Results in terms of Marginal Effects

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public</td>
<td>0.02***</td>
<td>0.02***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0.05</td>
<td>0.05</td>
<td>-2.71</td>
<td>-2.72</td>
</tr>
<tr>
<td>Age</td>
<td>0.05***</td>
<td>0.05***</td>
<td>-0.31</td>
<td>-0.30</td>
</tr>
<tr>
<td>Public Servants in Family</td>
<td>0.20***</td>
<td>0.19**</td>
<td>-2.05</td>
<td>-2.01</td>
</tr>
<tr>
<td>Log Family Income</td>
<td>-0.10***</td>
<td>-0.10***</td>
<td>-0.69</td>
<td>-0.67</td>
</tr>
<tr>
<td>Caste</td>
<td>0.21***</td>
<td>0.21***</td>
<td>1.15</td>
<td>1.10</td>
</tr>
<tr>
<td>Worker’s True Performance</td>
<td>-0.01***</td>
<td>-0.02**</td>
<td>-0.20</td>
<td>-0.18</td>
</tr>
<tr>
<td>Charity</td>
<td>0.01</td>
<td>0.01</td>
<td>0.15</td>
<td>0.16</td>
</tr>
<tr>
<td>Incentive Scheme</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.39</td>
<td>0.29</td>
</tr>
<tr>
<td>Constant</td>
<td>25.12**</td>
<td>25.97**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>74</td>
<td>74</td>
<td>74</td>
<td>74</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.126</td>
<td>0.129</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pseudo R-sq</td>
<td>0.620</td>
<td>0.623</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log Likelihood</td>
<td>-19.32</td>
<td>-19.18</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

† Note: Numbers in the parentheses represent the Standard errors. *** p<0.01, ** p<0.05, * p<0.1. The dependent variable in Col (1) - (2) is a dummy variable indicating public or private sector aspirants and the corresponding marginal effects are reported from probit regression. The dependent variable in Col (3) - (4) are corrupt earnings and the marginal effects from OLS are reported. For details of the variable definitions please refer to Table 5.
Figure 2: Corrupt Earnings by private and public sector aspirants
<table>
<thead>
<tr>
<th>Definition</th>
<th>Private</th>
<th>Public</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male =1 if gender is male</td>
<td>0.52</td>
<td>0.73</td>
<td>-0.21</td>
</tr>
<tr>
<td>Age</td>
<td>22.8</td>
<td>24.5</td>
<td>1.7</td>
</tr>
<tr>
<td>Caste&lt;sup&gt;a&lt;/sup&gt; Caste Category</td>
<td>0.03</td>
<td>0.76</td>
<td>-0.73</td>
</tr>
<tr>
<td>Public Servants in Family</td>
<td>0.36</td>
<td>0.51</td>
<td>-0.14</td>
</tr>
<tr>
<td>Log Family Income</td>
<td>11.27</td>
<td>10.26</td>
<td>1.01</td>
</tr>
<tr>
<td>Worker’s True Performance</td>
<td>16</td>
<td>12</td>
<td>-4</td>
</tr>
<tr>
<td>Charity</td>
<td>4.06</td>
<td>3.36</td>
<td>-0.70</td>
</tr>
<tr>
<td>Incentive Scheme&lt;sup&gt;b&lt;/sup&gt; Choice of wage scheme for the addition task</td>
<td>2.03</td>
<td>1.95</td>
<td>-0.08</td>
</tr>
<tr>
<td>Number of subjects&lt;sup&gt;c&lt;/sup&gt;</td>
<td>33</td>
<td>41</td>
<td>-8</td>
</tr>
</tbody>
</table>

<sup>a</sup>=0 for upper castes, =1 for other backward castes, =2 for Scheduled Castes and =3 for Scheduled Tribes.

<sup>b</sup>=1 for fixed wage, =2 for piece rate wage, =3 for tournament wage.

<sup>c</sup>8 students had appeared for UPSC exam once, 8 twice and 4 twice or more.
Appendix 2: Instructions

Instructions for Supervisors

Introduction

You are now taking part in an economic decision making study. You will get a fixed participation fee of Rs. 300 but you can earn additional money depending on the decisions that you make. During the experiment you can earn money by receiving tokens. All tokens that you can earn in the experiment will be exchanged into Rupees at the end of the experiment. The exchange rate is:

\[ 1 \text{ Token} = \text{Rs. 50} \]

At the end of the experiment you will receive the amount of money that you have earned during the experiment in cash.

The experiment consists of several parts and a survey. The instructions will explain in detail what the respective part of the experiment is about. Please follow the instructions carefully. If you have any questions please let us know by raising your hand. Your question will be answered by us in private.

In this experiment you will have to solve few math problems. Use of helping devices like calculators, cell phones is strictly prohibited. If you use helping device, you will be immediately excluded from the experiment and from all payments. Please note that communication between participants is strictly prohibited during the experiment. Further instructions will be provided at the beginning of each step of the experiment.

Each of you will be given an identity number. Please do not lose your identity number. This entire experiment is anonymous. We will sometimes form pairs of two participants. The matching of two participants has been randomly determined in advance. You will never be informed of the identity of the participant with whom you have been matched. You will not know the names of your partner. We will use identity number for payment. If you lose your id number we wont be able to pay you. Please raise your hands if you have any questions.

Part 1 : Game 1

We divided the total number of participants in this experiment session randomly into two equal groups: Workers and Supervisors. Your role has been randomly assigned to you by us at the beginning of the experiment. All the participants in this room are Supervisors. All the workers are in another adjacent room.

Description of Game 1

In this part each one of you is a “Supervisor” and you are paired with a “Worker” who is in the other room. Each one of you will receive a booklet filled out by the worker you are paired with.
Nobody will ever be informed of the identity of the worker he/she is paired with. Your task is to count the number of correct answers in their Worker’s booklet and pay her 1 token for each correctly solved box. In order to do it, we provide you with the “Solution Manual” which contains the list of correct answers to the problems.

The Worker’s booklet

The worker’s booklet consists of simple math problem of the following type. It has boxes like this:

```
| 1.79 | 3.70 | 2.99 |
| 8.34 | 7.19 | 5.55 |
| 9.01 | 4.45 | 6.32 |
```

The worker has to find a pair of numbers in the box which add up to to 10. Having found the pair, they are required to encircle the corresponding numbers and put a tick into the box corresponding to “Got it”. In this example note that only 4.45 and 5.55 add up to 10. The booklet will contain 20 such boxes and if the worker solves all of them she earns 20 tokens. The identity number of the worker is written on top of the booklet.

Your Task

The Solution Manual given to you contains the correct answers to the problems. In each box the correct answers are underlined and bold. Your role is to find the number of correct answers in the worker’s booklet with the help of this manual. We will not monitor your work.

Description of Game 1

You will be paid 2 tokens for checking the worker’s booklet. Once you are done checking the booklet of the worker, throw the booklet in the basket along the door and proceed towards the token counter located outside the room. At the token counter, ask for the total number of tokens that you need to pay the worker plus your 2 tokens. You just need to mention only one number to the cashier.

You will then return to your seat and put the number of tokens that the worker has earned in the envelope and seal it. The workers identity number is written on the envelope. Then place the envelope containing the worker’s earnings on the table kept in front.

Finally you will redeem the tokens that are left with you from the token counter at the end of the experiment along with the other earnings that you will receive in the subsequent games we will play.

Note that when you ask for tokens at the token counter and when you pay the worker you may choose to do any one of the following:

1. You may **Over-report at the token counter** i.e. claim more tokens than you need and **Under-report the worker’s performance** i.e. put less tokens in the worker’s envelope than he deserves.

2. You may **Over-report at the token counter** i.e. claim more tokens than you need and be **Truthful with the worker’s performance** i.e. put the exact number of tokens that the worker deserves.
3. You may be **Truthful at the token counter** i.e. claim exact number of tokens you need and **Under-report the worker’s performance** i.e. put less tokens in the worker’s envelope than he deserves.

4. You may **Over-report at the token counter** i.e. claim more tokens than you need and **Over-report the worker’s performance** i.e. put more tokens in the worker’s envelope than he deserves.

5. You may be **Truthful at the token counter** i.e. claim exact number of tokens you need and be **Truthful with the worker’s performance** i.e. put the exact number of tokens that the worker deserves.

You will redeem these tokens from the token counter at the end of the experiment along with the other earnings that you will receive in the subsequent games we will play.

**Example to show how to calculate earnings**

The worker will get (number of correct answers you find in the booklet)*1 tokens. You earn 2 tokens.

For example, suppose the worker solves 19 problems correctly. Here the worker gets 19 tokens and you get 2 tokens.

In this case, you may ask for 19+2=21 tokens. **Note that you can ask any amount up to 22 tokens depending on the decision you make.**

You just need to mention 21 tokens to the cashier. **Please note that your earnings in this part is completely unrelated to that in the next part.**

**Game 1 begins now..**

Now please grade the booklet.

Please go out at the token counter and claim the total number of tokens you need.

Please put the worker’s earning in the envelope. Then put the booklet in the basket at the door and proceed towards the counter one by one.

If you have any questions regarding these instructions, please raise your hand. We will answer to your questions in private.

**Part 2**

Welcome back. Now we will play two games. We will randomly pick up one of these games by tossing a coin and the amount of tokens that you will earn in that particular game will be en-cashed at the end. If coin toss yields a Head then you will be paid based on your decision in Game 2 and if it is Tail then you will be paid based on your decision in Game 2. Notice that you have a fifty percent chance of being paid for Game 1 and a fifty percent chance of being paid for Game 2. However, since you do not know in advance which round will be used, your decisions in these two games are equally important.
Part 2: Game 1

In this game each one of you is given 10 tokens. We will provide you with a list of five charities. We ask you to decide how much would you like to give to one of these five charities given below. You can donate any amount including zero. For example if you decide to give 3 tokens then you get: 10 tokens - 3 tokens = 7 tokens. If this round is randomly selected by the coin toss then we will en-cash 7 tokens and pay you in cash at the end, while the 3 tokens will be en-cashed and sent to the charity you mentioned.

Please fill in the following in the sheet called Part 2.

“I want to give ______________ number of tokens to the ______________ charitable organization.

The list of potential charitable organization where you may donate the money to is given below with their respective address. We will calculate the total amount that has been donated to each charity right after the experimental session. A check will be written for each of the organization with the corresponding amount on it and put in a sealed, stamped and addressed envelope. The envelope will be posted by the end of the day. If the organization accepts online payment we shall pay online by the evening tonight.

1. Ramkrishna Mission. Contact address: The General Secretary, Ramakrishna Math and Ramakrishna Mission. P.O. Belur Math. Dist- Howrah. W.B. 711 202 India. Tel: (91) 33 - 2654 1144 / 1180 / 9581 / 9681. Fax (91) 33 - 2654 4346. E-mail: rkmhq@vsnl.com

2. Indian Red Cross Society. Contact address: Indian Red Cross Society, 1, Red Cross Road, New Delhi 110001, India. Tel: (+91-11) 23716441/2/3 Fax: (+91-11) 23717454;23717063 Web site : www.indianredcross.org

3. Prime Ministers National Relief Fund. Contact address : PMNRF, Prime Minister’s Office, New Delhi 110011


5. Missionaries of Charity. Contact address : 78 Acharya Jagadish Chandra Bose Road, Kolkata-700014, West Bengal , India. Telephone +91 33 22175267, +91 33 22640638.

Part 2: Game 2

In this game you will need to solve simple three digit addition problems. For example: What is 345 + 567?

You have 2 minutes to solve these addition problems. In these 2 minutes you will be given a booklet of 25 addition problems. You are encouraged to solve as many problems as you can.
Incentive Scheme

Before we start this game you need to make a decision. In this game we will pay you using either of the following three wage schemes. If this round is randomly selected for payment at the end we will use the wage scheme that you will choose now to pay you for this round.

1. Tournament Wage: You receive 8 tokens if you are among the top three performers.
2. Variable Wage: You will receive 0.4 tokens for each problem that you solve correctly during the work session.
3. Fixed Wage: You will receive a fixed amount of 4 tokens. It does not matter how many problems you solve.

Take a look at the example and decide your wage scheme.

Please indicate your choice of incentive scheme on the sheet called Part 2.

Task Session

Your 2 minutes of task session begins now. You are encouraged to solve as many problems as you can. After the 2 minutes is over we will ring a bell. Please bring your booklet in front and keep it inside the box.
Instructions for Workers

Introduction

You are now taking part in an economic decision making study. We will pay you Rs 300 for participating but you can earn additional money depending on the decisions you and the others make. During the experiment you can earn money by receiving tokens. All tokens that you can earn in the experiment will be exchanged into Rupees at the end of the experiment. The exchange rate is:

\[
1 \text{ Token} = \text{Rs. } 50
\]

At the end of the experiment you will receive the amount of money that you have earned during the experiment in cash. Please follow the instructions carefully. If you have any questions please let us know by raising your hand. Your question will be answered by us in private. In this experiment we will have to solve few math problems. Please do not use any helping devices like calculators, cell phones.

Please note that communication between participants is strictly prohibited during the experiment. Each of you will be given an identity number. Please do not lose your identity number. This entire experiment is anonymous. We will sometimes form pairs of two participants. The matching of two participants has been randomly determined in advance. You will never be informed of the identity of the participant with whom you have been matched. You will not know the names of your partner. We will use the identity number for payment. If you lose your identity number we won’t be able to pay you.

Please raise your hands once you have read the questions.

Part 1

We have divided the total number of participants in this experiment session randomly into two equal groups: Workers and Supervisors. All the participants in this room are Workers. All the Supervisors are located in another room.

Booklet

You will receive a booklet which consists of simple math problem of the following type. It has boxes like this:

\[
\begin{array}{ccc}
1.79 & 3.70 & 2.99 \\
8.34 & 7.19 & 5.55 \\
9.01 & 4.45 & 6.32 \\
\end{array}
\]

Your task will be to find a pair of numbers in the box which add up to 10. Having found the pair, they are required to encircle the corresponding numbers and put a tick into the box corresponding to “Got it”. In this example note that only 4.45 and 5.55 add up to 10. The booklet will contain 20
such boxes. The identity number of the worker is written on top of the booklet. You will receive 1 token for each correctly solved box.

We have randomly matched you with a Supervisor in the next room. Your booklet will be graded by that matched Supervisor. It is completely anonymous. Only the identity number is written on the booklet.

You can earn extra money based on how the supervisors grades your booklet.

Your Earning = (number of correct answers that the supervisor finds in the answer sheet)*1 tokens.

The supervisor is paid 2 tokens for grading the answers.

After checking the answers, the supervisor will ask for the tokens that is needed to pay you and him.

The supervisors may ask for (Your earning+2) tokens to the cashier. However they can also exercise one of the options given below.

1. He may **Over-report at the token counter** i.e. claim more tokens than he needs and **Under-report the worker’s performance** i.e. put less tokens in the worker’s envelope than he deserves.

2. He may **Over-report at the token counter** i.e. claim more tokens than he needs and be **Truthful with the worker’s performance** i.e. put the exact number of tokens that the worker deserves.

3. He may be **Truthful at the token counter** i.e. claim exact number of tokens he needs and **Under-report the worker’s performance** i.e. put less tokens in the worker’s envelope than he deserves.

4. He may **Over-report at the token counter** i.e. claim more tokens than he needs and **Over-report the worker’s performance** i.e. put more tokens in the worker’s envelope than he deserves.

5. He may be **Truthful at the token counter** i.e. claim exact number of tokens he needs and be **Truthful with the worker’s performance** i.e. put the exact number of tokens that the worker deserves.

This process is not monitored by anyone.

The supervisors will then make a decision and submit it. Your identity number will be written on the top of the envelope. You will then receive your corresponding payment.

You will have 10 minutes to complete the task. We will collect the booklet at the end of 10 minutes. Your time begins now....

Please be seated till we get your answer sheets graded and then you will receive your payment.

In the meantime please fill put this survey.
Exit Survey (Common for supervisor and worker)

Please fill out this exit questionnaire. Please write your id number on top.

Occupational choice

1. Do you want to join public/private sector service?
2. Are you preparing for civil service examination?
3. How many times have you attempted to take the civil service examination?
4. What is your preferred choice of service once selected?
5. Are you working at present? Which organisation are you working for?
6. Are you teaching in a college?
7. What is your expected salary for a future job?

Demographics

1. Gender
2. Age
3. What is your major or area of study?
4. Are you an undergraduate, Masters, MPhil or PhD student?
5. What is your home state?
6. Do you belong to Scheduled caste/ Scheduled tribe/Other backward class category?
7. Does any of your family member work in public sector and what is your relationship with that family member?
8. What is the approximate monthly income of your family?
9. What is your CGPA/GPA?
10. Are you involved in student politics? Which political party/parties are you affiliated with?
11. Do you receive any type of scholarship?
Exposure to corruption

1. How significant do you think that corruption is in different spheres of your life and lives of your friends and family around you?

(a) Political environment around you: 1 ≡ not significant, 2 ≡ somewhat significant, 3 ≡ significant and 4 ≡ very significant

(b) Business environment around you: 1 ≡ not significant, 2 ≡ somewhat significant, 3 ≡ significant and 4 ≡ very significant

(c) Job, recruitment and career progress: 1 ≡ not significant, 2 ≡ somewhat significant, 3 ≡ significant and 4 ≡ very significant

(d) Day to day affairs in your personal life: 1 ≡ not significant, 2 ≡ somewhat significant, 3 ≡ significant and 4 ≡ very significant

(e) Education system that you are or were part of: 1 ≡ not significant, 2 ≡ somewhat significant, 3 ≡ significant and 4 ≡ very significant

2. Were you ever asked to pay a bribe by a public sector/government official? 1. No 2. Yes

3. When in need of public health services was it common for you or your family to contact a relative, friend, or friend of a friend who worked in the health service and/or offer favours/gifts to health workers in order to improve the speed or quality of the health service? 1 ≡ not at all common, 2 ≡ not very common, 3 ≡ somewhat common and 4 ≡ very common

4. When trying to secure a job in the public sector, was it common for people you know in your social network to contact a relative, friend, or friend of a friend already working in a position of authority in the sector and/or offer favours/gifts to those in authority? 1 ≡ not at all common, 2 ≡ not very common, 3 ≡ somewhat common and 4 ≡ very common

5. When trying to resolve a problem in hands of the police, is it common for people whom you know in your neighbourhood to contact a relative, friend, or friend of a friend working in the police force and/or offer favours/gifts to police officers? 1 ≡ not at all common, 2 ≡ not very common, 3 ≡ somewhat common and 4 ≡ very common

6. How common is it for you or your family and friends to pay a bribe to expedite a public sector process like getting a passport, driving license etc.? 1 ≡ not at all common, 2 ≡ not very common, 3 ≡ somewhat common and 4 ≡ very common

Public Service Motivation

Mark the statements below on a scale of 1 to 4 where 1 ≡ strongly disagree, 2 ≡ disagree, 3 ≡ agree and 4 ≡ strongly agree

1. I have a special interest in developing public policy programs which are aimed at helping my country or community.

2. I regularly debate and discuss my views on public policies with others.

3. I greatly enjoy discussing political matters with others.
4. My greatest professional satisfaction is/will be when my ideas on public policies are implemented.

5. I get intensely interested in what is going on in my community.

6. I will prefer seeing public officials broaden the road in front of my house even if it means sacrificing my patio or verandah.

7. When public officials take an oath of office, I believe they accept obligations not expected of other citizens.

8. The huge financial incentives of a private sector job is well compensated by a low paying public sector job’s potential to implement transformational public policy.

9. I regularly participate in political rallies and other political process in my college/university (including college/university elections).

10. How many times did you vote in the last four general/state elections?
Chapter Five
Awareness programs and change in taste-based caste prejudice

Ritwik Banerjee and Nabanita Datta Gupta*

Department of Economics and Business,
Aarhus University

Abstract

Becker’s theory of taste-based discrimination predicts that relative employment of the discriminated social group will improve if there is a decrease in the level of prejudice for the marginally discriminating employer. In this paper we experimentally test this prediction offered by Becker (1971) in the context of caste in India, with management students (potential employers in the near future) as subjects. First, we measure caste prejudice and show that awareness through a TV social program reduces implicit prejudice against the lower caste and the reduction is sustained over time. Second, we find that the treatment reduces the prejudice levels of those in the left tail of the prejudice distribution - the group which can potentially affect real outcomes as predicted by the theory. And finally, a larger share of the treatment group subjects exhibit favorable opinion about reservation in jobs for the lower caste.

JEL Classification: C91, O1, J15

Keywords: Caste prejudice, taste-based discrimination, implicit association test, laboratory experiment, media influence

*Corresponding author: Email: ndg@econ.au.dk, Address: Fuglesangs Allé 4, 8210 Aarhus V, Denmark. We thank Niels-Hugo Blunch, Nishith Prakash, Manuela Ciotti, Emma von Essen and participants of the Aarhus University INCLUDE project’s start up workshop at ADRI, Patna for helpful comments. We also thank Sujoy Chakroborty, Bibek Banerjee, Raminder Kaur and Manas Pathak for their help in conducting the experiment. Funding from Danish Council for Independent Research | Social Sciences, grant no. DFF-1327-00037 is gratefully acknowledged.
1 Introduction

Caste-based\textsuperscript{1} discrimination in India has very deep roots and thus presents a good setting for studying the origins of discrimination. At the same time, quota-based affirmative action policy has been adopted in a larger scale in India than elsewhere with reservation of public sector jobs, political seats and slots in higher educational institutions for the lowest-caste groups in society. No reservation however exists in the private sector. Despite a social setting different from that in which racial prejudice against African Americans is observed, the issue of caste-based discrimination can be studied under the framework of the Beckerian theory (Becker, 1971) of taste biases.

Becker’s theory makes an important prediction - a decrease in prejudice level against a social group will lead to an increase in employment and reduction in wage inequality for that social group but it is not the average prejudice level but the prejudice level held by the marginally discriminating employer, which matters. However the implication of this prediction has not been analyzed in the context of caste in India. Furthermore, few studies are able to exogenously manipulate the prevalent prejudice levels and thereby cleanly identify effect of prejudice on wage and hiring outcomes. This problem is further compounded by the fact that the menu of possible interventions by which prejudice may be affected is rather limited.

In this paper we experimentally test the predictions offered by Becker in the context of caste, with management students (potential employers in the near future) as subjects. We design a unique lab-based experiment by which prejudice levels of the subjects are exogenously altered. Our experiment exposes groups of observationally equivalent Masters level business school students at a prominent business school in Delhi to either a TV program on caste injustices or an innocuous TV cartoon program. Their prejudice levels are then measured by a caste relevant Implicit Attitude Test (IAT; Rudman et al. (2001)) available through the Project Implicit website\textsuperscript{2} and their opinions regarding job reservation recorded.

Our results show that the treated group who watch the TV program show significantly lower implicit bias and state a significantly higher preference for private sector job reservation. These effects persist even 3 months after the treatment is carried out. We rule out that the changed behavior in our study is induced by a celebrity effect of a famous Bollywood movie star hosting the program. Instead, we show that the treatment shifts the prejudice distribution of our sample of highly-educated business school students (likely future employers) mainly at the lower tail. This fact, coupled with the additional finding that the treatment significantly increases a desire for

\textsuperscript{1}Caste is an age-old system of rigid social and ritual stratification of Indian society, implying the total exclusion of certain groups from the rights and opportunities for advancement. The most marginalized groups are the Scheduled Castes (SCs or Dalits) and the Scheduled Tribes (STs or tribals living in remote areas, or Adhivasis). Caste status is endogamous, rarely can be changed (except through religious conversion) and implies a rigid occupational specialization. The modern day caste-system is subdivided into almost 3,000 jati groupings, that are related in a complex way to the original varna subdivision (Deshpande, 2011).

\textsuperscript{2}Project Implicit is a multi-university research collaboration started in 1998 by aimed at systematically developing methods to elicit and measure automatic preferences.
private sector reservation indicates that caste-based discrimination in India by high-educated urban individuals is taste-based.

Thus our study is one of the first to characterize the nature of employer discrimination related to caste by exogenously manipulating prejudice levels in the lab. We also show that public awareness campaigns play an important instrumental role in reducing caste prejudice and its effect is persistent over time.

The rest of the paper is organized as follows: section 2 discusses the related literature, section 3 sketches a theoretical framework, section 4 lays out the experimental design and procedure, section 5 presents results, validation exercises and considers mechanisms and Section 6 provides a brief conclusion.

2 Related Literature

An extensive literature on racial discrimination largely looking at the black-white differences in the US has evolved over the years (see (Altonji and Blank, 1999)). Few empirical studies, however, directly measure taste-based prejudice and relate it labor market inequalities. Charles and Guryan (2008) who study the black-white wage gap in the US find that one quarter of the racial wage gap can be explained by prejudice. Carlsson and Rooth (2011) find similar results in the context of the ethnic wage gap in Sweden. In India, despite legislated policies on affirmative action, there still remains substantial employer discrimination by caste (documented by Siddique (2011), Banerjee et al. (2009)). But we know little about the nature of the employer discrimination and what strategies to pursue to ameliorate employment conditions of the low caste. Unlike most studies which rely on evidence for discrimination indirectly (e.g. Anwar (2012) and Fershtman and Gneezy (2001) look at differences in the belief distribution and altruistic preference respectively), IAT allows us to directly measure levels of prejudice. Rooth (2010) takes this idea to Swedish data and relates automatic preferences with discrimination in hiring immigrants in Sweden.

We also draw on another emerging literature showing that public awareness programs distributed via social media have a big impact on behavior. Such programs generally operate by changing the social norm as opposed to through purely economic channels such as price, income, quota etc. For instance in Brazil La Ferrara et al. (2012) finds that a television soap opera portraying small families lead to significantly lower fertility among women who are exposed to it. Such an effect has also been found by Kearney and Levine (2014) who combine Nielsen ratings, Google searches and Twitter tweets following a US TV program (16 and Pregnant) documenting the late-term pregnancies and early days of motherhood of teen mothers. They quantify its effects on teen childbearing outcomes and attribute a 5.7% reduction in teen childbearing in the 18 months following the program which amounts to around one third of the overall decline in teen childbearing over this period. While this study found a longer lasting effect of a TV social program on behavior, other programs (paid political advertising via television ads) had short-lived effects on (voter) preferences such as the
one analyzed by Gerber et al. (2011).

Psychologists have taken this a layer deeper by examining how such interventions affect our behavioral primitive - automatic preference, which in turn aggregates to social norms. Such automatic preference is unwitting, unintentional and uncontrollable and thereby different from explicit preference; yet it has real consequence in terms of affecting social judgment and real outcomes such as hiring decisions (Rooth, 2010). Rudman et al. (2001) found that diversity education administered by a likeable black professor over a period of 14 weeks reduced implicit prejudice by increasing contact with black students. Even in the short term, Dasgupta and Rivera (2008) find that situational exposure to positive gay role models reduced implicit prejudice against gay and lesbian people. In fact one of the early studies in Economics to use automatic preferences in understanding discrimination, Bertrand et al. (2005), notes that “manipulating” implicit preference and thereby remedying implicit discrimination is feasible and important as discriminating behavior can be partially reduced without requiring the discriminating agent to act against her will.

An IAT is a widely used computer based tool, designed to capture the automatic preferences vis-a-vis social groups, products or identity. It compares the speed of categorization in two different sorting conditions (Nosek et al., 2005). For example, the concepts “Low Caste” and “Bad” tend to be more strongly associated than the concepts “Low Caste” and “Good”; thus respondents are able to identify and categorize items faster in a condition in which items represent “Low Caste” and “Bad” than in which items representing “Low Caste” and “Good”. The test computes scores by calculating the difference in speed of association between the two sorting conditions. The details of the scoring algorithm can be found in Greenwald et al. (2003).

In this paper we target a group of likely future Indian employers – the graduates of two prestigious business schools in Delhi – and explore their implicit attitudes relating to caste and test whether we can experimentally affect such attitudes and more directly, hiring preferences. Our aim is to look deeper into the nature of employer caste bias in India, its consequences on real labor market outcomes and instruments by which caste bias can be reduced.

### 3 Theoretical Framework

In this section we revisit the seminal work of Becker (1971) on discrimination to examine what type of clear prediction can be derived in terms of employment and wage inequality in the presence of discrimination.

We assume that employers belong to the upper caste but workers may be hired from either low caste (l) or high caste (h). An employer’s utility is $V_i$ and it depends on the profit earned.

---

3For an excellent review of the literature on psychology see Hardin and Banaji (2012).

4Models of statistical discrimination (due to Arrow (1973)), the alternative theory of discrimination, cannot generate observed patterns in employment and unemployment (for a review of how the two strands of theory compare see Lang and Lehmann (2012)). Thus taste-based discrimination models remain the most relevant tool to study hiring preferences of employers.
Let us consider a thought experiment where due to some external intervention the distribution of \( d_i \) shifts to the left i.e. \( d_i \sim G(.) \) such that \( G(d_i) > F(d_i), \forall i \).

Prediction 1: The proportion of employers who will be willing to hire lower caste labor will be greater under \( G(.) \) than under \( F(.) \) i.e. \( G(d^*) > F(d^*) \).

Prediction 2: The wage inequality between lower and higher caste labor will be lower (i.e. \( w_l/w_h \) will be higher) under \( G(.) \) than under \( F(.) \).

Prediction 3: The extent of labor market inequalities (wage and employment) is driven by the discrimination level of the marginally discriminating employer.

Prediction 3 offers the important insight that wage and employment gaps should be systematically related not to the average prejudice level but to the marginal discriminator’s prejudice level. It is difficult to empirically validate this prediction as it is hard to obtain the distribution of prevailing prejudice, which is further compounded by the fact that prejudice itself is difficult to measure precisely. It is even harder to know who the marginal discriminator is. Becker provides a simple rule of thumb that under some (not-so-outrageous) conditions, the marginal discriminator is the one who has the \( p \)-th percentile prejudice level given that the \( p \) is the proportion of the discriminated group in the population. In other words, if the share of the lower caste in India’s population is 25% then the marginally discriminating upper caste employer is the one who has the 25th percentile prejudice level.

Closely tied to this theory, our study addresses three important questions. First, our experiment is one of the first to measure the (implicit) prejudice distribution among potential employers in the private corporate sector. Second, we aim to understand whether the instrument of awareness program is an effective policy instrument in so far as reducing the distribution of prejudice among potential employers. Third and most importantly, we examine which percentile the change in the prejudice distribution, if any, is coming from - this allows us to verify if the designed intervention is able to reduce the prejudice levels of the marginally discriminating employer. Getting more
lower caste labor hired by high caste employers is possible only when the change in distribution of $d_i$ is driven by a change in the marginally discriminating employer and not the average level of discrimination held by all employers. Finally, we cannot observe actual hiring decisions of the student subjects but we can proxy it by their opinion about job reservation. While admittedly the two are not the same, opinion about job reservation can be a good indicator of a future manager’s preference about a low caste person’s inclusion in the labor force. With that caveat, we shall analyze Prediction 1 i.e. if a shift in the distribution of $d_i$ affects opinion about reservation-based affirmative action.

4 Experiment Design and Procedure

Subjects are randomized into two treatment groups, which differ in the content of the video they watch. In one treatment they watch parts of an episode of a reality TV show - Satya Meva Jayate (SMJ), which documents the atrocities often meted out to lower caste people in India and the inequality of opportunity that follows. It further attempts to deliver a strong social message which indicts caste-based discriminatory behavior.

SMJ hosted by one of the most popular film stars in Bombay, was widely watched across India and dealt with a socially pressing issue in each episode. “Dignity for All” was one of the episodes and it dealt with caste-based social segmentation and discrimination in India. Subjects watch 30 minutes of the episode “Dignity for All” from the website of Satya Meva Jayate. In the control group subjects watch an episode of Tom and Jerry for thirty minutes. Tom and Jerry is chosen as a counterfactual as it has little cultural moorings and thus is likely to provide subjects with a neutral reference frame.

After watching the video for thirty minutes, subjects are directed to the Project Implicit website where they are asked to take the Caste - IAT. It is specifically designed to capture caste-based automatic preferences. In the test, subjects quickly categorize different word based stimuli (in our case it is surnames which denote caste) into left or right categories. If the stimulus belongs to the right (left) category, the right (left) key has to be pressed. An implicit prejudice against low caste will potentially show up as a response time differential. Based on this, IAT scores are computed and results generated. We ask the students to write down the results on the response sheet. It is important to note that the test scores are not observed, but what one observes is only the result of the test-taker. It is important to note that the results from the test are stated as one of the seven

---

5 The episode is one hour long, out of which 30 minutes including the most relevant parts are shown to the subjects.

6 Tom and Jerry, a video series with universal appeal, consists of an ongoing cat and mouse chase, where the less clever but stronger and larger cat tries in vain to capture the tiny yet clever mouse. It turns out that more often than not, it is the tiny mouse which outwits the cat in the series. Thus not only the *cartoonistique* manner of representation but also the the characters exhibit no hierarchical content and thus is an appropriate counterfactual for us. If anything, it is a show of solidarity with the weak (mouse) and therefore may only bias the IAT ranks upward and thereby negate the treatment effect.
alternatives. Since a subject sees one and only one of the following alternatives (see the Appendix for a screen-shot), it is impossible for her to falsify the true result and write something socially more desirable. In order to further prevent any falsification, the experimenter randomly matched the result on the computer screen with the one that the subject writes down on paper. The result is stated as one of the following seven alternatives -

- Strong automatic preference for Scheduled Caste compared to Forward Caste (level = 3)
- Moderate automatic preference for Scheduled Caste compared to Forward Caste (level = 2)
- Slight automatic preference for Scheduled Caste compared to Forward Caste (level = 1)
- No preference of for one caste over another (level = 0)
- Slight automatic preference for Forward Caste compared to Scheduled Caste (level = -1)
- Moderate automatic preference for Forward Caste compared to Scheduled Caste (level = -2)
- Strong automatic preference for Forward Caste compared to Scheduled Caste (level = -3)

For the purpose of empirical analysis, we convert the result into numerical scores levels as indicated above. The above design allows us to draw treatment effects between results from caste-based IAT score of those who watch the episode “Dignity for All” and that of those who watch Tom and Jerry.

In order to address whether the treatment effect is sustained over time, we followed the participants in the treatment over time. They take the caste-based IAT three months later again – this time without watching the relevant episode SMJ allowing us to map how the implicit preference change over time.

The experiment was conducted among MBA subjects at two reputed management training institutes in Delhi in India. MBA subjects provide an interesting subject pool for this experiment. The students are geared towards a private sector career where not only is there no caste-based reservation but representation of the lower castes is also small. The subject pool is also overwhelmingly upper caste and come from higher income strata. Unlike groups which are statedly biased and openly discriminate against lower castes, subjects in our sample represent that part of the population who most likely perceive themselves to be “neutral” and “unbiased” between castes but resist affirmative action mainly on grounds of efficiency.

Groups of students were randomly assigned to treatment or control. The instructions carefully explained the nature of the test. Since IAT required the test taker to associate different surnames with left and right categories, several examples of surnames and their caste affiliations were discussed. The experiment was done in two phases - in phase one (conducted in October, 2013 at International Management Institute), the treatment group had 29 subjects while the control had 30 subjects; in phase two (conducted in January, 2013 at Institute of Management Technology), the treatment and control groups had 18 and 17 subjects, respectively. Thus in all, the treatment and the control
Treatment 0 refers to the control group i.e. the group which watched Tom and Jerry. Treatment 1 refers to the treatment group i.e. the group which watched the Dignity for All episode of Satya Meva Jayate.

Groups had 47 subjects each. However, out of the 29 subjects who had appeared in the treatment group phase one, 25 subjects came back to retake the test after 3 months (attrition rate was only 13% and the average rank of those who were unavailable was 2). Subjects were paid Rs. 300 each for participating in the experiment for a session which lasted between 50 minutes and one hour. Subjects who took the IAT again three months later, were paid Rs. 200.

5 Results

Figure 1 shows the distribution of the IAT levels by treatment status. Visual inspection reveals that the IAT level distribution for the treated group has more variance and is clearly right-translated compared to the distribution for the control group. We formally test the difference in the distributions by treatment status. We perform the Jonckheere-Terpstra test which tests the null hypothesis that the distribution of the response variable does not vary across treatment groups. For our sample, the null is strongly rejected ($p$-value=0.007).

Descriptive statistics on the dependent variable and controls, separated by treatment status are

---

7Given this data and anecdotal investigations into the reasons for attrition, we believe that selection will not be an issue in the subsequent empirical analysis.
shown in Table 1 below. The dependent variable is the IAT Level which ranges from -3 (strong automatic preference for Forward Caste compared to Scheduled Caste) to 3 (strong automatic preference for Scheduled Caste compared to Forward Caste). Both Treatment and Control groups show an implicit preference for the Forward Caste. However, the mean difference in IAT levels between Treated and Control is positive and highly significant, meaning that the Treated are more positively inclined toward the Scheduled Caste\textsuperscript{8}. The magnitude of the difference corresponds to nearly one IAT level. Thus, while the Control group on average has a moderate automatic preference for Forward Caste compared to Scheduled Caste, the Treated group has a slight automatic preference for Forward Caste compared to Scheduled Caste.

Table 1: Descriptive Statistics

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Treatment</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>S.D.</td>
<td>Mean</td>
</tr>
<tr>
<td>IAT level</td>
<td>-2.26</td>
<td>1.05</td>
<td>-1.4</td>
</tr>
<tr>
<td>Share of Female</td>
<td>0.34</td>
<td>0.48</td>
<td>0.38</td>
</tr>
<tr>
<td>Age</td>
<td>23.81</td>
<td>1.66</td>
<td>24.3</td>
</tr>
<tr>
<td>Sec. exam level</td>
<td>86.93</td>
<td>6.08</td>
<td>87.06</td>
</tr>
<tr>
<td>Watched show before</td>
<td>0.32</td>
<td>0.47</td>
<td>0.51</td>
</tr>
<tr>
<td>Religiosity</td>
<td>2.53</td>
<td>0.88</td>
<td>2.6</td>
</tr>
<tr>
<td>Reservation for Public Sector</td>
<td>0.27</td>
<td>0.45</td>
<td>0.4</td>
</tr>
<tr>
<td>Reservation for Private Sector</td>
<td>0.04</td>
<td>0.2</td>
<td>0.19</td>
</tr>
<tr>
<td>N</td>
<td>47</td>
<td></td>
<td>47</td>
</tr>
</tbody>
</table>

† Note: *** p<0.01, ** p<0.05, * p<0.1 corresponding to mean difference using a t-test.

There are no significant differences according to treatment status in any of the control variables – i.e. share of female, age, secondary level exam level, religiosity\textsuperscript{9} or whether subjects had watched the show before. In both groups, around two thirds of the subjects are male, the mean age is around 24 years and the score on their (national) secondary level exam is 87\%, indicating that the subject pool is drawn from the higher end of the ability distribution. While 51\% of the Treatment group had watched the show before, only 32\% of the Control group had done so. The difference in the two shares is not, however, statistically significant\textsuperscript{10}. We can infer from these descriptive statistics that our subject pool, consisting of 1st and 2nd year business students drawn from business school in Delhi, is a homogeneous one.

\textsuperscript{8} The observed difference is not an institute specific feature - the average IAT levels in IMI are -2.3 and -1.2 while those in IMT are -2 and -1.6, for control and treatment groups, respectively.
\textsuperscript{9} Measured on an integer scale of 1 to 4 where 1 is Not at all religious and 4 is Highly religious.
\textsuperscript{10} We may be concerned that the Control group who was not exposed to the show may not have recalled watching it in the post-experiment survey. We cannot rule out that ‘Watched show before’ could be measured with error. This may potentially bias its effect downwards as well any variables correlated with it. The regression results are, however, robust to the addition of this variable, allaying any concerns about the presence of substantial measurement error.
Regression Results

Results from simple OLS regressions treating the IAT level as a continuous variable are presented in Table 2 below. Control variables are added in successive specifications but improve the fit only marginally, which is to be expected based on the findings of the means analysis above. The basic result is that being exposed to the treatment (watching the TV show on caste injustices) increases the IAT level by around 0.87, i.e. nearly a full IAT level, which corresponds roughly to the difference in raw means. The only other significant estimate is on previous exam level and shows that the higher the secondary exam level the lower is the IAT level\textsuperscript{11}. Thus, the more high achieving students show a larger caste bias - a percentage point increase in secondary exam score leads to an increase of about 5\% of an IAT level. The ordered nature of the dependent variable suggests an ordered categorical model (ordered probit) would be more appropriate than OLS. Model (6) estimates an ordered probit model and the main result remains unchanged. In a small sample such as this one, we are naturally concerned about the influence of outliers. The results of outlier tests based on the interquartile range showed no presence of serious outliers in the IAT level.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>0.851***</td>
<td>0.860***</td>
<td>0.866***</td>
<td>0.900***</td>
<td>0.865***</td>
<td>0.732***</td>
</tr>
<tr>
<td></td>
<td>(0.290)</td>
<td>(0.291)</td>
<td>(0.298)</td>
<td>(0.294)</td>
<td>(0.306)</td>
<td>(0.246)</td>
</tr>
<tr>
<td>Age</td>
<td>-0.0110</td>
<td>-0.0748</td>
<td>-0.0616</td>
<td>-0.0912</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.106)</td>
<td>(0.109)</td>
<td>(0.113)</td>
<td>(0.0896)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sec. Exam Score</td>
<td></td>
<td>-0.0529*</td>
<td>-0.0511*</td>
<td>-0.0461**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0268)</td>
<td>(0.0272)</td>
<td>(0.0211)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Watched show before</td>
<td>0.157</td>
<td>0.162</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.309)</td>
<td>(0.239)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Religiosity</td>
<td>-0.0558</td>
<td>-0.0341</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.166)</td>
<td>(0.129)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-2.255***</td>
<td>-2.182***</td>
<td>-1.916</td>
<td>4.162</td>
<td>3.781</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.205)</td>
<td>(0.230)</td>
<td>(2.562)</td>
<td>(3.981)</td>
<td>(4.127)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>94</td>
<td>94</td>
<td>94</td>
<td>94</td>
<td>94</td>
<td>94</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.085</td>
<td>0.090</td>
<td>0.091</td>
<td>0.129</td>
<td>0.132</td>
<td>0.056</td>
</tr>
</tbody>
</table>

\textsuperscript{†} Note: Numbers below the coefficients represent the Standard errors. *** p<0.01, ** p<0.05, * p<0.1. Outlier test showed 0.00\% severe outliers. R-squared in (6) refers to pseudo-R-squared.

\textsuperscript{11}It is worth noting though that the relation between IAT levels and secondary exam score is not strictly linear. This is evident from including a quadratic term of the exam score in the regression which shows that beyond 85\% exam score caste bias starts to decrease again (not reported).
Validation Exercise

Since the IAT test measures implicit or deep-seated bias, we may worry that simply being exposed to a one ½ hour TV show induces a potentially spurious or short-term change in the treated subjects’ preferences for the Scheduled Caste compared to the Forward Caste. We will explore this question in two ways: (1) We validate the findings above by comparing the differences in answers to two survey reported questions on whether the subject is in favor of public and private sector reservation for the lower castes. This also allows us to examine Prediction 1 from Section 2, namely a reduction in the discrimination leads to an increase in lower caste labor hiring (as proxied by job reservation opinion). (2) As mentioned earlier, we test the longer-run effect of the awareness program on implicit bias by asking treated subjects to take the IAT test 3 months after watching the TV show, but without watching the relevant episode SMJ again.

Regarding (1), unlike in the IAT which tests automatic associations, subjects have time to think about the questions on whether they are in favor of job reservation in the public and private sector. Thus their answers to these questions hence would reflect a more reasoned, conscious, and analytic belief.

40% of the subjects in the treatment group favored reservation in the public sector compared to 28% in the control group. A test of the difference showed that it was not statistically significant, Pearson Chi.sq(1)=1.706 (Pr=0.192). When it came to private sector reservation, 19% of the treated vs. 4% of the controlled were in favor. This difference was statistically significant at the 5% level, Pearson chi sq. (1)=5.045 (Pr=0.025). Since public sector reservation is already in place, the issue of introducing private sector reservation is possibly a more contentious one. In this case, we find that the treated group (who is otherwise identical to the control group) displaying a significantly stronger preference for reserving jobs in the private sector for the Scheduled Castes.

Regarding (2), Figure 2 shows the distribution of the change in the IAT level after 3 months for the subset of treated individuals who took the re-test. 60% of the retake sample shows either no change or a positive change (more positive towards Scheduled Caste) in their IAT level. This breaks down to 28% showing no change in their IAT level compared to the last test, and 32% having a change of >1 IAT levels compared to the last test. 40% show a deterioration of their IAT level meaning an increase in implicitly held caste bias. Half this group, however, levels only 1 IAT level lower. We can conclude that a clear majority of our sample show no change or improve their IAT level 3 months later. Only 25 treated subjects participated in the 3-month follow-up even though they represent 86% of the first treated group. Whether these individuals participated because they have a more positive attitude towards the Scheduled Caste than other treated subjects in their group or because they responded more to the financial incentive cannot, unfortunately, be determined. There are, however, no significant differences in the compositional characteristics of the follow-up group compared to the treated group they were drawn from.

Finally, we also found that retakers’ views on public and private sector reservation showed
Mechanisms

Our results quite convincingly show that the awareness promoting TV program reduces treated subjects’ implicit caste bias both in the short and longer-run. In this section, we explore some causal mechanisms. The reality show, Satya Meva Jayate (SMJ) was hosted by one of Bollywood’s most popular current film stars, Aamir Khan. It is possible that a celebrity effect is at play whereby subjects are strongly influenced to adopt the values and opinions endorsed by a celebrity. Celebrities are perceived to be more trustworthy and credible than non-celebrities (a halo effect) and may come across as possessing a more authentic connection to the behavior or product they are endorsing even though they do not possess any particular expertise in that area. For instance, individuals are far more likely to accept health advice from celebrities even though they may be ill-informed on these issues (Hoffman and Tan, 2013). This may arise in part through celebrities giving a clearer signal by inducing greater recall of the product and thereby enabling individuals to differentiate between otherwise similar products (Clark and Horstmann, 2013). Celebrities may also lead the herd in getting individuals to imitate their choices leading to an informational cascade (Bikhchandani et al., 1992). In our set-up, subjects watched the TV program individually at their own computer workstations and were not allowed to interact or to discuss the contents with other subjects ruling out herd effects through informational cascade. Neither do we believe that a strong

40% (25%) of retakers favored public (private) sector reservation, which matches almost exactly what we found 3 months earlier for the treatment group as a whole except that retakers were even more positive about private sector reservation (25% vs. 19%).
halo effect is present. If this were the case, subjects watching the show for the first time would be affected to a greater extent by this celebrity factor. As mentioned earlier, the control for having watched the show before was positive, but most likely due to low power, was insignificant in the regressions. Furthermore, estimating the IAT level regression on the two subsamples separately (watched before, did not watch before) produced very similar treatment effects - the estimate is 0.84 (Prob>|t|=0.09) for the sample who had watched the show before and 0.79 (Prob>|t|=0.04) for that who had not watched the test before. Formally conducting a Chow test of equality of the treatment effect and group effect across the two subsamples failed to reject the null F(2,90) = 0.24, Prob > F = 0.7905. The possibility that we are picking up an experimenter demand effect is minimal too since automatic preferences, as measured by IAT, cannot be altered consciously.

We believe, instead, that the treatment effect is a direct effect of the content of the program reducing taste-based discrimination against the lower castes in India. One of the central predictions of Becker’s theory of employer discrimination is that the relative hiring gap and the relative wage gap between majority and minority workers will be driven by the discriminatory preferences of the marginal employer who is just indifferent to hiring a majority or a minority worker, assuming firm sizes are all the same. This is because when the supply of minority workers is low, all minority workers will find employment with non-prejudiced employers who will pay them the same wage as majority workers. Where the wage gap arises is when the supply of minority workers exceeds the demand for them by non-prejudiced employers, forcing them to work (at lower wages) with prejudiced employers. The first prejudiced employer who is willing to hire a minority worker at a lower wage thus establishes the relative wage of minority and majority workers and hence the wage gap (Prediction 3). As it turns out, this is the not the most prejudiced employer but in fact the least prejudiced among the group of prejudiced employers.

Empirical support for this proposition has been found for the black-white wage gap in the U.S. (Charles and Guryan, 2008) and for the ethic wage gap in Sweden Carlsson and Rooth (2011). Charles and Guryan (2008) was the first to empirically apply Becker’s suggestion that the prejudice level of the marginal employer can be found at the p—th percentile of the majority employer prejudice distribution where p is the share of the workforce that consists of minority employees. Both papers access survey self-reported measures of prejudice and construct prejudice distributions on samples of individuals, not necessarily employers. Charles and Guryan (2008), however, do separate analyses for the high-educated (at least a college education) in their sample, who according to them are more likely to be employers than the average person. Individuals in our sample, similarly, are soon-to-be graduates of a prestigious business school in Delhi and represent a group of potential future employers. We estimate a quantile regression of the treatment effect on the prejudice level (measured as -1 times IAT level and thus the expected signs in Table 3 are opposite to that of Table 2) to check whether the treatment effect we find arising in the experiment originates in the lower tail of the prejudice distribution. The results appear in Table 3 below.

Here we can see that the treatment effect arises only at the left tail of the prejudice distribution
Table 3: Quantile Regression of the prejudice level on treatment

<table>
<thead>
<tr>
<th>Variables</th>
<th>10th</th>
<th>25th</th>
<th>40th</th>
<th>50th</th>
<th>60th</th>
<th>75th</th>
<th>90th</th>
</tr>
</thead>
<tbody>
<tr>
<td>treatment</td>
<td>-3***</td>
<td>-2**</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(1.032)</td>
<td>(0.765)</td>
<td>(0.911)</td>
<td>(0.861)</td>
<td>(0.795)</td>
<td>(0.211)</td>
<td>(0.0860)</td>
</tr>
<tr>
<td>Constant</td>
<td>1</td>
<td>2***</td>
<td>2***</td>
<td>3***</td>
<td>3***</td>
<td>3***</td>
<td>3***</td>
</tr>
<tr>
<td></td>
<td>(0.786)</td>
<td>(0.566)</td>
<td>(0.603)</td>
<td>(0.539)</td>
<td>(0.432)</td>
<td>(0.135)</td>
<td>(0.0549)</td>
</tr>
</tbody>
</table>

Observations 94 94 94 94 94 94 94

† Note: Numbers below the coefficients represent the Standard errors. *** p<0.01, ** p<0.05, * p<0.1. Each column shows the result from p-th quantile regression.

of potential employers. For the treated group the prejudice is 3 and 2 levels lower for the control group and significant at the 10th and 25th percentile respectively but not significant elsewhere in the distribution including the median. It should be pointed out the sample is overwhelmingly Hindu upper caste\(^{12}\), so the distribution approximately simulates the prejudice distribution of majority employers. The share of Scheduled Caste and Scheduled Tribe in the urban India’s population is 15.4% while that in rural India is 30%\(^{13}\). The proportion of lower caste in the workforce is approximately similar and thus awareness programs are capable of reducing the prejudice of at least the marginal employer. Further, the result from Table 3 above, coupled with the earlier finding that the treatment induced a desire to increase private sector reservation, implies that employer caste bias against the lower castes in India is taste-based because the findings accord with the predictions of Becker’s employer discrimination model. Unlike the papers mentioned earlier, however, we are only able to test this on the employers’ hiring preference and not for wage inequality. As Prediction 1 in the theory section shows, this is an equivalent test of the theory of taste-based discrimination.

6 Conclusion

We expose business school students at two prominent business schools in Delhi to either a TV program on caste injustices or an innocuous TV cartoon program. Following the viewings, we test subjects’ implicit caste biases via a caste IAT available through the Project Implicit website. Our results show that the treated group who watched the TV program showed significantly lower implicit bias and stated a higher preference for private sector job reservation. We also tested whether a longer-run treatment effect existed for a subset of the treated individuals and found that the lower bias levels either persisted or reduced even further 3 months after treatment for a majority of this group. That there is some evidence that the effect persists need further investigation in the future. Clearly, unlike priming, which is a subtle intervention with a short term effect, this intervention

\(^{12}\)There were 2 lower caste subjects and 1 Muslim subject. Leaving them out does not change significance in mean difference in IAT ranks between treatments \(p−value=0.01\).

\(^{13}\)Census 2011, Government of India.
seems to have a deeper bite. It is not entirely surprising though, since the intervention had a clear, strong moral content and may have set the subjects thinking about why they fared in the IAT as they did. However, the deeper processes involved in the observed long term persistence of effects need further investigations. We also argued that the effect does not come from the fact that a celebrity actor hosted the program as the treatment effect was identical across the group who had watched the show before and the group who watched the show for the first time. We bring further evidence that the treatment shifted the prejudice distribution of our sample of highly-educated business school students (likely future employers) mainly at the lower tail. This fact, coupled with the additional finding that the treatment significantly increased a desire for private sector reservation, indicates that caste-based discrimination in India by high-educated urban individuals is largely taste-based. It points to a potentially powerful role that awareness programs can play in reducing such bias.

References


Appendix 1

Figure 3: Screen shots from IAT

(a) Test Screen shots

The above picture shows four different screenshots from the IAT. Subjects were required to associate the words that came up with either the left or the right category.

(b) Result Screen Shot

Below is the interpretation of your IAT performance, followed by questions about what you think it means. The next page explains the task and has more information such as a summary of what most people show on this IAT.

Your data suggest a slight automatic preference for Forward Caste compared to Scheduled Caste.

The interpretation is described as 'automatic preference for Forward Caste if you responded faster when Forward Caste faces and Good words were classified with the same key than when Scheduled Caste faces and Good words were classified with the same key. Depending on the magnitude of your result, your automatic preference may be described as 'slight', 'moderate', 'strong', or 'little to no preference'. Alternatively, you may have received the following feedback: There were too many errors to determine a result.
Appendix 2 (Instructions and Procedural Details)

Procedural Details

The recruitment process was done by word of mouth and subjects were told that they would be participating in an experimental study on psychology/economics. The experiment was conducted in a large lecture hall and subjects watched the shows on their individual computers. They also had access to the website where the instructions were laid out and links to the videos, tests and surveys were provided. Subjects were sufficiently spaced out from one another and other standard protocols of experiments were followed. This is an entirely between subject study i.e. subjects who took part in one treatment did not take part in another.

Instructions for Satya Meva Jayate (Website link given below)

Welcome!

You are now taking part in an economic experiment. 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. It is prohibited to communicate with the other participants during the experiment. Should you have any questions please raise your hand and we will come to you. Please note down your identity number which you will require for payment and filling out your responses.

Overview

The study consists of an experiment and a survey.

Earnings

You will earn Rs. 200 for participating in this experiment.

Details

You will watch a video from youtube for the next half an hour. The video is part of the episode 'Dignity for all' taken from Satya Meva Jayate (SMJ). Following that you will be asked to take a computer based simple test. You will receive a feedback from the test which you must write down on the sheet of paper which is provided to you. After that you will fill out a survey questionnaire. In the survey questionnaire you will be asked to state your result from the test. You must report the result and fill out the rest of the questions.

Please follow the following simple steps.

Step 1: Please click here to go to the following video links. We will watch video 'Amir Speaks', parts of 1 (2:30 to 12:00), 4 (1:00 to 11:00) and 5 (0:00 to 3:00).

Step 2: Please click here and take the test.

Click on Caste and then proceed. Please skip the survey by clicking OK at the bottom of the page.

Please write down the result of the test on the response sheet provided to you.
Instructions for Tom and Jerry (Website link given below)

Welcome!
You are now taking part in an economic experiment. 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. It is prohibited to communicate with the other participants during the experiment. Should you have any questions please raise your hand and we will come to you. Please note down your identity number which you will require for payment and filling out your responses.

Overview
The study consists of an experiment and a survey.

Earnings
You will earn Rs. 200 for participating in this experiment.

Details
You will watch a video from youtube for the next half an hour. The video is an episode of the classic Tom and Jerry. Following that you will be asked to take a computer based simple test. You will receive a feedback from the test which you must write down on the sheet of paper which is provided to you. After that you will fill out a survey questionnaire. In the survey questionnaire you will be asked to state your result from the test. You must report the result and fill out the rest of the questions.

Please follow the following simple steps.

Step 1: Please click here to go to the youtube video. You will now watch the Tom and Jerry video for the next 24 minutes.

Step 2: Please click here and take the test.
Click on Caste and then proceed. Please skip the survey by clicking OK at the bottom of the page.

Please write down the result of the test on the response sheet provided to you.

Step 3: Please fill out one short survey. Survey for IAT click here.
PhD Theses since 1 July 2011

2011-4  Anders Bredahl Kock: Forecasting and Oracle Efficient Econometrics
2011-5  Christian Bach: The Game of Risk
2011-6  Stefan Holst Bache: Quantile Regression: Three Econometric Studies
2011:12  Bisheng Du: Essays on Advance Demand Information, Prioritization and Real Options in Inventory Management
2011:13  Christian Gormsen Schmidt: Exploring the Barriers to Globalization
2011:16  Dewi Fitriasari: Analyses of Social and Environmental Reporting as a Practice of Accountability to Stakeholders
2011:22  Sanne Hiller: Essays on International Trade and Migration: Firm Behavior, Networks and Barriers to Trade
2012-1  Johannes Tang Kristensen: From Determinants of Low Birthweight to Factor-Based Macroeconomic Forecasting
2012-2  Karina Hjortshøj Kjeldsen: Routing and Scheduling in Liner Shipping
2012-3  Soheil Abginehchi: Essays on Inventory Control in Presence of Multiple Sourcing
2012-4  Zhenjiang Qin: Essays on Heterogeneous Beliefs, Public Information, and Asset Pricing
2012-5  Lasse Frisgaard Gunnersen: Income Redistribution Policies
2012-6  Miriam Wüst: Essays on early investments in child health
2012-7  Yukai Yang: Modelling Nonlinear Vector Economic Time Series
2012-9  Henrik Nørholm: Structured Retail Products and Return Predictability
2012-10  Signe Frederiksen: Empirical Essays on Placements in Outside Home Care
2012-11 Mateusz P. Dziubinski: Essays on Financial Econometrics and Derivatives Pricing
2012-12 Jens Riis Andersen: Option Games under Incomplete Information
2012-13 Margit Malmmose: The Role of Management Accounting in New Public Management Reforms: Implications in a Socio-Political Health Care Context
2012-14 Laurent Callot: Large Panels and High-dimensional VAR
2012-15 Christian Rix-Nielsen: Strategic Investment
2013-1 Kenneth Lykke Sørensen: Essays on Wage Determination
2013-2 Tue Rauff Lind Christensen: Network Design Problems with Piecewise Linear Cost Functions
2013-4 Rune Bysted: Essays on Innovative Work Behavior
2013-5 Mikkel Nørlem Hermansen: Longer Human Lifespan and the Retirement Decision
2013-7 Mark Strøm Kristoffersen: Essays on Economic Policies over the Business Cycle
2013-8 Philipp Meinen: Essays on Firms in International Trade
2013-9 Cédric Gorinas: Essays on Marginalization and Integration of Immigrants and Young Criminals – A Labour Economics Perspective
2013-12 Paola Andrea Barrientos Quiroga: Essays on Development Economics
2013-13 Peter Bodnar: Essays on Warehouse Operations
2013-14 Rune Vammen Lesner: Essays on Determinants of Inequality
2013-15 Peter Arendorf Bache: Firms and International Trade
2013-16 Anders Laugesen: On Complementarities, Heterogeneous Firms, and International Trade
<table>
<thead>
<tr>
<th>Year</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>2013-17</td>
<td>Anders Bruun Jonassen: Regression Discontinuity Analyses of the Disincentive Effects of Increasing Social Assistance</td>
</tr>
<tr>
<td>2014-1</td>
<td>David Sloth Pedersen: A Journey into the Dark Arts of Quantitative Finance</td>
</tr>
<tr>
<td>2014-3</td>
<td>Lukas Bach: Routing and Scheduling Problems - Optimization using Exact and Heuristic Methods</td>
</tr>
<tr>
<td>2014-4</td>
<td>Tanja Groth: Regulatory impacts in relation to a renewable fuel CHP technology: A financial and socioeconomic analysis</td>
</tr>
<tr>
<td>2014-5</td>
<td>Niels Strange Hansen: Forecasting Based on Unobserved Variables</td>
</tr>
<tr>
<td>2014-6</td>
<td>Ritwik Banerjee: Economics of Misbehavior</td>
</tr>
</tbody>
</table>