2008s-26

Can Corruption Be Studied in the Lab? Comparing a Field and a Lab Experiment

Olivier Armantier, Amadou Boly

Série Scientifique
Scientific Series

Montréal
Septembre 2008

CIRANO
Allier savoir et décision

Centre interuniversitaire de recherche en analyse des organisations

Can Corruption Be Studied in the Lab?
Comparing a Field and a Lab Experiment[*]

Olivier Armantier[†], Amadou Boly[‡]

Résumé / *Abstract*

Le présent article essaie de tester la validité externe des expériences sur la corruption en quittant le laboratoire dans un pays développé pour le terrain dans un pays en développement, où la corruption importe encore plus. Dans notre expérience, un candidat offre un pot-de-vin à un correcteur afin d'obtenir une meilleure note. Nous trouvons que la direction et la magnitude de la plupart des effets de traitement sont statistiquement indistinctibles entre le laboratoire et le terrain. En particulier, augmenter la rémunération des correcteurs réduit la probabilité d'accepter le pot-de-vin aussi bien en laboratoire que sur le terrain. Enfin, nous identifions plusieurs micro-déterminants de la corruption (âge, religiosité, habilité).

Mots clés : corruption, économie expérimentale, expérience sur le terrain.

This paper makes an attempt at testing the external validity of corruption experiments by moving from the lab in a developed country, to where it matters the most, the field in a developing country. In our experiment a candidate proposes a bribe to a grader in order to obtain a better grade. We find the direction and the magnitude of most treatment effects to be statistically indistinguishable between the lab and the field. In particular, increasing the graders'.wage reduces in both environments the probability to accept the bribe. Finally, we identify several micro-determinants of corruption (age, religiousness, ability).

Keywords: *corruption, experimental economics, field experiments.*

Codes JEL : C91, C93, D73, I20.

[†] Federal Reserve Bank of New York, Université de Montréal, CIRANO, and CIREQ.
[‡] Département d'économie, Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal, Québec, Canada, H3C 3J7. E-mail: a.boly@umontreal.ca.

1 Introduction

The micro-determinants of corruption, as well as possible anti-corruption measures have recently been studied in lab experiments conducted in developed countries. If shown to be externally valid (i.e. to be relevant for the real world), then lab experiments could become one of the most effective tools to study corruption. One may wonder, however, whether the lab provides an appropriate setting to study corruption. Indeed, although non-monetary considerations (e.g. moral, ethical, legal, cultural) may be major determinants of corruption, they may be difficult to capture in lab experiments. In the present paper, we make an attempt at testing whether lab experiments conducted in developed countries can be used to study corruption where arguably it matters the most, the field in developing countries. To do so, we propose to compare the outcomes of a corruption experiment conducted in the lab in Montreal (Canada), and in the field in Ouagadougou (Burkina Faso).

There is ample evidence that corruption is now recognized as one of the most detrimental factors to economic and social development. First, corruption has become a prime concern for major international institutions. In particular, the International Monetary Fund considers that corruption "clearly is detrimental to economic activity and welfare".[1] Similarly, the World Bank declared that it "has identified corruption as the single greatest obstacle to economic and social development".[2] Second, several countries have intensified their fight against corruption. This may be best exemplified with the case of China where for the first time a high ranking official (the previous head of the Food and Drug Administration) was executed on July 10, 2007 after admitting to taking bribes. Likewise, Hong Kong created a legal precedent by implementing a "guilty until proven innocent" approach toward individuals accused of corruption. Finally, as discussed in a subsequent section, the last two decades have witnessed a sharp increase in the production of academic work related to corruption, not only in economics but also in sociology, political science, and law.

There is no comprehensive or widely accepted definition of corruption. It is generally agreed to include such activities as bribery, embezzlement, fraud, nepotism, extortion, or influence peddling. Corruption, however, is occasionally interpreted in a broader sense to encompass any activity lacking integrity, virtue, or moral. The definition of corruption is also sensitive to cultural factors. What may be considered corruption in a country, may simply reflect politeness or traditional gift exchange in a different country. Corruption is not necessarily illegal. For instance, although legal in several countries, lobbying and political contributions are deemed corrupt by most. Likewise, corruption is not necessarily considered immoral. For instance, favoritism toward one's own kind may be perceived as natural and justified. Corruption may be individual or collective, petty or grand, organized or incidental, political or bureaucratic. The latter is the form generally considered in economics where corruption is typically defined as an abuse of public office for private gain.

---

[1] http://www.imf.org/external/np/exr/facts/gov.htm.
[2] http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTPUBLICSECTORAND GOVERNANCE/EXTANTICORRUPTION/0„menuPK:384461~pagePK:149018~piPK :149093~theSitePK:384455,00.html).

Most agree that although corruption occurs in all countries, it is more prevalent and damaging in developing countries. Corruption affects all sectors of the economy, from tax collection to public contracting, from the justice to the education or health systems. In the present paper, we adopt the economists' definition of corruption, and we design an experiment related to bribery in the education system.

Due in part to its covered nature, the analysis of corruption has been challenging to economists. In fact, it may be argued that the theoretical and empirical approaches to corruption have been of limited practical impact to understand and combat corruption. Experimental economics, on the other hand, could become the most promising approach to understand the determinants of corruptibility, and test possible anti-corruption measures. Indeed, lab experiments offer the economist the possibility to overcome the unobservability of corrupt activities by generating hard data, while controlling both the environment and the characteristics of the subjects' population. The experimental literature on corruption, however, is in its infancy and its practical relevance will not be fully established as long as the question of external validity remains unaddressed. In other words, we need to test the extent to which the results of corruption experiments obtained in the lab can be extrapolated to real-life situations in the field. As argued by (e.g.) List (2006a) or Levitt and List (2007), a lab and a field experiment on corruption could produce different outcomes as some features of the experiment are unlikely to be exactly the same in the two environments. In particular, the nature of the game played, the stakes, or the subject pool may differ between the lab and the field. Testing the external validity of corruption experiments is therefore necessary and legitimate.

The present article may be considered a first step in this direction. Indeed, although the vast majority of experimental labs are located in developed countries, understanding and fighting corruption is considered most relevant to developing countries. It is therefore important to know whether the experimental results obtained in the lab in developed countries can be extrapolated to the field in developing countries. To address this question, we carried out a traditional lab experiment in Montreal, while in Ouagadougou, we conducted what Harrison and List (2004) refer to as a "natural field experiment" in which the subjects do not know they are participants in an experiment.[3] Our goal is to test whether the direction and/or the magnitude of various treatment effects are comparable when the same experiment is conducted in the lab and in the field. Note that we are not trying to demonstrate that one environment is superior to the other for the analysis of corruption. Likewise, we are not arguing that to be relevant lab experiments necessarily need to be fully replicable in the field. Like Harrison and List (2004), we believe that the lab and the field can complement each other, even when they produce slightly different outcomes. Our research project may then be interpreted as an effort to identify the dimensions in which the results obtained in the two environments concur or differ when analyzing corruption. We may then understand better how the lab and the field may be used as complementary approaches to study the determinants of corruption, as well as anti-corruption measures.

---

[3] As we shall see, our field subjects are not professional at their tasks. As a result, it may be argued that our field experiment does not fit exactly under Harrison and List's definition of a "natural field experiment".

Moving from the lab to the field without losing too much control is never an easy task. It is even more challenging in the case of corruption, as it is typically considered an illegal and immoral activity. To design our experiment we strived to minimize the possible losses of control under the constraints imposed by the lab and the field. The solution we propose aims at reproducing a corruption scenario in which a candidate offers a bribe to a grader in order to obtain a better grade. In short, subjects in the lab and in the field were asked to grade the same set of 20 exam papers in the same order. The 11th paper came with a money offer and a message stating: "Please, find few mistakes in my exam paper". The key difference between the two environments is that subjects in the field were informed they participated in an experiment only after grading was completed. We conducted four different treatments, each in the lab and in the field, by varying successively i) the amount of the bribe, ii) the wage paid to graders, and iii) the level of monitoring and punishment.

The experimental results indicate that the probability to accept a bribe decreases with the grader's age, religious fervor and ability at the grading task. Gender, however, seems to have no significant effect. Controlling for these individual characteristics, we find that increasing the wage paid to graders lowers their probability of accepting the bribe. Monitoring and possible sanctions appear to have no significant effect on the graders' acceptance behavior. The direction and the magnitude of the previous two treatment effects are statistically indistinguishable between the lab and the field. The two environments, however, differ in some dimensions. In particular, we find that increasing the amount of the bribe has no effect in the lab, while it exacerbates corruption in the field.

We also find that graders who accept the bribe are more likely to favor the briber by reporting fewer mistakes. Once we control for individual characteristics, we find that, unlike men, women respond to monitoring by failing the briber more often. In addition, graders who accept the bribe are more likely to help the briber when they receive a higher wage. These two treatment effects are similar in the lab and in the field. There is, however, a notable difference between the two environments: when the bribe is increased, accepters in the field are more likely to reciprocate by providing a passing grade to the briber. No such behavior is observed in the lab.

Finally, we find that accepting the bribe affects how subjects subsequently perform the grading task. Indeed, the grades reported for the remaining nine exam papers by subjects who accept the bribe are significantly less accurate and more inconsistent. In the monitoring treatment, however, subjects who accept the bribe do a better job at grading the last exam papers, possibly in an effort to lower their expected penalty. Once again, no significant difference between the lab and the field can be identified.

The remainder of the article is organized as follows. Section 2 briefly summarizes the theoretical, empirical and experimental literature related to corruption. The design of the experiment is presented in section 3, and discussed in section 4. The experimental results are analyzed in section 5. Finally, we discuss in section 6 the practical implications of our results.

2   The Economic Approach to Corruption

We provide here a short summary of the economic literature pertaining to corruption.[4] In doing so, we try to highlight the practical contributions, as well as the possible limitations of the theoretical, empirical, and experimental approaches.

2.1   Theoretical and Empirical Approaches

Most of the theoretical literature on corruption builds on Becker (1968) and Rose-Ackerman (1975), who were the first to analyze in a formal setting the economics of crime and corruption respectively. This literature combines elements of various models including principal-agent, resource allocation, rent seeking, governance, and contract theory. The first major theoretical debate concerned whether corruption can increase welfare. While initially seen as "grease-in-the-wheel" (Lui 1985), corruption has been mostly found to reduce welfare by acting as "sand-in-the-wheel".[5] Another major branch of the theoretical literature is concerned with evaluating possible corruption deterrents such as higher wages (Becker and Stigler 1974, Mookherjee and Png 1995, Sosa 2004), increased monitoring (Besley and McLaren 1993, Mookherjee and Png 1995, Acemoglu and Verdier 1998), stiffer penalties (Besley and McLaren 1993, Mookherjee and Png 1995, Acemoglu and Verdier 1998), and increase competition between potential bribees (Rose-Ackerman 1978, Shleifer and Vishny 1993). Most of this literature did not produce unambiguous predictions. As a result, it may be argued that the theoretical approach has been of limited practical impact to combat corruption. In addition, balancing the monetary costs and benefits of corruption may be considered too limiting of an approach (see e.g. Bukovansky 2006). In particular, accounting for non-monetary factors, such as ethical, moral, cultural and religious factors, may be necessary to better understand corruptibility and devise effective anti-corruption measures.

Corruption is by nature an illicit and secretive activity. As a result, it is virtually impossible to observe and measure directly. This lack of hard data partially explains the absence of rigorous empirical analyses of corruption prior to the mid 90's. To circumvent the observability problem, surveys aimed at evaluating corruption perceptions have been conducted around the world since the mid 1980's. Starting with the pioneer work of Mauro (1995), and Knack and Keefer (1995), these surveys have been used by economists to study corruption, giving rise to a burgeoning empirical literature. A first branch of this literature is concerned with the link between corruption and economic activity. The bulk of the research tends to conclude that corruption is harmful, as it appears to hinder growth (Méon and Sekkat 2005), widen income inequality (Gyimah-Brempong 2002), discourage investments (Wei 2000), and cause misallocations of public expenditures (Mauro 1997). A second branch of the empirical literature attempts to identify the causes of corruption. In particular, wealth and corruption have been found to be correlated, but the direction of the

---

[4]For more comprehensive surveys of the economic literature on corruption see Bardhan (1997), or Jain (2001). For a survey focusing more specifically on the theoretical approach, see Aidt (2003). For reviews of the empirical literature see Lanyi (2004), Dreher and Herzfeld (2005) or Seldadyo and Haan (2006). Finally, for surveys of laboratory experiments on corruption see Dusek, Ortmann and Lizal (2005), as well as Abbink (2006).

[5]See e.g. Shleifer and Vishny (1993), Banerjee (1997), Bliss and Di Tella (1997), or Ades and Di Tella (1999).

relationship appears to be ambiguous.[6] Corruption has also been found to be lower in countries open to foreign trade (Knack and Azfar 2003), with high human capital (Brunetti and Weder 2003), and a higher participation of women to the labour force (Swamy et al. 2001). Finally, several studies find that higher wages reduce corruption (e.g. van Rijckeghem and Weder 2001), while others find no support for this common hypothesis (Rauch and Evans 2000).

The empirical approach to corruption has been criticized on several fronts.[7] First and foremost, corruption indices are believed to suffer from significant measurement errors.[8] Second, the direction of causality, although often far from obvious, cannot be established with reduced form approaches. For instance, regression models cannot test whether corruption exacerbates poverty, or in contrast, whether corruption is more likely to breed in poor countries. Finally, there is a glaring void in the empirical literature when it comes to testing anti-corruption measures.[9] This void may be explained in part by the lack of micro-level data.

2.2 The Experimental Approach

The experimental approach to corruption is the most recent, with the first published paper dating back to the beginning of the decade (Frank and Schulze 2000). Although growing in popularity, there is to date, and to the best of our knowledge, less than ten experimental papers published on corruption. The bulk of these experiments has been conducted in the lab where two forms of corruption have been studied: embezzlement and bribery. While embezzlement experiments use dictator games to study corruptibility in a single decision making, bribery experiments build on the trust game literature to study corruption in a strategic environment.

These lab experiments produced the following results. Barr, Lindelow and Serneels (2004) as well as Jacquemet (2005) find a negative correlation between wages and corruption, while no such treatment effect has been detected in other studies (i.e. Frank and Schulze 2000, Abbink 2002, Schulze and Frank 2003). A deterrence effect is found in staff rotation (Abbink 2004), as well as in monitoring and punishment (Abbink, Irlenbusch and Renner 2002, Schulze and Frank 2003, Barr et al. 2004). The results in Frank and Schulze (2000), as well as Rivas (2006) suggest that women may be less corrupt than men, although Alatas et al. (2006a) suggest that the effect may depend on cultural factors. Likewise, tolerance with respect to corruption may differ across cultures (Cameron et al. 2006) and subject pools (Alatas et al. 2006b). Finally, the use of loaded instructions does not appear to generate a treatment effect (Abbink and Hennig-Schmidt 2006, Barr et al. 2004).

---

[6] While most studies find a negative correlation (Hall and Jones 1999, Fisman and Gatti 2002, Serra 2006), some find that wealth and corruption are positively related (Braun and Di Tella 2004, Frechette 2001).

[7] For a flavor of the debate pertaining to the measure of corruption with survey data see Golden and Picci (2005), or Kaufmann, Kraay, and Mastruzzi (2006).

[8] In particular, Johnston (2001) argues that perception based surveys are skewed toward petty bribery (as it is the most visible to the public), while the economic impact of grand corruption is much greater. Lanyi (2004) also notes that respondents may be reluctant to answer these surveys honestly if they have been directly involved in corrupt transactions. In particular, Murrell and Azfar (2005) estimate that roughly 40% of their sample was reticent in answering questions about corruption.

[9] There are a few notable exceptions including Brunetti and Weder (2003), Reinikka and Svensson (2004), or Ferraz and Finan (2008).

We are only aware of two published field experiments directly related to corruption.[10] As argued below, this void may be explained by the difficulty in designing field experiments on corruption. The first study, conducted by Bertrand et al. (2006) with subjects applying for a driver's license in New Delhi, combines experimental and survey methods to test the "grease-the-wheel" hypothesis. The second study, conducted by Olken (2007) in Indonesia, analyzes the efficacy of audit and grass-roots participation as anti-embezzlement measures in road construction projects.[11]

Although still not fully mature, the experimental approach to corruption may be considered very promising. Indeed, experimental economics has proved to be most useful in three situations, which each applies to corruption. First, when naturally-occurring data are scarce, or do not vary along certain desired dimensions. This is the case with corruption which is difficult to observe directly, and which is rarely observed under different wage, monitoring or punishment structures. In such cases, lab experiments enable the economist to generate data in order to identify the causal relationship between two variables of interest (e.g., between wages and corruption). Second, to identify the micro-determinants of behavior. Although crucial to understand and combat corruption, which micro-level factors explain corrupt behavior remain essentially an open question that neither the theoretical nor the empirical approach has been able to address adequately. As the economist controls both the experimental design and the characteristics of the subjects' population, the lab provides a unique framework to identify the micro-determinants of behavior. Third, lab experiments have proved to be a useful first step in the area of policy-making when a trial-and-error approach is either too costly or impossible to implement in the field.[12] As argued by (e.g.) Dusek et al. (2005) and Abbink (2006), lab experiments could constitute a cost effective "wind tunnel" to test potential policies aimed at curbing corruption.

The experimental literature on corruption, however, is in its infancy, and its practical relevance will not be fully established as long as the question of external validity remains unaddressed. In other words, we need to evaluate the extent to which the results of corruption experiments obtained in the lab can be extrapolated to real-life situations in the field. There are at least four potential reasons why the two environments could produce different results.[13] First, the stakes in the lab might differ from those in the field. In lab experiments, the stakes are essentially limited to "free money" provided by the experimenter. In the field, one of the party is entitled to the money,

---

[10]Some empirical studies have been conducted with data obtained after natural experiments (see e.g. Di Tella and Schargrodsky 2003, or Reinikka and Svensson 2004). We do not consider these to be field experiments as the researchers had no control over the design of the experiments. Recent unpublished field experiments on corruption include Castillo et al. (2007).

[11]Note that in both of these studies, corruption is not observed directly. As a consequence, alternative explanations of the results are possible. For instance, although they provide some evidence to the contrary, Bertrand et al. (2006) cannot exclude that, instead of resorting to corruption (as assumed by the authors), subjects simply exerted more effort to obtain their driver's licence.

[12]For examples of how experiments have helped in the design of various policies and markets, see Plott (1999), Roth (2002), Klemperer (2004), or Milgrom (2004).

[13]There are several examples in the literature showing that behavior in the lab does not necessarily extend to the field. For instance, List (2006b) finds significant discrepancies between the lab and the field when analyzing the behavior of sportscard dealers. Likewise, Henrich et al. (2001) find that behavior in ultimatum, dictator, and public goods games differs notably from the lab when conducted in the field in developing countries. See also, (e.g.) List (2006a) or Levitt and List (2007), for a discussion of potential factors that may explain why a lab and a field experiment may produce different outcomes.

and the stakes may not be purely monetary. For instance, corruption may lead to imprisonment, physical harm, and, in the most extreme cases, death. Similarly, corruption in the field may have moral or social implications that may be difficult to replicate in the lab. Second, the game played in the lab and in the field may be different. In the lab, corruption is often modeled as a one shot game played in a context-free environment between anonymous subjects. In the field, corruption may involve repeated interactions between parties who can identify each other, and whose decisions may have life long consequences. Third, the players may be different in the lab and the field. Although efforts have been made to extend the subjects' population, the roles in a lab experiment are typically assigned randomly to a self-selected group of students. In contrast, the distribution of roles in the field may be endogenous. For instance, it is possible that officials have learned to become corrupt, or that they have deliberately chosen their position because they are intrinsically more corruptible. Fourth, subjects in the lab know they are being scrutinized. As a result, Levitt and List (2007) argue that lab subjects may be inclined to make the "moral" choice when morality and wealth are competing objectives, as it is the case with corruption.

3   Experimental Design

We present in this section the basic experimental design, as well as the different treatments. The principal choices made when designing the experiment are discussed in the next section. The basic idea behind the experiment is to reproduce a corruption scenario in which a candidate proposes a bribe to a grader in order to obtain a better grade. As further explained in section 4, we concentrate on the graders' behavior in order to maintain as much control as possible over the experiment. In other words, although we have subjects acting as candidates, their role is essentially passive. We start with a description of the tasks asked to the different types of subjects, and then we turn to the different experimental treatments.

3.1   The Candidates (Montreal, Canada)

Subjects, called "candidates", were recruited to type a text on the computer as it was continuously dictated to them.[14] The text, based on a newspaper article in French, has 290 words and fits on two pages.[15] At the beginning of the session, each candidate was assigned to an isolated computer. Instructions were then read aloud, followed by questions.[16] We explained carefully what would, and what would not constitute a mistake. The subjects were also informed that at the end of the

---

[14] Such a dictation exercise is a classic test in the Francophone schooling system. It is typically conducted with pen and paper. The aim is to evaluate a candidate's spelling abilities, as well as its knowledge of the French grammar (many words in French have the same pronunciation, but different spelling). Such a test is administered several times a year to French students between the ages of 8 and 14. In addition, it is one of the requirements to obtain a secretary's diploma, and it is part of the entry-exam to certain civil servant positions. For lack of a better word, the pages on which a candidate typed the dictated text will be referred to hereafter as an "exam paper" or simply a "paper".

[15] As the field experiment was conducted in a Francophone country, we only recruited French speaking subjects to conduct the different laboratory and field sessions.

[16] A copy of the materials used for the experiment (i.e. the original text, the candidates' exam papers, as well as the instructions given to the candidates, the lab graders, and the field graders) may be found on one of the author's website at www.amaboly.com.

dictation, they would not be allowed to spell-check or modify their papers in any way. We told the candidates that we would decide whether their paper would be spell-checked by an experimenter, or by various subjects called "graders". Finally, we explained that a candidate's payment would depend in part on the number of mistakes the grader(s) would report. The lower the number of mistakes reported, the higher the payment.

Each candidate was also asked whether he would be willing to send some of the graders a money offer (explicitly referred to as "a bribe"), accompanied by the following message: "Please, find few mistakes in my exam paper".[17] We explain to the candidates that if they accepted to do so, then their payoffs may not depend exclusively on the number of mistakes reported. Instead, they may also be affected positively or negatively by each grader's decision to accept or reject the bribe. Finally, the candidates were informed that even if they accepted to do so, we would not necessarily send the message and the bribe to the graders.[18]

As justified in the subsequent section, we deliberately left the candidates' instructions partly ambiguous. In particular, we did not explain how we would select the papers to be graded by experimental subjects. Likewise, we did not specify the precise way in which the candidates' payoffs would be calculated. We also remained ambiguous about the amount of the bribe that would be proposed to the graders, as well as the exact consequences on the candidate's payoff when a grader accepts or rejects the bribe. The candidates were only informed that they would receive three payments: $C\$20$ payable immediately after the conclusion of the typing session, and two additional amounts to be paid respectively three and six months later. This delay allowed us to complete the various grading sessions, both in the lab and in the field. The candidates were told that each of the additional amounts could vary between $C\$20$ and $C\$60$ depending on the number of mistakes reported, and possibly, on the graders' decisions to accept or reject the bribe.

We conducted two typing sessions in March 2007 at the CIRANO's Bell Laboratory for Experimental Economics located in Montreal. Each session lasted roughly an hour, and included respectively 11 and 12 subjects. All 23 subjects accepted to send a bribe to the graders. On average, the candidates received a total payment of $C\$70.39$, with a maximum of $C\$121.44$ and a minimum of $C\$60.00$.

We now describe how we constructed the set of 20 exam papers to be graded by the other experimental subjects. In order to control the distribution of mistakes, we only selected 7 out of the candidates' 23 papers.[19] Out of these 7 papers, we chose a "bribe paper" with 20 mistakes. To complete the set of 20 exam papers, we made up 13 papers with various numbers of mistakes. As the papers would be graded in a specific order, we constructed and ordered the set of 20 exam papers in a precise way. First, we decided to place the bribe paper in 11th position. Second, we arranged so that the first and last set of 10 papers each has a symmetric and roughly identical distribution of mistakes. In particular, they have the same average (15.5), the same median (15.5),

---

[17]The literal translation from French is "If you please, may you find few spelling mistakes in my paper".

[18]Immediately after reading the instructions, subjects were given the possibility to leave the laboratory with $C\$10$ without having to type the text. None elected to do so.

[19]After we eliminated the papers with too many skipped words and too many mistakes, the selection process was made purely on the ground of convenience (i.e., to generate an appropriate distribution of mistakes).

and roughly the same standard deviation (6.8 versus 6.7). Third, we decided on a passing grade of 15 mistakes, meaning that if all mistakes were detected and reported, then half the papers would fail.[20] Finally, the exam papers were only identified by a 10-character code combining digits and letters. The first two digits, going from 01 to 20, identified the order in which the graders were asked to grade the papers. For the lab sessions, we only gave the graders the two pages of text. For the field sessions, we added a front page so as to look like a legitimate exam. This front page included in particular the identification code, as well as the instructions given to the candidates.

3.2   Field Graders (Ouagadougou, Burkina Faso)

The field experiment took place in July 2007 in Ouagadougou (Burkina Faso) during the national exams' period.[21] Burkina Faso is a landlocked country in West Africa with over 13 million inhabitants, among which 1.4 million live in the capital Ouagadougou. A former French colony, the country became independent in 1960. Since the 1990s, the country has enjoyed political stability with multiparty elections. Burkina Faso has been categorized by the World Bank as a "low income country". In 2005, its per capita income was US$400, compared to an average of US$590 for low income countries, and US$750 for Sub-Saharan African countries.[22] Formerly called the "Republic of Upper Volta", it was renamed "Burkina Faso" on August 4, 1984, which may be translated into the "Land of Incorruptible People." Somewhat ironically, Transparency International ranks Burkina Faso among the most corrupt countries in the world (i.e. 105th out of 179).[23] All sectors of the economy seem to be affected by corruption. In particular, the educational sector was ranked as the 6th (out of 10) most corrupt public sector in the country in 2003.[24]

To hire graders, we used the service of a local recruiting firm (*Opty-RH*). Flyers placed around Ouagadougou proposed a part-time job consisting in grading exam papers. The offer stated that the job consisted of two sessions: a grading session lasting half a day, followed a week later by a debriefing session during which graders would be paid. Having a high school diploma and a form of identification were the only documentation required. People interested were asked to come register in person at the recruiting firm location. Upon registering, graders were given the day, the time, and the location of their two sessions. At no point were the field subjects informed that they were about to participate in an experiment.

The grading sessions took place in a high school located in the center of Ouagadougou. Upon arrival, the subjects were gathered in a large room. Instructions on how to grade the exam papers were read aloud, followed by questions. Each grader was then randomly assigned to a private room where he found an envelope containing the 20 exam papers properly ordered, a report sheet, a red

---

[20]Such a failure rate is common in most exams and admission tests in Francophone countries.

[21]National exams must be passed to move from primary to middle school, middle to high school, and high school to college. The exam period typically lasts from June to the end of July.

[22]http://web.worldbank.org/WBSITE/EXTERNAL/COUNTRIES/AFRICAEXT/BURKINAFASOEXTN/ 0„menuPK:343886~pagePK:141132~piPK:141107~theSitePK:343876,00.html.

[23]It may be noted, however, that Burkina Faso is one of the least corrupt West African countries. It ranks 4th out of 19 countries.

[24]For additional information on the extent of corruption in the education sector in Burkina Faso see the "Etat de la Corruption au Burkina Faso, Rapport 2005" published by the Réseau National de Lutte Anti-Corruption.

pen, and an answer book (i.e. a copy of the text without mistake). No information was given to the graders about the nature of the exam, the candidates, or the institution which administered the exam. The graders were explicitly instructed to grade the papers in the proper order. After spell-checking a paper, the graders had to report the number of mistakes both on the front page of the paper, and on the report sheet. Graders were made aware that a candidate would fail the exam when more than 15 mistakes are reported. In such cases, we asked the graders to check the "Fail" column on the report sheet next to the number of mistakes. The graders were also instructed not to leave their room under any circumstance until they were done grading the 20 papers. We told them that we would stop by their room every 15 minutes precisely to answer any potential question. Grading therefore took place behind closed doors, and the graders knew they would be undisturbed except at regular 15 minute intervals. Once their task completed, we gave the graders an "IOU", and reminded them to come back the following week for the debriefing and payment session.

To introduce the bribe, we wrote the candidate's message on an easily removable "post-it", and we taped it with a banknote on the second page of paper 11. We made sure that the message and the money were i) attached securely, ii) not visible unless the exam paper was opened to the second page, and iii) discovered before the grader started spell-checking the paper.[25] When a grader reported the bribe attempt during one of our visits, we asked him to write in bold on the paper "fraud attempt". We took the banknote and the message, and we instructed the grader to spell-check the bribe paper just like any other paper. It is important to note that the instructions given to the graders at the beginning of the session specified that any attempt at fraud by a candidate would be penalized by failure of the exam. This information was also available in bold on the front page of each paper.

In the debriefing sessions, field graders were first informed that they took part in an experiment. The nature of the experiment was explained, and information was provided about the objective of the research and the use of the data collected. In particular, we explained that the data would be fully anonymized, and that whatever decisions a subject made during the grading session would be without consequence. After answering all the graders' questions, we asked each subject whether he or she would accept to sign an ex-post consent form giving us the right to use the data we collected on him or her. We informed the subjects that they did not have to sign the consent form, in which case their data would be destroyed. They also knew that refusing to sign the consent form was without consequence on their payment. All subjects, in all of the treatments conducted in the field accepted to sign the ex-post consent form.[26] Finally, the subjects filled a short questionnaire, after which they were paid in cash in return for the "IOU".

---

[25] Recall that in the field, an exam paper consists of three pages: a front page, plus two pages of text. The bribe and the message were therefore attached to the first page of text. Pictures of the exam paper with the bribe, as well as pictures of the high-school where the experiment took place are available on one of the author's webpage at www.amaboly.com.

[26] This result may reflect the prevalence of petty corruption in Burkinabe's everyday life. It is also possible that some subjects immediately and publicly accepted to sign the consent form as a way to signal that they did not take the bribe.

3.3 Lab Graders (Montreal, Canada)

The grading sessions were conducted at CIRANO's Lab in the fall of 2007. To the extent possible, we strove to follow exactly the same protocol as in Burkina Faso. In particular, subjects had to grade the same set of 20 papers in the same order. The graders were also provided with an isolated work station, a pen, a report sheet, and an answer book. The sessions had no time limit, and the graders could leave the lab once their task completed.

Compared to the field, however, graders in the lab were provided with different information. First, they knew from the start that they were taking part in an experiment. The corruption nature of the experiment, however, was not revealed immediately. Subjects were just told that they had to grade 20 papers in a specific order. Second, the lab graders were also informed that some papers had been typed by real subjects, named "candidates", while others had been made up by the experimenters. The exact ratio of real candidates was not specified, and the graders were informed that nothing would enable them to identify whether a paper had been typed by a real candidate or an experimenter. Third, we partially explained to the lab graders how the number of mistakes they report for a paper affects the candidates' payoffs. Namely, if a grader reports more than 15 mistakes, then the paper is not remunerated. If the number of mistakes reported is 15 or less, then the payoff depends on the number of mistakes. The lower the number of mistakes reported, the higher the remuneration for the candidate.

Finally, bribery was introduced in a different manner than in the field. The 20 papers were divided in two packs of 10. After completing the first pack, the graders were given the remaining 10 papers, along with additional written instructions to be read privately. These instructions stated that paper 11 had been typed by a real candidate, and that this candidate had accepted to send a message and a money offer to the grader. The instructions then revealed to the grader the exact message and the offer. The graders were free to accept or reject the offer, and the consequence of each decision was explained. If the offer was accepted, then the amount was credited to the grader and debited from the candidate. The grader was then free to decide on the number of mistakes to report, knowing that paper 11 would then be remunerated like any other paper, i.e. according to the number of mistakes reported. If the offer was rejected, then paper 11 was not remunerated. Nevertheless, we instructed the graders to spell-check paper 11, as well as the 9 remaining papers. At the end of the session, subjects had to fill a short questionnaire after which they were paid in cash.

3.4 The Experimental Treatments

We conducted four different grading treatments. Each treatment was conducted both in the lab and in the field. In the Control treatment, subjects were paid a fixed amount for their grading, regardless of how they performed the task. In addition, graders in the Control treatment were offered a bribe. In the lab, the fixed amount, called a wage hereafter, was 250 Experimental Units (EU hereafter) and the bribe was $50EU$. The conversion rate in the lab was $C\$1 = 12EU$. In the

field, the wage was set at $5,000FCFA$, and the bribe was set at $1,000FCFA$.[27] Note that the bribe to wage ratio is the same in both environments. In addition, observe that the lab payoffs are standard, while the wage paid in the field has been selected so as to be credible. Indeed, $5,000FCFA$ represents about 1/6 of the monthly minimum wage in Burkina Faso, and it roughly corresponds to the amount our graders could actually expect for a similar half-day job.

The "High Wage" treatment is identical to the Control treatment except that the wage was 40% higher (i.e. $7,000FCFA$ in the field and $350EU$ in the lab). The "High Bribe" treatment is identical to the Control treatment except that the amount of the bribe was doubled (i.e. $2,000FCFA$ in the field and $100EU$ in the lab). Finally, the last treatment makes an attempt at studying the effect of monitoring. For practical and ethical reasons, we decided against confronting field graders who accepted the bribe. Instead, we introduced a mechanism aimed at monitoring the accuracy with which the graders perform their task. This indirect approach therefore makes it possible to detect and punish corrupt graders when they favor the briber. The monitoring mechanism was explained as follows. We told each grader that we would randomly pick and re-grade 5 of the 20 papers he spell-checked. Then, we would calculate the difference between the number of mistakes reported by the grader and the number of mistakes we found in the paper. Only the worst paper graded would be considered to determine the monetary penalty. More precisely, the penalty imposed in the field was $2,000FCFA$ when the difference was between 3 and 5 mistakes, $3,000FCFA$ when the difference was between 6 and 9 mistakes and, $4,500FCFA$ when the difference exceeded 10 mistakes.[28] The penalties imposed in the lab were proportional.[29] Except for the risk of being penalized, the "Monitoring" treatment is otherwise identical to the Control treatment.

In total, 166 (respectively, 125) subjects participated in the four treatments conducted in the field (respectively, in the lab). More precisely, in the field (lab), 37 (30) subjects participated in the Control treatment, 40 (31) in the High Wage treatment, 45 (32) in the High Bribe treatment, and 44 (32) in the Monitoring treatment. On average, the total earning of a field grader (a lab grader) was $6,432.43FCFA$ (C33.24$) in the Control treatment, $8,375.00FCFA$ (C36.44$) in the High Wage treatment, $7,155.56FCFA$ (C41.31$) in the High Bribe treatment, and $4,954.55FCFA$ (C23.98$) in the Monitoring treatment.

4   Comments on the Experimental Design

To some extent, the design we proposed may be interpreted as the outcome of an optimization problem under constraints: we tried to minimize the possible losses of control when moving from the lab to the field, subject to the constraints imposed by the two environments. We now discuss some

---

[27]The Franc CFA is the currency used in Burkina Faso. In July 2007, the conversion rate was roughly C1$ for $442.9 \ FCFA$.

[28]Consider a perfect grader, i.e. able to identify all the mistakes in every paper. The expected payoffs when this grader accepts the bribe and reports a passing grade (i.e. 15 mistakes) is $(1,000-2,000/4) = 500FCFA$ plus his wage. The expected payoff when this grader accepts the bribe and reports 11 mistakes is $(1,000-3,000/4) = 250FCFA$ plus his wage.

[29]More precisely, the penalty imposed in the lab was $100EU$ when the difference was between 3 and 5 mistakes, $150EU$ when the difference was between 6 and 9 mistakes and, $225EU$ when the difference exceeded 10 mistakes.

of the issues we faced when designing our experiment, and some of the solutions we implemented to address these issues.

Running a field experiment on corruption is complicated by the fact that corruption is an illegal activity. As a result, we must be careful not to ask field subjects to take actions for which they could be prosecuted. To circumvent this problem, we created in Ouagadougou a private and closed environment in which to observe behavior. A second constraint was to find a real life activity which may credibly lend itself to corruption in the field. As mentioned earlier, corruption in education appears to be prevalent in Burkina Faso, and bribing a grader is not uncommon.[30] Another issue was to prevent contamination within the subject pool. In particular, we did not want the word to spread that the grading task was fake, or that bribes were present in the exam papers. We took several actions in order to mitigate this problem: i) we conducted the field experiment in a large city; ii) we tried to recruit a geographically diverse subject pool; iii) we conducted all the sessions within a ten day period; iv) between the sessions in which a bribe was offered, we conducted some additional sessions (not reported here) without bribe; and v) field subjects were informed that they participated in an experiment only after all sessions had been carried out.[31]

Our design also reflects some of the constraints imposed by traditional lab practices. In particular, we did not resort to the use of deception techniques (as defined by Hey 1998). This explains why, although their role is essentially passive, we used real candidates to type the text. As a result, lab graders knew that their decisions could truly impact the financial well being of other subjects. Likewise, following lab conventions we chose to preserve the subjects' anonymity, although it may be argued that the ability to identify the other party is a key feature of real life corruption.

We took a number of measures in an attempt to mitigate the possible losses of control when moving from the lab to the field. In particular, we decided to concentrate exclusively on the graders' decisions. As a result, we were able to control in both environments the amount of the bribe, the distribution of mistakes in the 20 exam papers, and the number of mistakes in the bribe paper. We also decided to introduce the bribe with a short written message in order to prevent face-to-face communication and informal bargaining, which could have been influenced by the briber's personal characteristics (e.g. gender, ethnicity). Finally, we had to choose what information to provide the lab and the field graders about the consequences of reporting the bribe attempt. We decided to provide information that we felt was of comparable nature in both environments. Namely, we told the field graders that reporting the bribe would be punished by failure of the exam, and we told the lab graders that this would provide no payoff for the candidate.

Before we conclude this section, we must acknowledge that the experiment we designed does not allow us to tackle entirely the question of external validity for corruption experiments. In particular, as the subjects recruited in the field are not professional graders, we cannot test whether

---

[30]In fact, a Burkinabe's newspaper ("Le Pays") reported on March 7, 2006 that two students were caught in a bribery attempt comparable to the one in our experiment. Incidentally, the first author was once offered (and obviously declined!) US$100 as a "Christmas gift" by a first year graduate student in his class, before the grades had been turned in.

[31]Informal conversations during the debriefing sessions suggest that, until it was revealed to them, the wide majority of field subjects truly believed that they were hired for a real grading task.

14

the endogeneity of the subject pool play a role in explaining corrupt behavior. Likewise, our design only allows us to analyze one side of the market, i.e. we observe the behavior of the bribees but not of the bribers. Finally, our one-shot game experiment may not fully capture corruption in the field, which may be learned and may involve repeated interactions. Nevertheless, following Levitt and List (2007), one may argue that our design includes one of the most important features for external validity. Namely, in contrast with the lab, subjects in the field acted without knowing they were participating in an experiment. As a result, our paper may be considered an important first step in testing the external validity of corruption experiments.

5   Experimental Results

In this section, we present the results of the lab and field experiments. To ease the presentation, we divide the analysis in three parts. First, we analyze the grader's decision to accept or reject the bribe. Then, we look at the number of mistakes reported for paper 11, depending on whether or not the grader accepted the bribe. Finally, we test whether accepting the bribe affects the way subjects subsequently perform their grading task for the remaining nine exam papers.

5.1   The Decision to Accept the Bribe

We start with a brief presentation of descriptive statistics. Then, we conduct an econometric analysis to identify treatment effects, and test for possible differences between the lab and the field. Table 1 gives the frequency of the bribe acceptance in the different treatments, both in the lab and in the field. Let us first concentrate on the results obtained in the lab. In the Control treatment, 67% of the subjects accepted the bribe. In other words, nearly one out of three graders essentially refused "free money" despite the fact that i) they did not incur any risk of being caught and ii) accepting the bribe had no negative externality on any other candidate. This rejection rate is slightly higher than in comparable lab experiments.[32] We conjecture two potential explanations for this result. First, unlike previous corruption experiments, our design requires real efforts from the briber and the bribee. Second, since most graders are university students, they can personally relate to the grading task, and therefore, they may be less tolerant toward bribery. Table 1 also indicates that in the lab, increasing the wage reduces the probability of accepting the bribe, while monitoring the graders' work or proposing a higher bribe seem to leave this probability unchanged.

As for the field, we can see in Table 1 that the probability of acceptance in the Control treatment is slightly below 50%. In contrast with the lab, this relatively high rejection rate may not be attributable solely to intrinsic motivations. Indeed, it is unlikely that all field graders believed that they faced absolutely no risk of being caught. In addition, they may have been under the impression that helping the briber had negative externalities on the other candidates and/or on the institution which administered the exam. Table 1 also indicates that in the field, providing a

---

[32]For instance, only 9.4% of the subjects in Frank and Schulze (2000) acted honestly even though corruption had a negative externality on an actual public entity (i.e. a film club). Likewise, the rejection rate was only 13.1% in Cameron et al. (2006) in an environment with negative externalities and possible punishment.

higher wage and monitoring the graders' work seem to lower slightly the probability of accepting the bribe. In contrast, the bribe is accepted more often when the amount proposed is larger.

We now turn to the econometric analysis. To impose as little structure as possible, we adopt a semiparametric approach to model the grader's decision to accept the bribe. More specifically, we specify a binary response model of the form:

$$A_i = \mathbb{I}\left(\beta' X_i + U_i \geq 0\right) \tag{1}$$

where A_i is a dummy variable taking the value 1 when grader i accepts the bribe; X_i is a vector of explanatory variables; β is a vector of parameters to be estimated; $\mathbb{I}(\cdot)$ is the indicator function satisfying $\mathbb{I}(z) = 1$ when z is true and $\mathbb{I}(z) = 0$ otherwise; and U_i is an unobserved error term. To estimate β without imposing any distributional assumption on U, we adopt the smoothed maximum score estimator developed by Horowitz (1992) (see also Horowitz 1998 and 2002).[33] This estimator can accommodate arbitrary heteroskedasticity of unknown form, it is asymptotically normal and, under some smoothness conditions, its convergence rate can be made arbitrarily close to $\sqrt{N}$. In other words, in terms of the estimator theoretical properties, there is virtually no cost in using this semiparametric approach versus a more conventional parametric approach.

In addition to treatment and field dummies, we include several individual characteristics to estimate the model in (1). In particular, we control for the grader's age and gender. We also include a measure of the grader's religiousness. This was obtained from the post-experiment survey in which we asked the subjects how often they go to a church, a mosque or any other place of worship.[34] This variable has 5 categories, ranging from 0 (never) to 4 (every day). To capture a grader's ability at the grading task we include two variables measuring the grader's precision and improvement over the first ten exam papers.[35] These measures are valid instruments since the graders (both in the field and in the lab) are unaware of the presence of corruption until they reach paper 11. Finally, the econometric model accounts for the time the grader took to complete the grading task.[36]

The descriptive statistics provided in Table 2, indicate that the individual characteristics vary

---

[33]If the distribution of U is normal (respectively, logistic), then (1) is the traditional binary probit (respectively, logit) model. Both of these distributional assumptions, however, are rejected by our data. As shown by (e.g.) Horowitz (2002) such distributional misspecifications may lead to severely biased estimates of β. Therefore, we prefer to rely on a more robust semiparametric approach. A well known drawback of the smoothed maximum score estimator, however, is that it does not produce marginal effects. This is a second order issue here since, to address the questions raised in this paper, we need to correctly identify treatment effects, not marginal effects.

[34]We did not anticipate the survey responses to play such a role in explaining behavior. In hindsight, we should have collected additional personal information about e.g. the subjects' wealth and occupation.

[35]To measure a grader's precision, we averaged over the first ten exam papers the absolute deviation between the number of mistakes he or she reported and the true number of mistakes. The variable "Precision" was then set equal to the opposite of this average. "Precision" is therefore a negative number, and a grader is considered more precise when this variable increases toward 0. To measure a grader's improvement, we calculated for each of the first ten exam papers the absolute percentage deviation between the number of mistakes he or she reported and the true number of mistakes. We then regressed for each grader the opposite of this variable on a constant and the exam number. The variable "Improvement" was then set equal to the estimated slope in this regression. A grader is therefore considered to have improved at the grading task when "Improvement" is positive and large.

[36]The parameter β in (1) is only identified up to a multiplicative factor. Following the conventional approach, the scale normalization consists here in setting the parameter associated with one of the variables equal to 1. In the estimations that follow, the variable selected for the scale normalization is the time the grader took to complete the grading task.

markedly between the field and the lab. In particular, observe that although it took the field subjects longer to complete their task, they are slightly less accurate at grading the first 10 papers. Note also that the average lab grader is less religious, and more likely to be a woman. By controlling these individual characteristics, the econometric analysis allows us to disentangle intrinsic differences between the lab and the field, from differences in the composition of the subject pools. In other words, we will be able to test whether two individuals with identical observable characteristics behave differently in each environment. In addition, the econometric analysis will allow us to test which of these individual characteristics may be considered micro-determinants of corruption.

The results of the estimation are reported in Table 3.[37] In terms of individual characteristics, we find that an older, a more religious or a more precise grader is significantly less likely to accept the bribe.[38] It is worth noting that, to the best of our knowledge, this is the first study to identify religious fervor, and ability as micro-determinants of corruption. Table 3 also indicates that gender does not appear to influence significantly the decision to accept the bribe. This result is somewhat surprising as previous lab experiments suggest that women are less corruptible (Frank and Schulze 2000, and Rivas 2006). It also contrasts with a commonly held belief among practitioners that women are less susceptible than men to accept bribes.[39]

Before testing for possible differences between the lab and the field, we look at general treatment effects. We can see in Table 3 that, compared to the Control treatment, increasing the wage paid to graders reduces significantly the probability of accepting the bribe. This result is consistent with the lab experiments of Barr et al. (2004) and Jacquemet (2005), and with several empirical analyses.[40] It is also in line with the views of numerous practitioners and international institutions, who often recommend to pay civil servants up to, or even above their private sector alternative as a mean to deter corruption.[41] Table 3 also indicates that proposing a higher bribe and monitoring do not affect significantly the grader's decision to accept the bribe. The first result contrasts with previous lab experiments in which a positive relationship between the bribe level and corruption has been identified (Abbink et al. 2002, Jacquemet 2005). The lack of efficacy of monitoring is not that surprising. Indeed, recall that the type of monitoring we implemented was not aimed at catching the graders who accept the bribe. Instead, it was designed to catch corrupt graders who reciprocate by giving a good grade to the briber.

To identify possible differences between the lab and the field, we included in the econometric

---

[37]The estimates presented in the paper are bias corrected. To account for the finiteness of the sample, the standard deviations of the estimates, as well as the distributions of the test statistics have been evaluated by bootstrap.

[38]Note that these results do not appear to be specific to the environment in which the data was collected. In particular, although the value of the parameters are affected, the direction and the significance of the individual characteristics remain unchanged when we estimate the model with the data collected solely in the lab, or solely in the field.

[39]For instance, the police department in Mexico City decided in 1999 to dispatch women traffic officers at sensitive intersections because they were deemed less corruptible than their male counterparts. See e.g. the August 15, 1999 New York Times article available on-line at http://query.nytimes.com/gst/fullpage.html?res=940CE7DA1239F936A2575BC0A96F958260&n= Top/Reference/Times%20Topics/Subjects/P/Police%20Brutality%20and%20Misconduct).

[40]See e.g. van Rijckeghem and Weber (2001) or Alt and Lassen (2003).

[41]Singapore and Hong Kong are often presented as successful examples of such a policy. Indeed, these countries are typically ranked among the least corrupt, and they are known to pay high salaries to their civil servants. In particular, the prime minister of Singapore is paid several times more than the U.S. president.

model a dummy variable for each treatment conducted in the field. We only find a single significant difference between the two environments. Namely, a higher bribe increases the probability of acceptance in the field, while it appears to have no effect in the lab. This result may indicate that subjects have different price elasticities in the lab and in the field. It may also simply reflect a pure level effect. Indeed, although the bribe is raised by the same factor in the two environments, the amount of the raise is different in the field and in the lab. The lack of statistical differences between the two environments, also implies that the direction and the magnitude of the "High Wage" treatment effect are the same in the lab and in the field. This result is remarkable as it suggests that, at least in some dimensions, corruption experiments conducted in the lab in a developed country and in the field in a developing country are fully consistent. The econometric analysis also reveals that the differences in the acceptance rate between the lab and the field (see Table 1) can be explained in large part by the composition of the subject pool in each environment. In other words, once we control for the subjects' observable characteristics, there does not seem to be any genuine intrinsic difference between the two environments, except in the "High Bribe" treatment.

5.2   The Decision to Report a Failing Grade

Do graders who accept the bribe tend to favor the briber? To address this question we concentrate here on the most relevant decision made by the grader: whether or not to report a failing grade (i.e. more than 15 mistakes) for paper 11.[42] The descriptive statistics in Table 1 indicate that graders who accept the bribe tend to fail paper 11 less often, regardless of the treatment or environment. Observe, however, that they do not systematically report a passing grade. In other words, despite being more detrimental to the briber than reporting the bribe, a non-negligible number of graders act opportunistically by taking the bribe and doing nothing in return.

To confirm that corrupt graders favor the briber, we estimate a binary response model similar to (1) in which the dependent variable takes the value 1 when the number of mistakes reported by the grader for paper 11 exceeds 15 mistakes. The results are reported in Table 4 (column labelled "All Data").[43] Observe first that after controlling for individual characteristics and treatment effects, we can confirm that graders who accept the bribe are significantly less likely to fail the bribe paper, and that their behavior does not vary significantly between the lab and the field.

In terms of the influence of individual characteristics, we find that, all else equal, women and more religious graders are less likely to fail the briber. In contrast, despite playing a role in explaining the decision to accept the bribe, the age of the grader appears to have no effect. Finally, and not

---

[42]An analysis based on the actual grade reported for paper 11 yields essentially the same conclusions.

[43]The econometric models for the decision to accept the bribe and the decision to report a failing grade have been estimated separately. Although this is consistent with the timing of the events, it is also conceivable that a grader checked the bribe paper before accepting the bribe. To test this hypothesis we estimated a bivariate binary response model in which the two decisions are modeled jointly. We find that the correlation between the error terms in each equation is not significantly different from zero, thereby rejecting the hypothesis of a joint decision. This does not imply, however, that we consider the two decisions to be uncorrelated. Indeed, by including a dummy for the bribe acceptance when estimating the decision to report a failing grade, we are only imposing that the error terms in the two models are independent.

surprisingly, the probability to find more than 15 mistakes in paper 11 is positively correlated with the grader's overall ability. As we shall see in the results presented next, most of these individual effects are consistent across graders, regardless of their decision to accept or reject the bribe.

No general treatment effect emerges from this econometric estimation with pooled data. As shown next, this may be explained by the fact that the treatments essentially affected the subjects who accepted the bribe. Observe, however, that even after controlling for their initial grading precision, women are significantly more likely to fail paper 11 in the monitoring treatment. This result is consistent with Frank and Schulze (2000) and Schulze and Frank (2003), who find that women are more responsive to monitoring and punishment. Note also that this behavior is robust as it is not significantly affected by the environment in which the experiment was conducted, and, as shown below, by the grader's decision to accept or reject the bribe.

To gain a better understanding of behavior, we now divide our sample in two groups depending on whether or not the grader accepted the bribe.[44] We can see in table 4 that the estimation results obtained for rejecters are consistent with those just presented.[45] In contrast, the behavior of accepters seems to be influenced by different factors. In particular, among accepters, an older more able male is more likely to fail the briber. In terms of treatment effects, we find that providing a higher wage decreases the probability that an accepter reports a failing grade. In other words, although graders are less likely to take the bribe when they receive a higher wage, those who accept tend to reciprocate more often by giving the briber a passing grade. Once again, the parameter associated with the "High Wage" dummy variable in the field is not significant. In other words, the direction and the magnitude of the "High Wage" treatment effect is statistically indistinguishable between the lab and the field. The only significant difference between the two environments is related to the effect of a higher bribe. Indeed, compared to the Control treatment, accepters in the field are more likely to reciprocate by providing a passing grade to the briber, while increasing the bribe does not influence the grade reported by accepters in the lab.

5.3   Corruption and Subsequent Performance

We now test whether accepting the bribe affects how well a subject subsequently grades the remaining nine exam papers. To do so, we exploit the panel structure of the data collected in the experiment to estimate a model of the form:

$$Y_{i,t} = \alpha' X_{i,t} + U_{i,t} \tag{2}$$

---

[44]To simplify, we refer to graders who accept the bribe as "accepters", and graders who reported the bribe as "rejecters". Note also that by splitting our sample we may introduce unobserved heterogeneity. Recall, however, that our semi-parametric estimation method can accommodate arbitrary heteroskedasticity of unknown form. In fact, we find essentially the same results when the models for accepters and rejecters are estimated jointly.

[45]The results obtained for rejecters should be interpreted with some degree of caution. Indeed, although they were instructed to grade paper 11 like any other exam paper, subjects both in the lab and in the field were informed that reporting the bribery attempt would result in failure for the candidate. One may therefore wonder whether the rejecters gave their best effort when grading the bribe paper. Two pieces of evidence seem to refute this hypothesis. First, an econometric analysis indicates that rejecters graded paper 11 with similar accuracy as the other first 10 papers. Second, the grades reported by rejecters for paper 11 in the Control treatment are not significantly different from those reported in an additional treatment we conducted in which no bribe was provided.

where the grading quality is defined as $Y_{i,t} = -|M_{i,t} - M_{0,t}|$, $M_{i,t}$ is the number of mistakes reported by subject i for exam paper t ($t = 12, ...20$), and $M_{0,t}$ is the actual number of mistakes in exam paper t.[46] The vector of explanatory variables $X_{i,t}$ includes the two variables measuring the initial ability of grader i over the first ten exam papers (i.e. "Precision" and "Improvement" defined in footnote 36). As we shall see, these variables capture most of the variation in grading quality across subjects.[47] To test whether grader i keeps improving as he or she did over the first ten exam papers, we control for i) the exam number t, and ii) the exam number multiplied by the value of variable "Improvement" for subject i. We also include in $X_{i,t}$ the time the subject took to complete the grading task, as well as various dummy variables for the decision to accept the bribe, the decision to fail the bribe paper, the environment, and the treatments. Finally, to control for possible grader specific random effects, we model the error term as $U_{i,t} = \varepsilon_i + V_{i,t}$, where $Var(V_{i,t}) = \sigma^2$, $Var(\varepsilon_i) = \sigma_a^2$ when subject i is an accepter and $Var(\varepsilon_i) = \sigma_r^2$ when subject i is a rejecter.

The results reported in Table 5 indicate that the parameters associated with the variables "Precision" and, to a lesser extent, "Improvement" are highly significant. This therefore confirms that most of the variation in grading quality over last nine papers may be explained by the subjects' initial abilities. The trend parameter is negative and significant, thereby indicating an overall decline in grading quality over the last 9 exam papers. The parameter associated with the variable "Improvement * t" is close to, but significantly greater than zero. In other words, we find a persistence in improvement, whereby (all else equal) subjects whose initial grading improved (respectively, deteriorated) over the first 10 exam papers, keep improving (deteriorating) after the bribe paper. Once we control for differences in initial ability, we find that subjects who accept the bribe are significantly less precise when grading the last nine exam papers. This lack of precision is even more pronounced among accepters who gave the briber a passing grade. Moreover, the standard error of the individual specific effect is significantly larger for the accepters than for the rejecters (i.e. $\hat{\sigma}_a > \hat{\sigma}_r$). In other words, the grading of accepters, and more specifically accepters who helped the briber, becomes more inconsistent and less accurate. To explain this result, we conjecture that accepters may prefer to appear incompetent rather than corrupt. Observe also that accepters do a significantly better grading job over the last nine exam papers in the monitoring treatment. In other words, it appears that accepters best respond to monitoring in an effort to lower their expected penalty.[48] Finally, we are once again unable to detect any significant difference between the lab and the field. Indeed, none of the parameters associated with variables controlling for the environment are significantly different from zero.

---

[46] To eliminate possible "exam paper" specific effects, the dependent variable $Y_{i,t}$ has been centered by subtracting its mean calculated over all graders.

[47] In particular, the estimations of alternative specifications indicate that accounting for "Precision" and "Improvement" is sufficient to capture general treatment effects, individual characteristics, as well as general differences between the lab and the field. These variables are therefore not included in the model estimated in this section.

[48] We are unable to detect a significant difference between the accepters who did or did not fail the briber. This result may be partially explained by the fact that accepters tend to report fewer mistakes for the bribe paper, even when they fail the briber. Doing a better grading job over the last nine exam papers is therefore a best response for both kinds of accepters.

6 Discussion

As argued by several international institutions (e.g. the IMF or the World Bank) corruption is one of the most detrimental factors currently afflicting the economies of developing countries. Due in part to its secretive nature, economists, have had limited success in their effort to understand and combat corruption. Recently, the micro-determinants of corruption, as well as possible anti-corruption measures have been tested in laboratory experiments conducted in developed countries. If shown to be externally valid (i.e. to be relevant for the real world), then laboratory experiments could become one of the most effective tools to study corruption. One may wonder, however, whether the insights gained in the lab in a developed country can be extrapolated to where it matters the most, the field in a developing country.

In an attempt to address this question, we conducted the same corruption experiment in the lab in Montreal (Canada), and in the field in Ouagadougou (Burkina Faso). The key difference between the two environments is that subjects in the field acted without knowing they were participating in an experiment. In short, our design aimed at reproducing a corruption scenario in which a candidate proposes a bribe to a grader in order to obtain a better grade. We conducted four different treatments, each in the lab and in the field, by varying successively i) the amount of the bribe, ii) the wage paid to graders, and iii) the level of monitoring and punishment.

An econometric analysis of the data collected in the lab and in the field reveals several micro-determinants of corrupt behavior. In particular, we find that the probability to accept a bribe decreases with the grader's age, religious fervor and ability at the grading task. In addition, our results suggest that women may be more responsive to monitoring and punishment. To the best of our knowledge, this is the first study to identify religious fervor, and ability as micro-determinants of corruption.

Once we control for these individual characteristics, we find the direction and the magnitude of several treatment effects to be statistically indistinguishable between the lab and the field. In particular, increasing the grader's wage reduces the probability that he will accept the bribe in both environments. In other words, we do not identify any intrinsic difference between the two environments, in the sense that the behaviors of two individuals with identical observable characteristics are not statistically different in the field and in the lab. This result is encouraging as it suggests that, at least in some dimensions, the results of corruption experiments conducted in the lab in a developed country carry over to the field in a developing country.

The outcomes of our experiment, however, differ in some dimensions when conducted in the lab or in the field. In particular, we find that doubling the amount of the bribe proposed to the grader has no effect in the lab, while it makes the grader more corrupt in the field. As suggested in Section 5, this result could simply reflect the fact that the amount of the increase was different in each environment. To test this hypothesis, and to further our understanding of the differences between the two environments, we are planning on conducting a lab experiment in Ouagadougou.

References

Abbink, K. (2002): "Fair Salaries and the Moral Costs of Corruption," Forthcoming in *European Journal of Political Economy*.

Abbink, K. (2004): "Staff Rotation as an Anti-Corruption Policy: An Experimental Study," *European Journal of Political Economy*, 20(4), 887-906.

Abbink, K. (2006): "Laboratory Experiments on Corruption," in *The Handbook of Corruption*, ed. by S. Rose-Ackerman. Edward Elgar Publishers, Cheltenham, UK, and Northampton, US.

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

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

Acemoglu, D., and T. Verdier (1998): "Property Rights, Corruption and the Allocation of Talent: A General Equilibrium Approach," *Economic Journal*, 108(450), 1381-1403.

Ades, A., and R. Di Tella (1999): "Rents, Competition, and Corruption," *American Economic Review*, 89(4), 982-993.

Aidt, T. S. (2003): "Economic Analysis of Corruption: A Survey," *Economic Journal*, 113(491), 632-652.

Alatas, V., L. Cameron, A. Chaudhuri, N. Erkal, and L. Gangadharan (2006a): "Gender and Corruption: Insights from an Experimental Analysis," Forthcoming in *Southern Economic Journal*.

Alatas, V., L. Cameron, A. Chaudhuri, N. Erkal, and L. Gangadharan (2006b): "Subject Pool Effects in a Corruption Experiment: A Comparison of Indonesian Public Servants and Indonesian Students," Forthcoming in *Experimental Economics*.

Alt, J. E., and D. D. Lassen (2003): "The Political Economy of Institutions and Corruption in American States," *Journal of Theoretical Politics*, 15(3), 341-365.

Banerjee, A. V. (1997): "A Theory of Misgovernance," *Quarterly Journal of Economics*, 112(4), 1289-1332.

Bardhan, P. (1997): "Corruption and Development: A Review of Issues," *Journal of Economic Literature*, 35, 1320-1346.

Barr, A., M. Lindelow, and P. Serneels (2004): "To Serve The Community Or Oneself: The Public Servant's Dilemma," Working paper, University of Oxford.

Becker, G. S. (1968): "Crime and Punishment: An Economic Approach," *Journal of Political Economy*, 76(2), 169-217.

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

Bertrand, M., S. Djankov, R. Hanna, and S. Mullainathan (2006): "Obtaining a Driving License in India: An Experimental Approach to Studying Corruption," *Quarterly Journal of Economics*, 122(4), 1639-1676.

Besley, T., and J. McLaren (1993): "Taxes and Bribery: The Role of Wage Incentives," *Economic Journal*, 103(416), 119-141.

Bliss, C., and R. Di Tella (1997): "Does Competition Kill Corruption?," *Journal of Political Economy*, 105(5), 1001-1023.

Braun, M., and R. Di Tella (2004): "Inflation, Inflation Variability, and Corruption," *Economics and Politics*, 16(1), 954-1985.

Brunetti, A., and B. Weder (2003): "A Free Press Is Bad News for Corruption," *Journal of Public Economics*, 87(7-8), 1801-1824.

Bukovansky, M. (2006): "The Hollowness of Anti-Corruption Discourse," *Review of International Political Economy*, 13, 181-209.

Cameron, L., A. Chaudhuri, N. Erkal, and L. Gangadharan (2006): "Propensities to Engage in and Punish Corrupt Behavior: Experimental Evidence from Australia, India, Indonesia and Singapore," Working paper, Department of Economics, University of Melbourne.

Castillo, M., R. Petrie, M. Torero, and A. Viceisza (2007): "Corruption in the Mail Sector: A Field Experiment in Peru," Working Paper, Georgia Institute of Technology.

Di Tella, R., and E. Schargrodsky (2003): "The Role of Wages and Auditing during a Crackdown on Corruption in the City of Buenos Aires," *Journal of Law and Economics*, 46(1), 269-292.

Dreher, A., and T. Herzfeld (2005): "The Economic Costs of Corruption: A Survey and New Evidence," Working paper 0506001, Public Economics, EconWPA.

Dusek, L., A. Ortmann, and L. Lizal (2005): "Understanding Corruption and Corruptibility Through Experiments: A Primer," *Prague Economic Papers*, 14(2), 147-162.

Ferraz, C., and F. Finan (2008): "Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes," *Quarterly Journal of Economics*, 123, 703-745.

Fisman, R., and R. Gatti (2002): "Decentralization and Corruption: Evidence Across Countries," *Journal of Public Economics*, 83, 325-345.

Frank, B., and G. G. Schulze (2000): "Does Economics Make Citizens Corrupt?," *Journal of Economic Behavior & Organization*, 43(1), 101-113.

Frechette, G. R. (2001): "An Empirical Investigation of the Determinants of Corruption: Rent, Competition, and Income Revisited," Working paper, Ohio State University.

Golden, M. A., and L. Picci (2005): "Proposal for a New Measure of Corruption, Illustrated with Italian Data," *Economics and Politics*, 17(1), 37-75.

Gyimah-Brempong, K. (2002): "Corruption, Economic Growth, and Income Inequality in Africa," *Economics of Governance*, 3(3), 183-209.

Hall, R. E., and C. I. Jones (1999): "Why Do Some Countries Produce So Much More Output Per Worker Than Others?," *Quarterly Journal of Economics*, 114(1), 83-116.

Harrison, G. W., and J. A. List (2004): "Field Experiments," *Journal of Economic Literature*, 42, 1009-1055.

Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, and R. McElreath (2001): "In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies," *American Economic Review*, 91(2), 73-78.

Hey, J. D. (1998): "Experimental Economics and Deception: A Comment," *Journal of Economic*

Psychology, 19, 397-401.

Horowitz, J. (1992): "A Smoothed Maximum Score Estimator for the Binary Response Model," *Econometrica*, 60, 505-531.

Horowitz, J. (1998): "Bootstrap Methods for Median Regression Models," *Econometrica*, 66, 1327-1351.

Horowitz, J. (2002): "Bootstrap Critical Values for Tests Based on the Smoothed Maximum Score Estimator," *Journal of Econometrics*, 111, 141-167.

Jacquemet, N. (2005): "Corruption as Betrayal : Experimental Evidence on Corruption Under Delegation," Working paper 0506, Groupe d'Analyse et de Theorie Economique (GATE).

Jain, A. K. (2001): "Corruption: A Review," *Journal of Economic Surveys*, 15, 71-121.

Johnston, M. (2001): "The Definitions Debate: Old Conflicts in New Guises," in *The Political Economy of Corruption*, ed. by A. K. Jain, pp. 11-31. Routledge, London.

Kaufmann, D., A. Kraay, and M. Mastruzzi (2006): "Measuring Corruption: Myths and Realities," Working paper, World Bank.

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

Knack, S., and O. Azfar (2003): "Trade Intensity, Country Size and Corruption," *Economics of Governance*, 4(1), 1-18.

Knack, S., and P. Keefer (1995): "Institutions and Economic Performance: Cross-Country Tests Using Alternative Institutional Measures," *Economics and Politics*, 7(3), 207-227.

Lanyi, A. (2004): "Measuring the Economic Impact of Corruption: A Survey," Working paper 04/04, IRIS, University of Maryland.

Levitt, S., and J. A. List (2007): "What Do Laboratory Experiments Measuring Social Preferences Tell Us about the Real World," *Journal of Economic Perspectives*, 21(2), 153-174.

List, J. A. (2006a): "Field Experiments: A Bridge between Lab and Naturally Occurring Data," *Advances in Economic Analysis & Policy*, 6 (2).

List, J. A. (2006b): "The Behavioralist Meets the Market: Measuring Social Preferences and Reputation Effects in Actual Transactions," *Journal of Political Economy*, 114, 1-37.

Lui, F. T. (1985): "An Equilibrium Queuing Model of Bribery," *Journal of Political Economy*, 93(4), 760-781.

Mauro, P. (1995): "Corruption and Growth," *Quarterly Journal of Economics*, 110(3), 681-712.

Mauro, P. (1997): "The Effects of Corruption on Growth, Investment, and Government Expenditure: A Cross-Country Analysis," in *Corruption and the Global Economy*, ed. by K. A. Elliott, pp. 83-107. Institute for International Economics, Washington, D.C.

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

Méon, P. G., and K. Sekkat (2005): "Does Corruption Grease or Sand the Wheels of Growth?," *Public Choice*, 122(1), 69-97.

Mookherjee, D., and I. P. L. Png (1995): "Corruptible Law Enforcers: How Should They Be Compensated?," *Economic Journal*, 105(428), 145-159.

Murrell, P., and O. Azfar (2005): "Identifying Reticent Respondents: Assessing the Quality of Survey Data on Corruption and Values," Working papers 05-001, University of Maryland, Department of Economics.

Olken, B. A. (2007): "Monitoring Corruption: Evidence from a Field Experiment in Indonesia," *Journal of Political Economy*, 115(2), 200-249.

Plott, C. R. (1999): "Policy and the Use of Laboratory Experimental Methodology in Economics," in *Uncertain Decisions Bridging Theory and Experiments*, ed. by L. Luini, pp. 293-315. Kluwer Academic Publishers, Boston.

Rauch, J. E., and P. B. Evans (2000): "Bureaucratic Structure and Bureaucratic Performance in Less Developed Countries," *Journal of Public Economics*, 75(1), 49-71.

Reinikka R., and J. Svensson (2004): "The Power of Information : Evidence from a Newspaper Campaign to Reduce Capture," Policy research working paper series 3239, World Bank.

Rivas, F. (2006): "An Experiment on Corruption and Gender," Discussion paper, Working paper 0806, Universitat Autonoma de Barcelona.

Rose-Ackerman, S. (1975): "The Economics of Corruption," *Journal of Public Economics*, 4(2), 187-203.

Rose-Ackerman, S. (1978): *Corruption: A Study in Political Economy*. New York: Academic Press.

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

Schulze, G. G., and B. Frank (2003): "Deterrence versus Intrinsic Motivation: Experimental Evidence on the Determinants of Corruptibility," *Economic governance*, 4, 143-160.

Seldadyo, H., and J. de Haan (2006): "The Determinants of Corruption: A Literature Survey and New Evidence," Working paper, University of Groningen.

Serra, D. (2006): "Empirical Determinants of Corruption: A Sensitivity Analysis," *Public Choice*, 126(1), 225-256.

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

Sosa, L. A. (2004): "Wages and Other Determinants of Corruption," *Review of Development Economics*, 8(4), 597-605.

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

van Rijckeghem, C., and B. Weder (2001): "Bureaucratic Corruption and the Rate of Temptation: Do Wages in the Civil Service Affect Corruption, and by How Much?," *Journal of Development Economics*, 65(2), 307-331.

Wei, S.-J. (2000): "How Taxing is Corruption on International Investors," *Review of Economics and Statistics*, 82(1), 1-11.

Table 1
Descriptive Statistics

Treatment		Control		High Wage		High Bribe		Monitoring	
Environment		Lab	Field	Lab	Field	Lab	Field	Lab	Field
Number of Subjects		30	37	31	40	32	45	32	44
% of Graders who Take the Bribe		0.67	0.48	0.48	0.37	0.66	0.69	0.66	0.41
% of Graders who Report more than 15 mistakes for Paper 11	All	0.67	0.62	0.62	0.42	0.66	0.42	0.66	0.68
	Accepters	0.60	0.50	0.40	0.27	0.61	0.35	0.67	0.72
	Refusers	0.80	0.74	0.69	0.52	0.72	0.57	0.64	0.65

Table 2
Subject Pool Characteristics

	Age		Female		Religiousness		Time (in Min)		Precision		Improvement	
	Lab	Field	Lab	Field	Lab	Field	Lab	Field	Lab	Field	Lab	Field
Average	26.26	24.86	0.41	0.16	0.83	2.68	100.21	140.74	11.82	11.28	-0.19	0.17
Standard Deviation	6.32	2.24	0.49	0.36	1.06	1.2	17.30	24.93	2.09	2.25	2.22	2.12
Min	18	20	0	0	0	0	50	70	4	4	-6.40	-6.90
Max	54	33	1	1	3	4	160	223	18	20	6.60	5.24

Table 3
Binary Response Model for the Decision to Accept the Bribe

Variable	Parameter Estimate	Variable	Parameter Estimate
Constant	0.901 (0.667)	"High Wage" Treatment	-1.382** (0.684)
Female	-0.752 (0.734)	"High bribe" Treatment	-0.686 (0.653)
Age	-0.689** (0.349)	"Monitoring" Treatment	-0.432 (1.116)
Religiousness	-1.128*** (0.437)	Field * Control Treatment	-1.103 (0.749)
Precision	-0.943** (0.478)	Field * "High Wage" Treatment	-0.560 (0.694)
Improvement	-0.455 (0.437)	Field * "High bribe" Treatment	1.431** (0.726)
		Field * "Monitoring" Treatment	0.130 (0.925)

Significance: * = 10%, ** =5%, *** = 1%

Table 4
Binary Response Model for the Decision to Report a Failing Grade

Variable	Parameter Estimate			Variable	Parameter Estimate		
	All Data	Accepters	Rejecters		All Data	Accepters	Rejecters
Constant	1.341** (0.563)	1.948*** (0.605)	4.864* (2.888)	"High Wage" Treatment	-0.203 (0.527)	-2.679*** (0.828)	-1.818 (2.173)
Female	-1.932*** (0.542)	-3.123*** (0.886)	-3.469* (2.147)	"High bribe" Treatment	0.229 (0.618)	0.653 (0.752)	-1.338 (2.435)
Age	0.093 (0.408)	1.275** (0.540)	-0.208 (0.366)	"Monitoring" Treatment	-0.465 (0.546)	-1.116 (0.825)	-4.562 (3.072)
Religiousness	-0.679** (0.234)	0.301 (0.366)	-1.182** (0.454)	Field * Control Treatment	-0.292 (0.695)	-0.356 (0.660)	-2.170 (2.959)
Precision	0.692** (0.364)	1.572*** (0.402)	1.139** (0.573)	Field * "High Wage" Treatment	-0.708 (0.636)	0.500 (2.821)	-2.355 (2.146)
Improvement	0.006 (0.285)	0.593** (0.265)	0.047 (0.532)	Field * "High bribe" Treatment	-0.134 (0.779)	-1.599** (0.745)	-1.387 (2.502)
Accept	-1.252** (0.551)	—	—	Field * "Monitoring" Treatment	0.584 (0.790)	1.413 (1.628)	0.383 (0.762)
Accept * Field	0.264 (0.632)	—	—	Female * "Monitoring"	2.067** (1.048)	3.990* (2.277)	4.596* (2.846)
				Female * Field * "Monitoring"	2.234 (1.709)	-0.537 (2.446)	—

Significance: * = 10%, ** =5%, *** = 1%

Table 5
Grading Quality after the Bribe Paper

Variable	Parameter Estimate	Variable	Parameter Estimate
Constant	4.378*** (0.773)	Accept * "High Wage" Treatment	0.575 (0.369)
Precision	0.598*** (0.084)	Accept * "High bribe" Treatment	0.454 (0.336)
Improvement	0.347** (0.177)	Accept * "Monitoring" Treatment	0.934** (0.365)
t	-0.090** (0.042)	Field	-0.472 (0.358)
Improvement * t	0.055** (0.020)	Field * Accept	0.172 (0.352)
Time	-0.282 (0.367)	Field * Pass	-0.045 (0.242)
Accept	-1.684*** (0.404)	σ	2.457*** (0.110)
Accept * Pass	-0.785** (0.252)	σ_a	1.219*** (0.106)
		σ_r	0.930*** (0.128)

Significance: * = 10%, ** =5%, *** = 1%