Can Corruption Be Studied in the Lab?

Olivier Armantier* Amadou Boly†

August 2010

Abstract

We investigate the external validity of corruption experiments, and more generally the universality of corrupt behavior, by conducting an experiment in three different environments: a lab in a developed country, a lab in a developing country, and the field in a developing country. In the experiment, a candidate proposes a bribe to a grader to obtain a better grade. We find the direction and magnitude of several treatment effects to be statistically indistinguishable across the three environments. In particular, increasing the graders' wage reduces the probability of accepting the bribe, but promotes reciprocation. Likewise, women are found to respond to monitoring and punishment in all three environments. Finally, we find that some individual characteristics (age, religiousness, ability) are correlated with corrupt behavior regardless of the environment. Our results are therefore consistent with the hypotheses that i) lab experiments on corruption have empirical relevance, and ii) some attitudes toward corruption may be universal.

Keywords : Corruption, Experimental Economics.
JEL Classifications : C91, C93, D73.

*Federal Reserve Bank of New York, CIRANO, CIREQ. E-mail: olivier.armantier@ny.frb.org.
†United Nations Industrial Development Organization.

We are grateful to Omar Al-Ubaydli, Jim Engle-Warwick, Gérard Gaudet, Glenn Harrison, Julie Heroux, Abraham Hollander, John List, Claude Montmarquette, and Soiliou Daw Namoro. We also thank seminar and conference participants at the Federal Reserve Bank of New York, SUNY Stony Brook, UC Santa Barbara, NYU, the Federal Reserve Bank of Boston, the ESA meeting in Tucson, the CSAE conference in Oxford, the IMEBE in Alicante, the New Techniques in Development Workshop in Canberra, the CEA in Vancouver, and the SCSE in Montebello. We would like to thank the recruiting firm Opty-RH for its help with hiring subjects in Burkina Faso. Finally, the authors gratefully acknowledge financing from the CRSH. As always, all remaining errors are ours. The views expressed here do not necessarily reflect those of the Federal Reserve Bank of New York, the Federal Reserve System, or the United Nations Industrial Development Organization.

1. Introduction

Corruption is now recognized as one of the most detrimental factors to economic and social development.[1] Because of its secretive nature, however, the analysis of corruption has been challenging to development economists. To circumvent the absence of hard data, public perception surveys have been used since the mid 1990's to conduct empirical analyses. This approach, however, has been criticized in part because of the potential for significant measurement errors.[2] As an alternative, lab experiments on corruption have recently been conducted.[3] This approach offers two key advantages: first, corrupt behavior is unambiguously observed at the individual level; second, the researcher controls both the environment and the characteristics of the subjects' population. As a result, the researcher is in a unique position to test possible corruption deterrents (e.g. higher wages, monitoring and punishment) and to identify micro-determinants of corruption (e.g. gender, religion). Lab experiments have therefore the potential to become a powerful tool in development economics to understand and fight corruption.

Although promising, the experimental literature on corruption is in its infancy and its practical relevance will not be fully established as long as the question of external validity remains unaddressed. A number of factors (e.g. different stakes, environments, populations) have been identified to explain why lab and field experiments do not always produce identical results (Harrison and List 2004, List 2006). Of particular interest here is the issue of scrutiny. Indeed, an inherent difference between lab experiments and the field, is that subjects know upon entering the lab that their behavior will be monitored and analyzed. Because of this scrutiny, Levitt and List (2007) argue that lab subjects may be more inclined to make the "moral" choice when morality and wealth are competing objectives, as it is the case with

[1] In particular, the World Bank states on its website that it "has identified corruption as the single greatest obstacle to economic and social development" (http://www1.worldbank.org/publicsector/anticorrupt/index.cfm).

[2] For a more general discussion of the issues related to the measure of corruption with perception-based surveys, see Golden and Picci (2005), or Kaufmann et al. (2006).

[3] Lab experiment conducted in developed countries include Frank and Schulze (2000), Abbink et al. (2002), Schulze and Frank (2003), Abbink (2004), Rivas (2008), Barr and Serra (2009). Lab experiment conducted in developing countries include Alatas et al. (2009a,b), Cameron et al. (2009), Barr et al. (2009). Although field experiments have become increasingly popular in development economics, few have been conducted on corruption in large part because they present major methodological challenges. We are aware of only three field experiments, all conducted in developing countries, directly related to corruption (Bertrand et al. 2006, Olken 2007, Castillo et al. 2009). Note also that these field experiments suffer from one of the same drawbacks as empirical studies as the act of corruption is not observed directly and must therefore be inferred.

corruption. In addition, although non-monetary considerations (e.g. moral, ethical, legal) may be major determinants of corruption, they may be difficult to capture in the lab.

More generally, one may wonder about the universality of corrupt behavior. The empirical literature in development economics has now clearly shown that the level of corruption may vary across countries and cultures depending (e.g.) on education levels, religious denominations, ethnic or language heterogeneity (see e.g. La porta et al. 1999, Treisman 2000, Fisman and Gatti 2002). Little is known, however, about whether corrupt behavior in different countries has the same determinants and responds to the same factors.[4] This issue is not only interesting in itself, but it is also particularly important to establish the practical relevance of lab experiments on corruption. Indeed, although understanding and fighting corruption are considered a priority for developing countries, lab experiments have been conducted primarily in developed countries where most labs are located. So, even if lab experiments can be shown to have empirical relevance for the field, we need to know where these lab experiments should be conducted.

The present paper raises two questions: First, do lab experiments on corruption conducted in developed and developing countries produce similar outcomes? Second, to what extent the results of corruption experiments obtained in the lab can be extrapolated to the field? In an attempt to address these questions, we carried out a corruption experiment in three different environments: a lab in developed country, a lab in developing country, and the field in a developing country. The experimental design essentially reproduces a corruption scenario in which a candidate offers a bribe to a grader to obtain a better grade. In short, subjects were asked to grade 20 exam papers. The 11^{th} paper came with a money offer and a message saying: "Please, find few mistakes in my exam paper." To determine whether subjects behave differently when they know they are being observed, subjects in the field were informed they had participated in an experiment only after grading was completed. We conducted four different treatments in each of the three environments by varying factors susceptible to promote or deter corruption: i) the amount of the bribe, ii) the wage paid to graders, and iii) the level of monitoring and punishment. We then tested whether the direction and the magnitude of the observed treatment effects differed across the

[4]Cameron et al. (2009) conducted a two-treatment experiment (the bribe is welfare enhancing in one treatment and welfare reducing in the other) to compare corruption propensities and punishments across subjects in Australia, India, Indonesia and Singapore. Our cross-cultural comparison differs in several important ways (e.g. we study corruption deterrents, our experimental task is framed and requires real efforts from the briber and the bribee), and it should therefore be considered a complement to Cameron et al. (2009).

three environments.

The remainder of the paper is organized as follows. The design of the experiment is presented in section 2. The experimental results are analyzed in section 3. Finally, we discuss the practical implications of our results in section 4.

2. Experimental Design

To conduct the same lab experiment in two different countries does not present any major methodological challenges. Moving from the lab to the field, however, is much more difficult, especially when studying corruption an illegal, covert, and immoral activity. In particular, we had to find a real-life activity that could i) credibly lend itself to corruption in the field without asking subjects to take actions for which they could be prosecuted, and ii) be replicated in the lab. In addition, given the difficulty in conducting a field experiment on corruption, we had to maintain as much control as possible over the experimental task to facilitate the comparison of behavior between the lab and the field. Finally, because of the inherent differences between the lab and the field, we had to adjust the way the experiment was implemented in each environment. In this section, we describe the experimental task we designed to meet these challenges, and we discuss how the experiment was implemented in the lab and in the field.

2.1. The Task

The experimental task given to subjects in all three environments consisted in spell-checking 20 "*dictées*." A dictée is a classic exam in the Francophone schooling system, whereby candidates type a text as it is continuously dictated to them.[5] The text, based on a newspaper article in French, has 290 words and fits on two pages.[6] To type some of the exam papers, we recruited 23 subjects (called "candidates" hereafter) in Montréal whose payoffs depended in large part i) on the number of mistakes

[5]The aim is to evaluate one's spelling ability and knowledge of French grammar (many words in French have the same pronunciation but different spelling). Such a test is administered several times a year to students between the ages of 8 and 14. In addition, it is one of the requirements for obtaining a secretary's diploma, and it is part of the entry exam to several civil servant positions. Although traditionally conducted with pen and paper, a dictée may also be typed. In particular, secretary's exams in Burkina Faso are still often typed on a typewriter and spell-checked by graders as in our experiment.

[6]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 the website of one of the authors, https://sites.google.com/site/olivierarmantier/.

subsequently identified, and ii) on their decision about whether or not to offer graders a bribe.[7] To control the distribution of mistakes in the set of papers to be graded, we selected only 7 out of the candidates' 23 papers.[8] Out of these 7 papers, we chose a "bribe paper" with 20 mistakes. Note that this bribe paper had been typed by a candidate who elected to bribe the graders. 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 ordered the set of 20 exam papers in a precise way. First, we decided to place the bribe paper in 11^{th} position.[9] Second, we ensured that the first and last set of 10 papers each had a symmetric and roughly identical distribution of mistakes, with the same average (15.5), the same median (15.5), 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 (including the bribe paper) would fail.[10] Finally, the exam papers were identified only 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 gave the graders only the two pages of text. For the field sessions, we added a front page so that it would look like a legitimate exam. This front page included in particular the identification code, as well as the instructions given to the candidates.

2.2. The Controlled Field Experiment (Ouagadougou)

Since the experimental design was essentially shaped by what is feasible in the field, we first describe how the controlled field experiment was implemented.[11] The field experiment took place in Ouagadougou, the capital of Burkina Faso. This former French colony located in West Africa is classified as a "low-income country" by the

[7] Since the role of the candidates is essentially passive, we refer the reader to appendix A where details are provided about the candidates and the typing sessions.

[8] After we eliminated the papers with too many skipped words and too many mistakes, the selection was made on the ground of convenience to generate an appropriate distribution of mistakes.

[9] Observe that by focusing exclusively on the graders' behavior we were able to fix the amount of the bribe, the distribution of mistakes in the 20 exam papers, the position of the bribe paper in the set of 20 exam papers, and the number of mistakes in the bribe paper. Maintaining this amount of control over the experiment makes it easier to compare the graders' decision to accept the bribe across the three environments.

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

[11] We summarize here the main features of the controlled field experiment conducted in Ouagadougou. For a more comprehensive description of the experimental procedures and a discussion of the issues raised by the implementation of the protocol in the field, we refer the reader to Armantier and Boly (2009) where additional treatments are analyzed.

World Bank and is known to suffer from a serious corruption problem.[12] In Burkina Faso, part-timers are regularly hired to grade exams.[13] Following this practice, we asked a local recruiting firm (Opty-RH) to place flyers around Ouagadougou proposing a part-time job as an exam grader during what is known as the "national exams' period."[14] In addition to a form of identification, the graders had to possess a university diploma or a proof of enrollment at a university. People interested in the part-time job were invited to register in person at the recruiting firm location. After validating their credentials, they were randomly assigned to a treatment and a session. At no point, however, were the subjects informed they were about to participate in an experiment.

The grading sessions took place in a high school located in the center of Ouagadougou.[15] Upon arrival, the subjects were gathered in a large room. We read aloud instructions on how to grade the exam papers and answered any 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 pen, and an answer book (i.e. a copy of the text without mistake). No information was given about the nature of the exam or the candidates. 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 told that a candidate would fail the exam when more than 15 mistakes are reported. In such cases, the graders had 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 had finished grading the 20 papers. We told them we would stop by their room every

[12]The corruption index reported by Transparency International for Burkina Faso has systematically been well below five (3.5 in 2008). Note also that bribing a grader is not uncommon in Burkina Faso. For instance, a Burkinabe's newspaper ("*Le Pays*") reported on March 7, 2006 that two students were caught in a bribery attempt similar to the one in our experiment.

[13]In fact, roughly 50% of the field subjects we interviewed reported having previously taken part in similar actual part-time grading jobs.

[14]Students must pass national exams to move from primary to middle school, from middle to high school, and from high school to college. This is the main exam period in Burkina Faso and it typically lasts from June to the end of July. To prevent contamination within the subject pool (e.g. through communications between sessions), we conducted all the experimental sessions within a week and we tried to recruit a geographically diverse subject pool. See Armantier and Boly (2009) for further details on the recruiting procedure.

[15]Note that every feature of the experiment (from the recruiting, to the design of the task, to the way the grading sessions were conducted) was chosen so as to appear like a regular part-time grading job in Burkina Faso. None of the subjects, even those with prior experience with such part-time grading jobs, questioned the credibility of the experiment.

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.

To introduce the bribe, we handwrote "Please, find few mistakes in my exam paper" on an easily removable "post-it" and taped it with a banknote on the second page of paper 11.[16] 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.[17] When a grader reported the bribe attempt during one of our visits, we asked him to write in bold "fraud attempt" on the paper. We took the banknote and the message, and instructed the grader to spell-check the bribe paper just like any other paper. Note that the graders were told in the instructions that any candidate that attempted fraud would be penalized by failure of the exam. In other words, the graders knew the implications of reporting the bribe. Finally, the field subjects were informed that they had taken part in an experiment only after grading was completed, in a debriefing session during which they were paid in cash.[18]

2.3. The Lab Experiments (Montréal and Ouagadougou)

We conducted the same lab experiment in Montréal and in Ouagadougou. The lab experiment in Ouagadougou was conducted roughly a year after the field experiment with a different set of subjects. The recruiting procedure was essentially the same, that is, the same recruiting firm posted flyers around town and we imposed the same educational restriction. The only difference is that the flyers used for the lab experiment specified that we were looking for subjects to participate in an economic experiment. The lab experiment in Montréal was conducted at CIRANO's Bell Laboratory for Experimental Economics using the lab's standard recruiting procedures (i.e. random E-mail solicitations from the subjects' database).

[16] This approach prevents face-to-face communication and informal bargaining which could have been influenced by the briber's personal characteristics (e.g. gender, ethnicity). Keeping the briber anonymous therefore facilitates the comparison of behaviors across the three environments.

[17] Recall that an exam paper in the field experiment 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 papers, including the exam paper with the bribe, as well as pictures of the high-school where the experiment took place are available on the first author's webpage at https://sites.google.com/site/olivierarmantier/.

[18] To close the loop, we could have conducted the same field experiment in Montréal. We elected not to do so because it would have been difficult to convince subjects in Montréal that they were hired for a real grading task. Indeed, unlike Burkina Faso, hiring part-timers to grade exams, and in particular dictées, is extremely rare in Canada.

The task for the lab experiments was the same as in the field. The subjects had to grade the same set of 20 papers in the same order. To the extent possible, we also tried to follow the same protocol when implementing the lab and the field experiments. The lab graders were provided with an isolated work station, a pen, a report sheet, and an answer book. As in the field, the lab sessions had no time limit, and the graders could leave the lab once their task was completed. Finally, the same experimenter supervised all the lab and field sessions.

To assess whether the experimenter's scrutiny affect behavior, and consistent with previous lab experiments on corruption, lab subjects were informed from the start that they were taking part in an experiment. The corruption nature of the experiment, however, was not revealed immediately. As in the field experiment, subjects were just told they had to grade 20 papers in a specific order. To prevent deception (as defined by Hey 1998), the lab graders were informed that some papers had been typed by real candidates, while other indistinguishable papers had been made up by the experimenters. The exact ratio of real candidates was not specified, and the subjects were asked to grade each exam as if typed by a real candidate. We also explained to the lab graders that the consequence of their grading decisions were purely monetary. Namely, the graders knew that a real candidate would receive no payment when more than 15 mistakes are reported. Otherwise, the graders were told that the lower the number of mistakes reported between 0 and 15, the higher the remuneration for the real candidate.

To introduce the bribe in the lab we divided the 20 papers into 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 agreed to send a message and a money offer to the grader. The instructions then revealed to the grader the exact message and the offer.[19] The grader was free to accept or reject the offer, and the consequence of each decision was explained. The grader was told that if he accepts the offer, then the amount will be debited from the candidate and credited to him. The grader was then free to decide on the number of mistakes to report, knowing that paper 11 would then be remunerated according to the number of mistakes reported like any other paper. The grader was also told

[19] As an alternative, we could have followed the field experiment's protocol by inserting without further instructions the money and a request for a better grade directly inside paper 11. We decided against this practice for essentially three reasons. First, we did not want to deceive lab subjects. Second, it is inconsistent with the way previous corruption experiments have been conducted in the lab. Third, to test the experimenter's scrutiny hypothesis, we need the subjects to know unambiguously that the experimenter is aware of the corruption attempt.

that if he rejects the offer, then paper 11 will not be remunerated. Nevertheless, we instructed the grader to spell-check paper 11, as well as the 9 remaining papers. At the end of the session, subjects had to fill out a short questionnaire after which they were paid in cash.

To conclude this section, note that although subjects faced the same experimental task in all three environments, the way the experiment was implemented had to be adjusted to reflect some of the basic differences between the field and the lab. In particular, in every lab experiment on corruption in the literature, subjects are told they are taking part in an experiment in which they will not be deceived (even temporarily), and they are provided with explicit instructions which explain the exact consequences of their actions and that these consequences are purely monetary. Our objective here is not to test how behavior may be affected by each adjustment made when moving from the field to the lab. Instead, we argue that, for the purpose of this paper, the protocol adjustments should be considered jointly as a reflection of the inherent differences between the lab and the field. In other words, we believe that the comparison of our lab and field experiments is informative to address one of the two questions raised in the introduction: to what extent the results of corruption experiments obtained in the lab can be extrapolated to the field?

2.4. The Experimental Treatments

We conducted four treatments in each of the three environments, that is, in the lab in Montréal, in the lab in Ouagadougou, and in the field in Ouagadougou. Each of the treatment conducted in each environment consisted of three sessions with different groups of subjects. In the Control treatment, the subjects were paid a fixed amount (called a wage hereafter) regardless of how they performed the grading task. In addition, graders in the Control treatment were offered a bribe.[20] In the lab, the wage was 250 Experimental Units (EU hereafter) and the bribe $50EU$. The conversion rate was $C\$1 = 12EU$ in Montréal and $20FCFA = 1EU$ in Ouagadougou.[21] In the field, the wage was $5,000FCFA$ and the bribe $1,000FCFA$. Three features of the design are worth noting. First, the bribe to wage ratio is the same in the lab and field experiments. Second, the payoffs for the lab and the field experiments conducted in Ouagadougou are identical. Third, because we did not want the field subjects to know they were participating in an experiment, we selected the wage

[20] The objective is to compare corrupt behavior in different environments. For a comparison of grading quality when subjects are not offered a bribe see Armantier and Boly (2009).

[21] The Franc CFA is the currency used in Burkina Faso. The conversion rate was roughly $C\$1$ for 400 $FCFA$ at the time the lab and field experiments were conducted.

so that it would not raise suspicion: $5,000FCFA$ corresponds to the amount our subjects could expect for a similar part-time job of grading.[22]

Following the economic literature on corruption, we selected treatment variables that have long been studied for their ability to promote or deter corruption: bigger bribes (e.g. Rose-Ackerman 1975, Abbink et al. 2002), higher wages (e.g. Becker and Stigler 1974, van Rijckeghem and Weder 2001, Sosa 2004), and monitoring and punishment (e.g. Rose-Ackerman 1978, Mookherjee and Png 1995, Schulze and Frank 2003). 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 and punishment. For obvious 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 in the instructions as follows. We told each grader that we would randomly pick and regrade 5 of the 20 papers he spell-checked. Only the worst of the 5 papers regraded was considered for the monetary penalty. More specifically, we calculated the difference between the number of mistakes reported in the worst regraded paper and the actual number of mistakes. 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.[23] The penalties imposed in the field were proportional.[24] Except for the risk of being penalized, the "Monitoring" treatment is identical to the Control treatment.

A total of 422 subjects participated in the experiments, with a minimum of 125 subjects per environment (see Table 1), 30 subjects per treatment (see Table 2), and 10 subjects per session. The subject pool both in Montréal and in Ouagadougou was

[22] It is difficult to compare the wages paid in Montréal and in Ouagadougou. The wages are larger in Montréal based on the official exchange rate, but those paid in Ouagadougou provide more purchasing power.

[23] Consider a perfect grader, i.e. able to identify all the mistakes in every paper. The expected payoff when this grader accepts the bribe and reports a passing grade (i.e. 15 mistakes) is $(50 - 100/4) = 25EU$ plus his wage. The expected payoff when this grader accepts the bribe and reports 11 mistakes is $(50 - 150/4) = 12.5EU$ plus his wage.

[24] 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.

comprised mostly of past or current university students, with a median age of roughly 25 (see Table 1). There were, however, marked differences between the subjects in the two locations. In particular, we can see in Table 1 that subjects in Ouagadougou were significantly more religious, less likely to be a woman, and took longer to complete the grading task.[25] In an effort to measure ability, we also calculated each subject's grading precision and improvement over the first 10 exam papers.[26] We can see in Table 1 that, on average, subjects in each of the three environments missed roughly 20% of the spelling mistakes in the first 10 exam papers and did not exhibit any significant grading improvement.

3. Experimental Results

3.1. The Decision to Accept the Bribe

We start with a brief presentation of descriptive statistics. Then, we conduct an econometric analysis to test whether treatment effects differ across environments when we control for the observable differences between the subject pools. In Table 2 (column 1) we report the frequency of bribe acceptance in each treatment and environment. Observe first that the results obtained in the lab and in the field in Ouagadougou are remarkably similar in all four treatments. In particular, roughly 50% of the Ouagadougou subjects accepted the bribe in the Control treatment. In other words, half the graders essentially refused "free money" despite the absence of explicit risks and negative externalities.[27] Table 2 also indicates that, compared

[25] The measure of religiousness was obtained from the post-experiment questionnaire in which we asked subjects how often they go to a church, a mosque or any other place of worship. This variable has 5 categories, ranging from 0 (never) to 4 (every day).

[26] Formally, $Precision_i = \frac{1}{10}\sum_{t=1}^{10}\Delta_{i,t}$, where $\Delta_{i,t} = -\left|\frac{R_{i,t}-M_t}{M_t}\right|$, $R_{i,t}$ is the number of mistakes reported by subject i for paper t, and M_t is the true number of mistakes in paper t. Note that $Precision_i \leq 0$, and grader i is considered more precise when $Precision_i$ increases toward 0. To obtain the improvement measure, we estimate for each subject i the regression model $\Delta_{i,t} = \alpha_{0,i} + \alpha_{1,i} \cdot t$ (for $t = 1, ..., 10$), and we set the variable $Improvement_i$ equal to the estimated slope $\widehat{\alpha}_{1,i}$. Grader i is then considered to have improved at the grading task when $Improvement_i$ is positive and large. Finally, note that these measures of ability are valid instruments since the graders both in the lab and in the field were unaware of the presence of corruption until they reached paper 11.

[27] This 50% rejection rate is substantially higher than in comparable lab experiments (e.g. the rejection rate is 9.4% in Frank and Schulze 2000 and 13.1% in Cameron et al. 2009). We conjecture that framing may explain this difference, as our subjects are or used to be university students and therefore, they may be less tolerant toward the type of corruption implemented in our experiment.

to the Control treatment, the probability of accepting the bribe in Ouagadougou is lower when graders are provided with a higher wage and, to a lesser extent, when their work is monitored. In contrast, lab and field subjects in Ouagadougou accepted the bribe more often when the amount proposed was larger.

The lab experiments conducted in Montréal and in Ouagadougou produced different probabilities of acceptance in most treatments. In particular, observe in Table 2 (column 1) that the rejection rate in the Control treatment was significantly smaller (at the 10% significance level) in Montréal than in Ouagadougou, as only one in three Montréal subjects rejected the bribe. Despite these differences in the level of bribe acceptance, the directions of some of the treatment effects are similar in the lab experiments conducted in Montréal and Ouagadougou. In particular, higher wages appear to curb down the bribe acceptance in both locations. In contrast with the lab experiment conducted in Ouagadougou however, proposing a higher amount to Montréal subjects does not seem to affect their decision to accept the bribe.

To test these treatment effects more formally, and to determine whether the differences between the lab experiments in Montréal and Ouagadougou may be explained by the characteristics of the subjects in each location, we estimate a probit model for a grader's decision to accept the bribe in each environment. The results reported in Table 3 suggest that the same individual characteristics shape corrupt behavior in all three environments. Specifically, we find that an older, more religious, more competent or more deliberate grader is significantly less likely to accept the bribe. These individual effects appear to be robust as they are highly significant in each of the three environments. In addition, we find that only some of these individual effects (indicated by shaded cells in Table 3) differ significantly in magnitude across the three environments. These results are remarkable as they suggest that some micro-determinants of corrupt behavior are identical in a developed and in a developing country, but also in the lab and in the field. Finally, we find that gender does not appear to influence significantly the decision to accept the bribe. This result is somewhat surprising because previous lab experiments suggest that women are less corruptible (Frank and Schulze 2000, Rivas 2008). It also contrasts with a commonly held belief among practitioners that women are less susceptible than men to accepting bribes.[28]

Regarding treatment effects, Table 3 confirms that subjects in all three environments are significantly less likely to accept the bribe when they receive a higher wage. In addition, we find the magnitude of this treatment effect to be statistically

[28]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.

indistinguishable across the three environments.[29] More specifically, we find that a 40% increase in the grader's wage leads to a 29% to 35% reduction in the probability that the bribe will be accepted. Although the magnitude is not directly comparable, the direction of this treatment effect is consistent with several lab experiments (Jacquemet 2005, Barr et al. 2009) and empirical analyses (van Rijckeghem and Weder 2001, Alt and Lassen 2003). Our results are also in line with the views of numerous practitioners and international institutions that often recommend paying civil servants above their private sector alternative as a means of deterring corruption.[30] Finding a treatment effect with the same direction and magnitude in all three environments is noteworthy, as it suggests that, at least in some dimensions, corruption experiments conducted in the lab in developed countries can be consistent with behaviors observed both in the lab and in the field in developing countries.

Table 3 also indicates that graders reject the bribe slightly more frequently when monitored, although the effect is not statistically significant in any of the three environments. This absence of treatment effect is not necessarily surprising since the type of monitoring we implemented was not aimed at catching the subjects who accept the bribe. Instead, it was designed to catch corrupt graders when they reciprocate by reporting fewer mistakes for the briber's paper.

Finally, the results in Table 3 confirm the existence of a significant difference between the lab experiments in Montréal and Ouagadougou. Indeed, we find that offering a higher bribe increases by 38% its probability of acceptance in the lab in Ouagadougou, while it appears to have no effect in the lab in Montréal. Although this result may indicate that subjects have different price elasticities in each country, we conjecture that it may reflect a pure level effect. Indeed, although the bribe is raised by the same factor in the two countries, the absolute increase in purchasing power is substantially larger in Ouagadougou.

3.2. The Number of Mistakes Reported for the Bribe Paper

Although they have no financial incentive to reciprocate, it appears that the graders who accept the bribe respond to the briber's request for reporting fewer mistakes.[31] Indeed, a comparison of column 2 and 3 in Table 2 reveals that in all but 1 of the

[29] We also find that, once we control for individual characteristics, lab subjects in the Control treatment are not significantly more corruptible in Montréal than in Ouagadougou.

[30] Singapore and Hong Kong are often presented as successful examples of such a policy. These countries are typically ranked among the least corrupt, and they are known to pay high salaries to their civil servants.

[31] Note that virtually all legal systems around the world consider that bribery starts with the acceptance of a bribe. Reciprocation is rarely a necessary condition.

12 treatments conducted (the monitoring treatment conducted in the lab in Ouagadougou) the subjects who accepted the bribe reported fewer mistakes for the bribe paper (paper 11) than the subjects who rejected the bribe. In addition, an average rejecter reported a failing grade (more than 15 mistakes) for the briber in all 12 treatments (Table 2, column 3), while the average number of mistakes reported by accepters was typically below 15 (Table 2, column 2).[32] In terms of treatment effects, no clear pattern emerges for the rejecters in each of the three environments. In contrast, compared to their respective Control treatments, accepters in each environment appear to report fewer mistakes for paper 11 in the High Wage and High Bribe treatments, and slightly more mistakes in the Monitoring treatment.

The picture becomes clearer when we control for individual effects by comparing the number of mistakes a grader reports for paper 11 and paper 7 (which both contain 20 mistakes). Indeed, a comparison of columns 4 and 5 in Table 2 suggests that, compared to a rejecter, an accepter in each of the 12 treatments is more likely to find fewer mistakes in the bribe paper than in paper 7. Note also that, regardless of the environment, the propensity for an accepter to reciprocate seems to be more pronounced in the High Wage treatment. Indeed, the proportion of accepters who do not report as many mistakes in paper 11 as in paper 7 roughly doubles when subjects receive a higher wage (Table 2, column 4). The treatment effect is even stronger for the High Bribe treatments conducted in the lab and in the field in Ouagadougou. Indeed, an accepter in Ouagadougou is nearly three times more likely than in the Control treatment to report fewer mistakes for the bribe paper than for paper 7 (Table 2, column 4). Although more modest in magnitude, the direction of the High Bribe treatment effect appears to be similar in the experiment conducted with lab subjects in Montréal. Finally, in all three environments, monitoring and punishment did not seem to affect in any predictable manner the relative number of mistakes reported for paper 7 and 11. As documented later on however, the subjects' response to monitoring and punishment was highly gender specific.

To test these treatment effects more formally, and to identify possible environment specific effect, we report in Table 4 the results of regressions in which the endogenous variable is the number of mistakes reported for the bribe paper.[33] Not surprisingly,

[32] A formal comparison with a treatment in which no bribe was offered provides no evidence that subjects who rejected the bribe were affected by the corruption attempt when grading the briber's paper (Armantier and Boly 2009).

[33] Although the bribe was offered before the subject had a chance to grade the paper, it is possible that the acceptance and grading decisions were made jointly, in which case "Accept" would be endogenous in the regressions reported in Table 4. To test this hypothesis, we estimated a simultaneous equations model in which the model in Section 3.1 (for a grader's decisions to accept the bribe) is estimated jointly with the model in Section 3.2 (for the number of mistakes reported

the parameters associated with the variables "Precision" and "Improvement" are greater than zero and highly significant in each environment, thereby indicating that subjects with better grading ability find more mistakes in Paper 11. Note also that the parameters of "Accept" and "Female*Accept" are significantly lower than zero in all three environments. This therefore confirms that subjects who accepted the bribe reciprocated by reporting fewer mistakes for paper 11, but it also reveals a gender difference in the intensity of reciprocation. Indeed, we can see in Table 4 that, while male accepters reported between 1.3 and 2.3 fewer mistakes for paper 11, female accepters omitted an additional 2.6 to 4.3 mistakes. This result is consistent with several experiments showing that women tend to reciprocate more than men (see Croson and Gneezy 2009 for a survey). In contrast, despite playing a role in explaining the decision to accept the bribe, the other individual characteristics appear to have no influence on the number of mistakes identified in the bribe paper.

While none of the treatment variables is significant for the rejecters, we find several treatment effects for the accepters (see Table 4). In particular, accepters in all three environments reported fewer mistakes when provided with a higher wage. The impact of a wage increase on corruption therefore appears to be ambiguous in our experiment: while it lowers the probability of accepting the bribe, it also seems to promote reciprocation toward the briber. The results in Table 4 also suggest that the impact of monitoring and punishment on the number of mistakes reported is gender specific. While no effect is found for men who accept the bribe, we find that female accepters did not reciprocate as much in the Monitoring treatments, as they reported between 5.1 and 6.6 more mistakes than their counterparts in the Control treatments. 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.

Observe that, up to this point, both the direction and the magnitude of the wage and the monitoring treatment effects on accepters are statistically indistinguishable in all three environments. The only significant difference we can identify is related to the impact of a higher bribe. Indeed, we can see in Table 4 that, although the parameters corresponding to the High Bribe treatment are negative in all three environments when a subject takes the bribe, they are significantly different from zero only for the lab and the field experiments conducted in Ouagadougou. In other

by the grader for paper 11). The results reported in the supplementary materials are essentially similar to those reported in Tables 3 and 4, thereby providing no support to the joint decision hypothesis. Moreover, an analysis based on the graders' decision to fail the briber (i.e. report more than 15 mistakes) produces similar results (see Armantier and Boly 2009 for such an analysis with the field data).

words, compared to the Control treatment, lab and field accepters in Ouagadougou report fewer mistakes for the briber, while increasing the bribe does not influence significantly how accepters grade the bribe paper in the lab in Montréal.

4. Discussion

As argued by several international institutions, corruption is one of the most detrimental factors currently afflicting the economies of developing countries. Because naturally occurring data on corruption are scarce and do not vary along certain desired dimensions (e.g. under different wage or monitoring structures), development 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 lab experiments conducted in developed countries. This approach could become increasingly popular in development economics as it enables the analyst to overcome the unobservability of corrupt behavior in a controlled and cost-effective environment. First, however, it is essential to establish whether or not the insights gained in these lab experiments can be extrapolated to developing countries.

In an attempt to address this issue, we conducted a corruption experiment in three different environments to compare the behavior of i) lab subjects in a developed and in a developing country, and ii) lab and field subjects in a developing country. In short, we hired subjects to grade exam papers, one of which came with a bribe offer and a demand for a better grade. We then recorded the frequency with which the bribe was accepted and the number of mistakes the subjects reported for the bribe paper. The experiment consisted of four different treatments where we varied successively i) the amount of the bribe, ii) the wage paid to graders, and iii) the level of monitoring and punishment. The experiment was conducted in three environments: the lab in Montréal (Canada), the lab in Ouagadougou (Burkina Faso), and the field in Ouagadougou. The key difference between the lab and the controlled field experiment is that subjects in the field acted without knowing they were participating in an experiment.

The results obtained in the lab and in the field in Ouagadougou are remarkably similar. In particular, the probability that a subject accepts the bribe in any given treatment is virtually identical in both environments. The lab experiments conducted in Ouagadougou and in Montréal produced slightly different results. However, once we control for observable differences between the subjects in each country (e.g. gender, age, ability), we find that the direction and the magnitude of several treatment effects are statistically indistinguishable in the lab experiments conducted in Mon-

tréal and in Ouagadougou. In particular, we find that paying higher wages has an ambiguous effect on corruption: it reduces the probability that a bribe will be accepted, but it promotes reciprocation toward the briber among those who accept. Likewise, regardless of the environment, we find that women who accept the bribe respond to monitoring and punishment by reporting more mistakes for the briber. The outcomes of the lab experiments, however, differ in some dimensions between the two locations. Indeed, doubling the amount of the bribe proposed to the grader has no effect in Montréal, while it promotes bribe taking and reciprocation in Ouagadougou. As argued in Section 3.1, this result may reflect a stake effect, as the magnitude of the bribe increase was smaller in real terms in Montréal. Finally, our results suggest that some micro-determinants of corruption may be universal. Indeed, we find that, regardless of the environment, corrupt behavior in our experiment can be linked to the same individual characteristics (e.g. age, religiousness, ability). Our results therefore support Cameron et al. (2009) who find that propensities to accept bribe are comparable across countries and cultures.

In addition to providing some insights into the universality of corrupt behavior, our results provide some elements to answer the two questions raised in the introduction. First, despite differences in the levels of corruption, we find that, once we control for the subjects' observable characteristics, the lab experiments conducted in Ouagadougou and Montréal produced several treatment effects that were statistically indistinguishable both in direction and magnitude. In other words, we do not find evidence of major intrinsic differences between developed and developing countries, in the sense that several treatment variables seem to affect equally the behaviors of two lab subjects with similar characteristics. Second, the lab and the field experiments conducted in Ouagadougou produced almost identical results, thereby suggesting that the propensity to accept bribes does not change fundamentally when subjects know their actions are being observed. This therefore suggests that the experimenter's scrutiny may not be a major concern for lab experiments on corruption, as it does not appear to generate any relevant biases.

Our paper therefore provides some support to the external validity of corruption experiments, but it should not be considered a definitive answer to the question. Indeed, although our subjects faced the same task in all three environments, the implementation of the experiment had to be adjusted when moving from the field to the lab because of the inherent differences between the two environments. Subsequent studies will be necessary to identify more precisely possible differences in behavior between the lab and the field. In addition, the experiment we conducted cannot be considered to fully capture corruption in the field. In particular, our anonymous one-shot game contrasts with the field where corruption typically involves repeated

interactions between parties who can identify each other and whose decisions may have life long consequences. More generally, the very specific task, stakes, and subject pool in our experiment may not be representative of corruption in the field. Nevertheless, our results may be considered encouraging as they suggest that lab experiments, and in particular lab experiments conducted in developed countries, can be informative to understand and combat corruption in developing countries.

References

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

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

Alatas, V., L. Cameron, A. Chaudhuri, N. Erkal, and L. Gangadharan (2009a): "Gender and Corruption: Insights from an Experimental Analysis," *Southern Economic Journal*, Vol. 75, Issue 3, 663–680.

Alatas, V., L. Cameron, A. Chaudhuri, N. Erkal, and L. Gangadharan (2009b): "Subject Pool Effects in a Corruption Experiment: A Comparison of Indonesian Public Servants and Indonesian Students," *Experimental Economics*, Vol.12, Issue 1, 113-132.

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.

Armantier O. and A. Boly (2009): "A Controlled Field Experiment on Corruption," Working paper, Université de Montréal.

Barr, A., M. Lindelow, and P. Serneels (2009): "Corruption in Public Service Delivery: An Experimental Analysis," *Journal of Economic Behavior and Organization*, 72(1), 225-239.

Barr, A. and D. Serra (2009): "The Effects of Externalities and Framing on Bribery in a Petty Corruption Experiment," *Experimental Economics*, Vol 12 (4), 488-503.

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.

Cameron, L., A. Chaudhuri, N. Erkal, and L. Gangadharan (2009): "Propensities to Engage in and Punish Corrupt Behavior: Experimental Evidence from Australia, India, Indonesia and Singapore," *Journal of Public Economics*, Vol. 93, Issue 7-8, 843-851.

Castillo, M., R. Petrie, M. Torero, and A. Viceisza (2009): "Corruption in the Mail Sector: A Field Experiment in Peru," Working Paper, Georgia Institute of Technology.

Croson, R. and U. Gneezy (2009): "Gender Differences in Preferences," *Journal of Economic Literature*, 47(2): 448–74.

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.

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.

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

Hey, J. D. (1998): "Experimental Economics and Deception: A Comment," *Journal of Economic Psychology*, 19, 397-401.

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

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

La Porta, R., F. Lopez-de Silanes, A. Shleifer, and R. Vishny (1999): "The Quality of Government," *Journal of Law, Economics, and Organization*, 15(1), 222-279.

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

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

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

Rivas, F. (2008): "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.

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

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

Treisman, D. (2000): "The Causes of Corruption: A Cross-National Study," *Journal of Public Economics Buletin*, 76, 399-457.

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.

4.1. Appendix A: The Candidates (Montréal, Canada)

Subjects, called "candidates", were recruited to type a "*dictée*." At the beginning of the typing session, each candidate was assigned to an isolated computer. Instructions were then read aloud, followed by questions. We explained what would and what would not constitute a mistake. The subjects were also informed that, at the end of the 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 several subjects called "graders." Finally, we explained that a candidate's payment would depend in part on the number of mistakes reported by the experimenter or the graders. 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." We explained to the candidates that if they accepted to offer a bribe, 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 agreed to offer a bribe, we would not necessarily send the message and the bribe to the graders.[34]

To keep as much control as possible over the experiment, we deliberately left the candidates' instructions partly ambiguous.[35] In particular, we did not explain how we would select the papers to be graded by experimental subjects. Likewise, we did

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

[35]According with standard lab practices, however, we did not explicit mislead the candidates. We simply withheld some information.

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 to the candidate's payoff when a grader accepts or rejects the bribe. The candidates were told they would received several payments: $C\$20$ payable immediately after the conclusion of the typing session and three additional amounts paid after the completion of the grading sessions in each of the three environments. The candidates knew that each of the additional amounts could vary between $C\$20$ and $C\$60$, depending on the average number of mistakes reported by the graders, and, when relevant, on the number of graders who accepted and rejected the bribe offer.

We conducted two typing sessions in Montréal at CIRANO's Bell Laboratory for Experimental Economics. Each session lasted roughly an hour and included respectively 11 and 12 subjects. All 23 subjects agreed to send a bribe to the graders. On average, the candidates received a total payment of $C\$96.48$, with a maximum of $C\$171.41$ and a minimum of $C\$80.00$.

Table 1
Characteristics of Subject Pools
Averages and Standard Deviations (in Parenthesis)

	N	Age	Female	Religiousness	Time (in Min)	Precision	Improvement
Lab Montréal	125	26.264 (6.324)	0.408 (0.493)	0.832 (1.063)	100.208 (17.303)	-0.190 (0.100)	-0.068 (1.493)
Lab Ouagadougou	133	24.872 (3.014)	0.195 (0.398)	2.496 (1.241)	164.707 (34.905)	-0.184 (0.101)	0.358 (1.591)
Field Ouagadougou	164	24.933 (2.381)	0.152 (0.361)	2.671 (1.254)	140.396 (26.674)	-0.231 (0.121)	-0.045 (1.350)

A cell shaded in dark (respectively, light) gray indicates that the characteristic has a distribution significantly different at the 5% level (respectively, 10% level) than in the lab in Ouagadougou according to Wilcoxon rank-sum test

Table 2
Descriptive Statistics

			1	2	3	4	5
	Treatment	Number of Subjects	% of Graders who Accept the Bribe	Average # of Mistakes Reported for Paper 11		% of Graders who Report Fewer Mistakes for Paper 11 than for Paper 7	
				Accepters	Rejecters	Accepters	Rejecters
Lab Montréal	Control	30	0.667	14.950	16.700	0.350	0.300
	High Bribe	32	0.656	14.809	15.636	0.429	0.273
	High Wage	31	0.484	13.933	16.250	0.600	0.438
	Monitoring	32	0.656	15.000	15.455	0.333	0.455
Lab Ouagadougou	Control	33	0.485	16.375	17.294	0.250	0.235
	High Bribe	33	0.667	15.227	16.545	0.727	0.364
	High Wage	33	0.364	14.916	17.476	0.571	0.381
	Monitoring	34	0.412	17.500	16.450	0.286	0.250
Field Ouagadougou	Control	36	0.500	14.722	16.444	0.278	0.222
	High Bribe	45	0.689	13.097	15.429	0.710	0.357
	High Wage	39	0.359	12.929	15.680	0.643	0.240
	Monitoring	44	0.409	15.667	15.923	0.333	0.269

A cell shaded in dark (respectively, light) gray indicates that the variable has a distribution significantly different at the 5% level (respectively, 10% level) than in the same treatment in the lab in Ouagadougou according to Wilcoxon rank-sum test

Table 3
The Decision to Accept the Bribe

Variable	Environment					
	Lab Montréal		Lab Ouagadougou		Field Ouagadougou	
	Parameter	Marginal Effect	Parameter	Marginal Effect	Parameter	Marginal Effect
Constant	0.882** (0.387)	—	0.044 (0.229)	—	0.006 (0.233)	—
Female	-0.484 (0.316)	-0.126 (0.083)	0.170 (0.322)	0.072 (0.137)	-0.572 (0.395)	-0.196 (0.135)
Age	-0.097*** (0.027)	-0.038 (0.010)	-0.083** (0.036)	-0.035 (0.015)	-0.368*** (0.048)	-0.149 (0.020)
Time	-0.032*** (0.007)	-0.011 (0.003)	-0.009** (0.004)	-0.005 (0.002)	-0.019** (0.009)	-0.009 (0.003)
Religiousness	-0.708*** (0.155)	-0.192 (0.041)	-0.306*** (0.110)	-0.116 (0.043)	-0.479*** (0.102)	-0.161 (0.034)
Precision	-0.491** (0.192)	-0.178 (0.069)	-0.674*** (0.178)	-0.250 (0.065)	-0.249** (0.122)	-0.120 (0.059)
Improvement	-0.523*** (0.126)	-0.191 (0.046)	-0.750*** (0.149)	-0.258 (0.052)	-0.104 (0.084)	-0.040 (0.032)
High bribe Treatment	-0.239 (0.420)	-0.084 (0.147)	1.169*** (0.395)	0.384 (0.131)	0.732** (0.331)	0.293 (0.132)
High Wage Treatment	-1.023*** (0.406)	-0.352 (0.141)	-0.863** (0.382)	-0.314 (0.139)	-0.741** (0.327)	-0.295 (0.130)
Monitoring Treatment	-0.343 (0.418)	-0.121 (0.146)	-0.345 (0.351)	-0.139 (0.141)	-0.251 (0.320)	-0.099 (0.125)
Ln L	-50.698		-65.610		-76.109	
N	125		133		164	

In each environment, the subject of reference is a man in the Control treatment. The standard deviations are robust and clustered at the session level. Significance: * = 10%, ** =5%, *** = 1%. A cell shaded in dark (respectively, light) gray indicates that the estimated parameter is significantly different at the 5% level (respectively, 10% level) than in the lab in Ouagadougou according to a Likelihood ratio test.

Table 4
Grade Reported for Paper 11

Variable	Environment		
	Lab Montréal	Lab Ouagadougou	Field Ouagadougou
Constant	17.172*** (0.816)	17.382*** (0.371)	16.425*** (0.406)
Female	-0.882 (0.748)	-0.292 (0.500)	0.082 (0.821)
Age	0.008 (0.047)	0.037 (0.065)	-0.049 (0.100)
Time	-0.031 (0.023)	-0.010 (0.014)	-0.018 (0.012)
Religiousness	-0.384 (0.276)	0.002 (0.158)	-0.375 (0.212)
Precision	1.444*** (0.323)	0.630*** (0.220)	0.962*** (0.251)
Improvement	0.837*** (0.231)	0.365** (0.114)	0.469** (0.186)
Accept	-2.294** (0.962)	-1.299** (0.531)	-1.821** (0.752)
Female * Accept	-3.554*** (0.970)	-2.580*** (0.755)	-4.323*** (0.841)
Age * Accept	0.111 (0.073)	0.095 (0.082)	-0.262 (0.138)
Time * Accept	0.018* (0.010)	-0.005 (0.009)	0.012 (0.014)
Religiousness * Accept	0.463 (0.431)	0.264 (0.238)	0.488 (0.288)
Precision * Accept	-0.022 (0.468)	-0.438 (0.309)	0.075 (0.261)
Improvement * Accept	0.395 (0.321)	0.256 (0.210)	0.150 (0.245)
Reject * High bribe Treatment	-0.146 (0.921)	-0.833 (0.613)	-0.944 (0.802)
Reject * “High Wage Treatment	-0.083 (0.886)	0.148 (0.522)	-0.595 (0.570)
Reject * Monitoring Treatment	1.258 (0.853)	-0.560 (0.509)	0.162 (0.682)
Accept * High bribe Treatment	-0.657 (0.498)	-1.304** (0.528)	-1.610** (0.646)
Accept * High Wage Treatment	-1.909** (0.776)	-1.424** (0.626)	-1.878** (0.763)
Accept * Monitoring Treatment	-0.817 (0.711)	0.530 (0.600)	-0.554 (0.860)
Female * Accept * Monitoring Treatment	6.562*** (1.502)	5.128*** (1.341)	4.946*** (1.322)
Adjusted R-squared	0.589	0.486	0.468
N	125	133	164

In each environment, the subject of reference is a man in the Control treatment. The standard deviations are robust and clustered at the session level. Significance: * = 10%, ** =5%, *** = 1%. A cell shaded in dark (respectively, light) gray indicates that the estimated parameter is significantly different at the 5% level (respectively, 10% level) than in the lab in Ouagadougou according to a Wald test.