Certainty and Severity of Punishment in Crime and Corruption Deterrence: An Experimental Study

Lin Jing* Yohanes E. Riyanto†

November 29, 2021

Abstract

We investigate experimentally the (relative) effectiveness of certainty and severity of punishment in deterring crime and corruption simultaneously in a game developed by Ortner and Chassang (2018). The experimental design features two different policy regimes: $\mathcal{HP}$ with a high certainty and low severity of punishment and $\mathcal{LP}$ the other way around. Within each regime, we examine whether there is a real deterrent effect by increasing the certainty or severity of punishment, and which one delivers a greater impact if there is any. We show that, in the regime $\mathcal{LP}$, neither increasing certainty nor increasing the severity of punishment deters crime or corruption effectively. In contrast, both increasing the certainty and severity of punishment significantly deter crime and corruption in regime $\mathcal{HP}$ where the certainty of punishment is high enough, and an increase in the certainty delivers a greater deterrent effect than that in the severity. In addition, we document the presence of the Cobra Effect in the regime $\mathcal{LP}$ when we intend to deter corruption by increasing the expected wage of the monitor. Last but not least, we explore the changes in extensive and intensive margins of crime and corruption, and we find a difference in celerity of deterrent effect between increasing the certainty and severity of punishment in regime $\mathcal{HP}$, specifically, subjects are promptly responsive to changes in the certainty of punishment while they are inertial to changes in the severity of punishment.

Keywords: Crime Deterrence, Corruption Deterrence, Asymmetric Information, Cobra Effect, Extensive Margin, Intensive Margin.

JEL Codes: D73, D82, K14, K42

*Division of Economics, School of Social Sciences, Nanyang Technological University, Singapore, 639818. Email: jing0014@e.ntu.edu.sg

†Division of Economics, School of Social Sciences, Nanyang Technological University, Singapore, 639818. Email: yeriyanto@ntu.edu.sg

Contents

1 Introduction

Since the seminal work by Becker (1968) , the certainty and severity of punishment have been the most commonly examined and used instruments to deter crime and corruption, and the author predicts that certainty of punishment would have a larger deterrent effect than severity according to the model given the fact that punishment is costly.

Some studies support Becker's prediction and show that certainty of punishment does have a larger deterrent effect against crime or corruption than severity of punishment both empiricailly (for example, see Chauncey (1975), Witte (1980), and Grogger (1991)) and experimentally (for example, see Barr, Lindelow, and Serneels (2009) and Mungan (2017)). However, some other studies show the opposite: an increase in the severity of punishment is more effective in deterring crime/corruption than an increase in the certainty of punishment (for example, Friesen (2012) and Banerjee and Mitra (2018)). Therefore, this issue requires more explorations in details to make the jigsaw more complete.

Given the fact that complete elimination of crime and corruption is impossible due to the extremely high marginal cost of increasing the certainty of punishment when it is already very high, the issue is more about how to control the crime and corruption at an acceptable level or an desired level[1]. Nagin (1998) points out that "for policy makers the issue is not whether the criminal justice system in its totality prevents crime but whether a specific policy, grafted onto the existing structure, will materially add to the preventive effect."

Therefore, between the two mostly discussed and examined policy instruments: certainty versus severity of punishment, the issue is more about which one the authority should manipulate to deter crime and corruption when it decides to do so at the current state. Furthermore, how does it work underneath? How do the extensive and intensive margins change when there is an deterrent effect? Specifically, it is important to investigate whether the policy intervention deters crime (corruption) by decreasing the population of those who commit a crime (corruption) or by decreasing the intensity of those who have committed a crime (corruption).

Empirical studies usually have several drawbacks in addressing the issue. First, they do not have an accurate measure of the deterrent power of an increase in certainty and severity of punishment, thus the result they have obtained might due to an incomparable increase in these two policy instruments rather than the difference in these two instruments themselves. Second, it is usually difficult for these studies to separate the effect of certainty and that of severity on deterring crime and/or corruption from each

[1] There are some concerns on the welfare effect of corruption deterrence since it might consume too many resources and thus the costs exceed the benefits. In addtion, the relationship between corruption and economic growth is still unclear. For example, Ang (2020) points out that China has achieved a fast economic growth while corruption is prevalence since China's Reform and Opening in 1978, and the US in late 19^{th} century experienced a very similar process. Therefore, it is very important to make the level of corruption under control. The authority should have effective policy instruments at hand such that they can effectively achieve the desired corruption level by manipulating these policy instruments.

other.

Experimental studies solve this problem in a better way with controlled lab experiments such that all the potential confounding factors are controlled except the parameter of interest. With controlled lab experiments, one can study the deterrent effect of increasing the certainty versus the severity of punishment separately in different treatments. However, current experimental studies do not address the above issue in the right way. For example, Armantier and Boly (2011) investigate the deterrent effect of increasing the wage against corruption, however, they do not have the element of detection and thus punishment in the treatments, therefore, any corruption deterrence is restricted to take effect via the official's reciprocity[2]. Banerjee and Mitra (2018) experimentally test the effectiveness of corruption deterrence between two commonly compared policy designs with the same expected payoff: a low probability of detection with high punishment versus a high probability of detection with low punishment. Furthermore, they also compare these two policy designs against a control treatment where the chance of corruption being detected and thus punished is zero, therefore, the comparison between the control and the above two policy designs might be problematic since there is not only a quantitative change but also a qualitative change in the punishment against corruption.

With a simple principle-agent model developed by Ortner and Chassang (2018), we examine the issue experimentally and provide some novel results. In the model, the agent decides whether to commit a crime or not and the benefit of committing a crime is π_A, then the monitor fully observes this and reports the decision of the agent to the authority. Once reported guilty, the agent will receive a punishment of k. The agent can make a bribery offer to the monitor in exchange for a report of innocence. However, any false report made by the monitor has a chance of p being detected by the authority and the monitor's expected wage $\mathbb{E}(w)$ is deprived. Therefore, the idea is to design a wage structure such that the bribery offer is not enough to compensate the monitor's expected cost of accepting it, and thus, both crime and corruption are deterred at the same time.

In order to investigate the deterrent effect of crime and corruption with certainty and severity of punishment as policy instruments, we separate our experimental treatments into two policy regimes: one with high probability of detection p and low expected wage $\mathbb{E}(W)$ which is denoted as regime $\mathcal{HP}$ and the other with low probability of detection and high expected wage which is denoted as regime $\mathcal{LP}$, while maintaining the expected (opportunity) cost of corruption the same. And then within each regime, we either increase the probability of detection or the expected wage to see which policy intervention has a larger deterrent effect against crime and corruption if there is any. We conduct all our experimental treatments with a between-subject design, therefore, we have six treatments in total and three treatments in each regime.

[2] Armantier and Boly (2011) report that a higher wage does not have an clear deterrent effect against corruption, specifically, it decreases the acceptance of bribe offer but it increases the official's corrupted behavior, namely, the official (grader in their experiment) gives more pass to the agent (exam takers). This is clearly a result led by the absent of detection and thus punishment of misreport.

This paper contributes to the understanding of crime and corruption deterrence in several significant ways: we do not only compare the relative effectiveness of different policy interventions in these two regimes, but we also try to answer how it works and where the effect should be attributed to.

First, we are able to investigate the deterrent effect against crime and corruption at the same time[3] by employing the model developed by Ortner and Chassang (2018). In this model, the rate of crime apprehension and conviction is 100% by assumption, thus, one would expected that a higher deterrent effect against corruption will result in a higher deterrent effect against crime immediately. However, the results show that the policy interventions that have an significantly evident deterrent effect against corruption have a significant but smaller deterrent effect against crime. This is largely due to the fact that the policy interventions are mainly on the monitor's side which would have a direct impact on corruption deterrence but an indirect impact on crime deterrence. The result suggests that the deterrent effect will diminish via transmission. A more direct policy intervention against crime (corruption) would yield a larger deterrent effect against crime (corruption) than that against corruption (crime).

Second, we are able to examine the relative effectiveness of certainty and severity of punishment in deterring crime and corruption within each policy regime, and we find significant differences between regime $\mathcal{HP}$ and $\mathcal{LP}$. In the $\mathcal{HP}$ regime, both increasing the probability of detection and the expected wage would significantly reduce the corruption rate although the magnitude is larger when we increase the probability of detection. In contrast, we fail to find any significant deterrent effect with either policy interventions in the $\mathcal{LP}$ regime. This contributes to the understanding of the whole picture that it requires the certainty of punishment (probability of detection) to be high enough so that either policy interventions would be able to deliver significant deterrent effect. In the domain of crime deterrence, the difference between these two regimes retains the same pattern though the magnitude is smaller.

In addition, we also document a potential Cobra Effect[4] in the $\mathcal{LP}$ regime when we increase the expected wage aiming to deter corruption. The famous Cobra Effect is the most typical representation of the perverse incentive effect where the provided incentive leads to the opposite of the intended outcome. In our experiment, when we increase the expected wage of the monitor, the expected opportunity cost of corruption increases which should lead to a lower corruption rate. However, our data shows that, in the $\mathcal{LP}$ regime, the corruption rate becomes higher when we increase the expected wage. Furthermore, a closer look at the data shows that this is due to an unexpected adjustment from the agent's side. When anticipating the expected wage of the monitor becoming higher, the agent responses by providing a higher

[3] To the best of our knowledge, exsiting experimental studies only investigate deterrent effect of certainty versus severity of punishment against either crime or corruption. For example, Friesen (2012) focuses on the issue of deterring crime, while Banerjee (2016); Banerjee and Mitra (2018) mainly address the issue of corruption deterrence.

[4] The Cobra Effect refers to the case where the provided incentive to address a problem actually makes it worse. The name is after an anecdote happened in India: The government offered a bounty for cobra tails in order to reduce the number of cobra in the town. However, the locals started to breed cobras to claim more bounties in the end which led to an increase in the number of cobra. The story is documented by Lucas and Fuller (2018), and is also discussed in many scenarios. For example, Lueck and Michael (2003) docmuments an act that intends to protect the habitat of endangered species leads to a decrease of the area of the habitats. Bajo-Buenestado and Borrella-Mas (2019) show that the effect of tax change on firms beyond borders becomes more prevalent after the authority discourages the residents do not cross the borders to buy.

bribery offer which dominates the increase in the expected wage of the monitor, and jointly lead to a higher corruption rate. Nonetheless, we do not observe a similar Cobra effect in the $\mathcal{HP}$ regime, where an increase in the expected wage does not induce the agent to provide a higher bribery offer, and on the contrary, the offer becomes lower. We would like to suggest that this is due to the high detection probability in the $\mathcal{HP}$ regime which translates the increase in the expected wage to a real deterrent power in effect.

Last but not least, through analyses on the changes of extensive and intensive margins, we are able to answer where the deterrent effect is attributed to (if there is any): an decrease in the number of people who choose to commit a crime/corruption, or an decrease in the frequency of crime/corruption conditional on one has committed a crime/corruption, or both. The results show a significant contrast between the $\mathcal{HP}$ and $\mathcal{LP}$ regimes. In the $\mathcal{HP}$ regime, the extensive margin of corruption decreases significantly and immediately when we increase the probability of detection, and so does the intensive margin. This suggests that, when we increase the probability of detection in the $\mathcal{HP}$ regime, the number of subjects who accepts a bribery offer decreases and those who are corrupted accept the offer less frequently. However, when we increase the expected wage in the $\mathcal{HP}$ regime, the extensive margin of corruption remains unchanged in the beginning periods and then decreases gradually over time. It would be very close to the extensive margin of corruption by increasing the probability of detection in the ending periods. This suggests that the deterrent effect against corruption by increasing the probability of detection takes effect immediately, while that by increasing the expected wage won't have any immediate effect. This demonstrates the difference in the celerity of deterrent effect between different policy interventions.

However, in the $\mathcal{LP}$ regime, there is no significant change in neither the extensive margin nor the intensive margin when we increase either the probability of detection or the expected wage. In addition, the intensive margin of corruption in the $\mathcal{LP}$ regime increases due to an increase in the intensive margin of bribery offers, which demonstrates the underlying mechanism of the Cobra Effect.

In terms of the deterrent effect against crime, there is no significant difference in the intensive margin across treatments in either the $\mathcal{HP}$ regime or the $\mathcal{LP}$ regime, however, the pattern for extensive margin remains the same as it is against corruption. In the $\mathcal{HP}$ regime, extensive margin of crime significantly decreases when the probability of detection increases and thus the crime deterrence takes effect immediately and significantly. Similar with corruption deterrence, the extensive margin of crime by increasing the expected wage remains unchanged in the beginning, and it decreases over time which ends up being close to the extensive margin by increasing the probability of detection.

The remainder of this paper is organized as follows. section 2 discusses related literature, section 3 presents the basic model that describes the stage game that we use in our experiments, and section 4 shows the experimental design and our hypotheses, followed by results in presented in section 5, and section 6 concludes.

2 Related Literature

Becker (1968) build the foundation of modern economic analysis on deterring criminal activities in an efficient way. He set up an economic framework trying to minimize the total social loss from criminal activities by finding the optimal resource allocations, and one of his important findings is that, with the assumption that punishment is costly which is true most of the time, a change in conviction rate should have a larger impact on criminal activities than a change in the magnitude of punishment .

Given the detection probability is usually far less than unity, later Becker and Stigler (1974) ask a question that "How can corrupt enforcement be discouraged when detection is uncertain?" And their suggestion is to "raise the salaries of enforcers above what they could get elsewhere, by an amount that is inversely related to the probability of detection, and directly related to the size of bribers and other benefits from malfeasance." By comparing benefits of malfeasance against the present value of all future streams of salaries as well as pensions after retirement, they claim that "Malfeasance can be eliminated, therefore, even when the probability of detection is quite low." Niehaus and Sukhtankar (2013) actually confirms this statement and show that a higher daily wage significantly deters theft behavior in piece-rate projects. On the other hand, Borcan, Lindahl, and Mitrut (2014) show that an unexpected wage cut in the public sector employees will result in an increased level of corruption.

Several empirical studies support Becker's prediction that a change in conviction rate should have a larger impact on criminal activities than a change in the magnitude of punishment. Chauncey (1975) examines the effectiveness of several different institutional designs on deterring skyjacking, and the author reports that a higher certainty of punishment leads to a lower crime incidents (rate), while a higher severity of punishment does not.

Using criminal records of post-released criminals, Witte (1980) show that an increase in both the certainty and severity of punishment can effectively deter criminal activities, with a generally larger deterrent effect by the increases in certainty of punishment than the same percentage increase in severity. .

With a large longitudinal dataset on official criminal records, Grogger (1991) shows that, at the individual level, increased certainty of punishment delivers a greater deterrence power against crime than increased severity of punishment whose effect is generally insignificant. One issue of this study is that, the increase in the probability of conviction might not be comparable in magnitude with the increase in the severity of punishment[5].

[5] In the paper, the measure of certainty of punishment is in percentages, while the measure of severity of punishment is the length of sentences, and the change might not be comparable with each other in terms of expected punishment. For example, suppose the certainty is 40%, the severity is 10 years, then a 1% increase in the certainty leads to an increase in

In a review paper, Doob and Webster (2003) conclude that "sentence severity has no effect on the level of crime". On one hand, they examine many studies that fail to support the hypothesis that variations in severity have an deterrent effect against crime with a particular focus on the studies investigating the overall deterrent effect of structural changes in sentencing laws. On the other hand, they also examine the studies that find some evidence on deterrent effect of increased severity of punishment and they point out that these studies do not arrive at their conclusions in a credible way due to certain serious methodological, statistical, or conceptual problems[6]. Later, Chalfin and McCrary (2017) also summarize that "there is far less evidence that crime responds to the severity of criminal sanctions."

Barr et al. (2009) also support this point of view in an experimental study. They examine the corruption behavior in service delivery. There is no bribery offers to be exchanged in the game, so the notion of corruption is mainly on the individual level (morally corrupted). They show that the corruption rate is lower when the detection probability is high (i.e., the service provider's effort is easier to be observed in their context), while a higher wage of the service provider has little effect on preventing corruption[7].

However, some experimental studies demonstrate the opposite: severity of punishment has a greater deterrent effect against crime or corruption than certainty of punishment. Friesen (2012) reports that increasing the severity of punishment is more effective in deterring crime than an equivalent increase in the certainty of punishment. There are three key parameters in the experimental design: the benefit of a crime, probability of detection, and the fine. The experiment lasts for 30 periods and they vary these three parameters from period to period which we consider to be problematic. Such a design is not clean enough to address the issue. In addition, the results are partially driven by the majority of their subjects being risk averse, and the risk preference is a significant predictor of crime rate in their regressions results.

Using a harassment bribery game, Banerjee and Mitra (2018) experimentally show that a low probability of detection with high fines reduces both the amount and the likelihood of bribe demand, while a high probability with low fines has no effect on bribe demand. However, this paper does not address the issue that we raise in this paper. As Nagin (1998) points out that "for policy makers the issue is not whether the criminal justice system in its totality prevents crime but whether a specific policy, grafted onto the existing structure, will materially add to the preventive effect", the question we focus here is more about which policy instrument — the certainty or severity of punishment — that the authority should manipulate to deter crime and corruption when it decides to do so at a particular state of policy regime. Banerjee and Mitra (2018) do have a control treatment, but there is no possibility of detection and thus no possibility of punishment in the control treatment, so the control treatment barely resembles

sentence of 0.1 year in expectation, while a 1 year increase in the severity leads to an increase in sentence of 0.4 year in expectation.

[6]For example,Kessler and Levitt (1999) suggest that sentence severity produces deterrent effect against crime, however, Doob and Webster (2003) point out that the study suffers from data selection problem. Specifically, Kessler and Levitt (1999) only uses odd-numbered years in their data collection, and no explanation is provided.

[7]The main reason (of the ineffectiveness of higher wages in this study) is that the wage does not stand as an opportunity cost of corruption. The service provider only lose the benefits they keep during the service-providing process if the corruption behavior is detected.

most empirical scenario. Therefore, the comparison between the other two treatments falls into two commonly compared paradigms: low probability of detection with high fines versus high probability of detection with low fines. In addition, they focus on the cases where the officials demand bribes from the agents, while we focus on the cases where the agents tries to bribe the official in exchange for a favorable result and the official cannot demand any bribe.

3 Theoretical Framework

In this paper, we examine the deterrence of crime and corruption simultaneously by adopting a theoretical framework developed by Ortner and Chassang (2018). Their major objective is to show that, in a principal-agent model, the government can deter corruption with a lower expected wage cost than that would have been paid to the monitor. They achieved this by introducing random distributed wages among the monitors and asymmetric information on the realization of wages between the monitor and the criminal agent. Among all the wage distributions that renders a non-positive payoff for the agent from committing a crime, the optimal one is the one with the lowest expected wage cost which renders a zero payoff from committing a crime (thus just fully deters criminal activities). As no one would commit a crime, no one would make any bribe offer in such a game, therefore, corruption is also deterred.

There are three players in the game: the principal, the agent, and the monitor. The principal in the game is a passive role that only used to illustrate the story[8]. The agent (he) first makes a choice of whether to commit a crime $c \in \{0, 1\}$ with $c = 1$ implying the agent choosing to commit a crime. He gets a constant positive payoff $\pi_A > 0$ if he chooses $c = 1$ and zero otherwise.

The criminal activity also renders a cost to the principal, and the principal does not observe the agent's action choice so he hires a monitor to observe and make a report $m \in \{0, 1\}$ on the agent's choice with $m = 1$ implying the monitor reports that the agent has chosen $c = 1$. The principal makes a judicial judgement according to this report m and imposes a punishment $k > \pi_A$ on the agent if and only if $m = 1$. The monitor can make any report at her will including a false report where $m \neq c$.

The principal also performs an audit on the monitor's report from time to time so that there is a chance $p \in (0, 1)$ that a false report can be detected. This makes the report partially verifiable. Ortner and Chassang (2018) argues that partial verifiability can happen in several different ways: for example, "accounting discrepancies, random rechecks, or tips from informed parties", as well as observable consequences from criminal activities.

[8] In our study, we focus on investigating: First, which policy design is more effective in deterring crime and corruption when they have the same deterrence power (which is defined clearly later on); Second, which policy intervention (raising detection rate vs the monitor's expected wage) is more effective in deterring crime and corruption. We actually do not care about the principal's expected payoff in our study since that is another issue: the principal's optimization problem given certain constraints.

The principal needs to pay the monitor a wage according to a statistical distribution with c.d.f. F_w. Whenever a misreport is detected, the monitor is fired by the principal and thus loses her wage. The principal cannot punish the monitor beyond her wage because of limited liability.

Corruption becomes an issue when the agent can make a monetary transfer $\tau > 0$ to the monitor. As long as the monitor accepts the transfer, and so corruption takes place, the monitor also agrees to destroy any criminal evidence that might be used to charge against the agent by reporting $m = 0$ although $c = 1$. Therefore, as long as the monitor reports $m = 0$, the principal can not punish the corresponding criminal agent due to the lack of evidence even if the principal knows that the agent has engaged in criminal activities later on.

The timing of moves in the game is as follows:

1. The principal draws a wage w randomly from a distribution with c.d.f. $F(w)$. The monitor gets to know the wage w but the agent does not. The c.d.f $F(w)$ is common knowledge[9].
2. The agent decides whether to commit a crime or not $c \in \{0, 1\}$.
3. The agent then decides whether or not to make a take-it-or-leave-it offer $\tau > 0$ to the monitor in exchange for the monitor reporting $m = 0$. Perfect commitment is assumed here so that the monitor will report $m = 0$ as long as he accepts the bribe offer made by the agent.
4. When there is no bribe offer made by the agent or the monitor chooses to reject the offer, the monitor chooses the report m that maximizes her payoff.

Figure 1 shows the structure of the game.

In such a game, when the agent chooses not to commit a crime ($c = 0$), his payoff is $\mathbb{E}(P_A) = 0$ no matter what the monitor reports, therefore, the monitor cannot make a credible threat of sending a report $m = 1$ when $c = 0$.

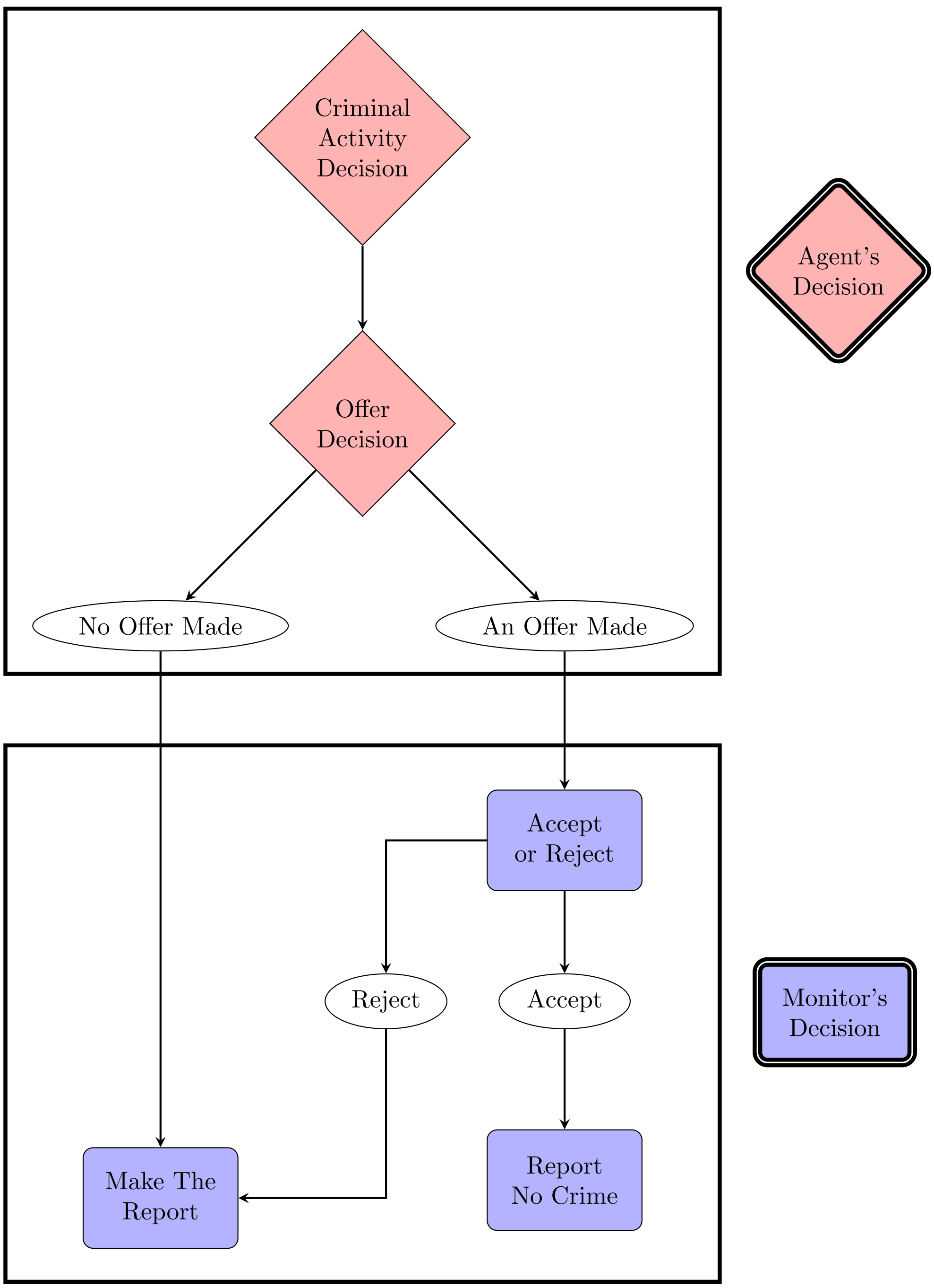
When he chooses to commit a crime ($c = 1$), he gets π_A and faces a potential punishment $k > \pi_A$ depending on the monitor's report m. Therefore, he has an incentive to make a bribe offer $\tau \in (0, \pi_A]$ expecting that the monitor will accept the offer and destroy all the criminal evidence by reporting $m = 0$.

From the monitor's perspective, the expected cost of sending a false report is pw when wage is w, so she accepts the offer if and only if $\tau \geqslant pw$. Thus, given a particular offer τ, the chance that the monitor is going to accept the offer is $\mathsf{prob}(\tau \geqslant pw) = F(\tau/p)$. Therefore, the agent's expected payoff of committing a crime is described as follows:

$$\mathbb{E}(P_A) = \pi_A - \tau F(\tau/p) - k\,(1 - F(\tau/p)) = \pi_A - k + (k - \tau)F(\tau/p) \tag{1}$$

[9]Two key assumptions are made here: First, perfect commitment of the principal on the wage distribution, and second, it is not possible for the monitor to disclose her wage information to the agent in a credible way.

Figure 1: Sequential Moves of The Game



As a result, in order to fully deter the crime and thus corruption, we just have to make sure that the agent can not gain anything from committing a crime no matter how much he would offer to the monitor, which is, find all the wage distributions $F(w)$ such that $\mathbb{E}(P_A) \leqslant 0$. In the spirit of deterring crime and corruption completely at the lowest expected wage cost, the optimal wage distribution $F_q^*(w)$ must satisfy $\mathbb{E}(P_A) = 0$.

According to Ortner and Chassang (2018), for any $p \in (0, 1)$, the unique optimal wage distribution $F_p^*(w)$ which just completely deters crime and corruption is characterized as follows:

$$F_p^*(w) = \frac{k - \pi_A}{k - pw}, \quad \forall w \in [0, \pi_A/p]. \tag{2}$$

Let's denote the set of all these optimal wage distributions as $\mathscr{F} = \{F_p^*(w) \mid p \in (0, 1)\}$. For the rest of the paper, we only consider any wage distribution $F_p^*(w) \in \mathscr{F}$, which implies that $(p, F_p^*(w))$ just fully deters crime and corruption.

The expected wage under $F_p^*(w)$ is:

$$\mathbb{E}(W_p^*) = \int_0^{\pi_A/p} w \cdot dF_p^*(w) = \frac{1}{p}\left[\pi_A - (k - \pi_A)\ln\left(\frac{k}{k - \pi_A}\right)\right] \tag{3}$$

According to Equation 3, $\mathbb{E}[W_p^*]$ is a monotonic decreasing function in p, which implies that, a lower value of p requires a higher level of expected wage $\mathbb{E}(W_p^*)$ in order to fully deter crime and corruption which is quite intuitive.

For any pair of $\{p', F_p^*(w)\}$, whether it can fully deter crime and corruption is determined by the expected payoff $\mathbb{E}(P_A)$ that an agent can acquire if he commits a crime. With the optimal wage structure characterized by Equation 2, the agent's expected payoff $\mathbb{E}(P_A)$ under $\{p', F_p^*(w)\}$ with a bribery offer τ becomes:

$$\mathbb{E}(P_A) = \pi_A - k + (k - \tau)F_p^*(\tau/p') = (k - \pi_A)\left(\frac{k - \tau}{k - \tau + \tau\left(1 - \dfrac{p}{p'}\right)} - 1\right) \tag{4}$$

From Equation 4, we know that, if $\left(1 - \dfrac{p}{p'}\right) < 0$, $\mathbb{E}(P_A) > 0$, which means that the agent can obtain a positive expected payoff from committing a crime and making a bribe offer afterwards, so $\{p', F_p^*(w)\}$ cannot fully deter crime and corruption. Furthermore, the lower the value of $\left(1 - \dfrac{p}{p'}\right)$ is, the higher the expected payoff the agent can obtain from committing a crime.

On the other hand, if $\left(1 - \dfrac{p}{p'}\right) \geqslant 0$, we have $\mathbb{E}(P_A) \leqslant 0$, thus the agent in expectation cannot gain anything from committing a crime, and the higher the value of $\left(1 - \dfrac{p}{p'}\right)$ is, the higher the expected loss the agent bears from committing a crime.

Therefore, for any pair of $\{p', F_p^*(w)\}$, we define *Deterring Power against Crime and Corruption* (hereafter DPCC) be[10]:

$$\text{DPCC} = \left(1 - \frac{p}{p'}\right) \tag{5}$$

and thus we have

- if DPCC < 0, $\{p', F_p^*(w)\}$ cannot deter crime and corruption;
- if DPCC $= 0$, $\{p', F_p^*(w)\}$ just fully deters crime and corruption;
- if DPCC > 0, $\{p', F_p^*(w)\}$ overly deters crime and corruption.

The absolute value of DPCC shows how powerful (when DPCC > 0) or powerless (when DPCC < 0) the pair $\{p', F_p^*(w)\}$ is in deterring crime and corruption.

4 Experimental Design & Procedures

With a proper measure of DPCC given by Equation 5, we can choose the parameters such that the policy design $\{p', F_p^*(w)\}$ in corresponding treatments has the same DPCC. Therefore, any observed treatment difference in terms of deterring crime and corruption is attributed purely to the policy design rather than the difference in DPCC.

4.1 Design and Treatments

We have two policy regimes under investigation in our experiment: $\mathcal{HP}$ and $\mathcal{LP}$. Within each regime, we have one control group and two treatments. Take regime $\mathcal{HP}$ as an example. The control group is denoted as HP, and the two treatments are obtained by increasing either the probability of detection p or the monitor's expected wage $\mathbb{E}(W_p^*)$ from the control, therefore, they are denoted as HPP or HPW respectively. After the manipulation of parameters in treatment HPP and HPW, we maintain the DPCC the same that is defined by Equation 5, so that HPP and HPW are comparable against each other. With such a design, we are able to investigate which policy intervention is more effective in deterring crime and corruption: raising the probability of detection p or raising the expected wage $\mathbb{E}(W_p^*)$.

Similarly for regime $\mathcal{LP}$, we have a control group denoted as LP, and two treatments denoted as LPP and LPW. In addition, we also maintain DPCC the same between the two control treatments HP and

[10]Though it is intuitive to use the opposite of the agent's expected payoff $(-\mathbb{E}(P_A))$ as a measure of DPCC, it is not as clean as this one since $-\mathbb{E}(P_A)$ depends on the agent's offer τ. Alternatively, we can also use the difference between the expected wages $\mathbb{E}(W_p'^*) - \mathbb{E}(W_p^*)$ as a measure of DPCC. Those different measures of DPCC are not going to change the results at all.

LP, and the same across the four treatments HPP, HPW, LPP, and LPW such that the deterrent effect against crime and corruption can also be compared between regime $\mathcal{HP}$ and $\mathcal{LP}$.

Therefore, we have six treatments in total with three of them in each policy regime, and they can be summarized in the following table:

Table 1: Summary of Treatments

Regime	Control	Raising p	Raising $\mathbb{E}(W_p^*)$
$\mathcal{HP}$	LP	LPP	LPW
$\mathcal{LP}$	HP	HPP	HPW

On top of maintaining DPCC the same for parallel treatments, the parametrizations have the following considerations. On one hand, for the two controls (LP and HP), we do not want to fully deter crime and corruption. We would like to choose the parameters such that there are spaces for the agent to commit a crime as well as for the monitor to be corrupted. Equivalently, we want DPCC < 0 for these two control treatments. On the other hand, for the other four treatments (LPP, LPW, HPP, HPW), we would like to choose the parameters such that crime and corruption are overly deterred. Equivalently, we want DPCC > 0 for these four treatments.

When choosing final values for the parameters, we first choose two policy designs that just fully deters crime and corruption, $\{p_L, F^*_{p_L}(w)\}$ and $\{p_H, F^*_{p_H}(w)\}$ with DPCC $= 0$, and we call them baseline policy designs. Then we decrease the level of p_L and p_H to obtain the parameters for our two control treatments while maintaining DPCC the same between them. After that, for each baseline, we obtain the other two treatments by either increasing the detection probability p or increasing the expected wage $\mathbb{E}(w)$ by adjusting the wage distribution $F^*_p(w)$ such that DPCC is positive and the same across these four treatments.

The baseline policy designs have $p_L = 1/3$ in $\mathcal{LP}$ regime and $p_H = 2/3$ in $\mathcal{HP}$ regime[11]. For the wage distributions, we take discrete ones so that it is easier for the experimental subjects to understand. Together with $k = 40$ and $\pi_A = 20$, the corresponding optimal wage distributions are as follows:

Table 2: The Wage Distribution in Baselines

Detection Rate	Wage	Wage Distribution			
		(1/2	1/6	1/6	1/6)
1/3	$F^*_{1/3}$	0	30	48	60
2/3	$F^*_{2/3}$	0	15	24	30

[11] The choice of these two values is mainly for the sake of convenience so that the parameters in our experiment are mostly integers. In addition, the resulted probability of detection is 25% in control LP and 50% in control HP, which is in line with the literature. Banerjee and Mitra (2018) uses 20% in the LP and 40% in the HP.

From the baselines, we decrease p_L and p_H by 25% to obtain the two controls. And then we increase either the probability of detection or the wage distribution in each control by 50% to obtain two treatments in each regime. The parametrization of all treatments is shown in Table 3.

Table 3: The Parametrization of All The Treatments

Treatment	Detection Rate	Wage	DPCC*	Wage Distribution			
				(1/2	1/6	1/6	1/6)
Baseline	1/3	$F^*_{1/3}$	0	0	30	48	60
LP	1/4	$F^*_{1/3}$	-1/3	0	30	48	60
LPP	3/8	$F^*_{1/3}$	1/9	0	30	48	60
LPW	1/4	$F^*_{2/9}$	1/9	0	45	72	90
Baseline	2/3	$F^*_{2/3}$	0	0	15	24	30
HP	1/2	$F^*_{2/3}$	-1/3	0	15	24	30
HPP	3/4	$F^*_{2/3}$	1/9	0	15	24	30
HPW	1/2	$F^*_{4/9}$	1/9	0	22.5	36	45

*If DPCC = 0, it just fully deters crime and corruption. If DPCC < 0, it does not deter crime and corruption. If DPCC > 0, it overly deters crime and corruption.

In control LP, the monitor's expected cost of accepting a bribe offer is at most $1/4 \times 60 = 15$, so a risk neutral monitor should always accept an offer $\tau \geqslant 15$. The agent's expected payoff of committing a crime when $\tau \geqslant 15$ is thus $\pi_A - \tau = 20 - \tau$, which is maximized at $\tau = 15$. For any offer $\tau < 15$, the agent's expected payoff is

$$\mathbb{E}(P_A) = 20\left(\frac{40-\tau}{40-4/3\tau} - 1\right) = 20\left(\frac{3}{12-0.4\tau} - \frac{1}{4}\right) > 0$$

It is easy to note that $\mathbb{E}(P_A)$ is a monotonic increasing function in τ. Therefore, in control LP, assuming risk-neutral and rationality, the theory would predict that the agent always chooses to commit a crime $c = 1$ and make a bribe offer $\tau = 15$, and the monitor always accepts the offer and report $m = 0$. This is consistent with the purpose of the design that we should observe crime and corruption in control LP. Similarly the theoretical prediction in control HP is exactly the same.

For treatment LPP, the agent gets payoff 0 if he chooses $c = 0$. The agent's expected payoff of choosing $c = 1$ is

$$\mathbb{E}(P_A) = 20\left(\frac{40-\tau}{40-3/4\tau} - 1\right) = 20\left(\frac{1}{3} - \frac{16}{48-0.9\tau}\right) < 0 \quad \forall \tau \in (0, 20]$$

which implies that the agent always suffers a loss in expectation if he chooses $c = 1$. As a result, the agent will never choose to commit a crime and thus the monitor has no opportunity to be corrupted. This is also consistent with the intention of the design that we should not observe any crime or corruption in LPP treatment. The same theoretical prediction holds for treatment LPW, HPP, and HPW.

4.2 Hypotheses

Based on the theoretical predictions that we have obtained given the set of parameters, we propose the following hypotheses:

Hypothesis 1A: *In controls LP and HP, crime and corruption are pervasive. Specifically, the agent always chooses to commit a crime* $c = 1$ *and makes an offer* $\tau = 15$*. The monitor always accepts the offer and makes a false report* $m = 0 \neq c$*.*

Hypothesis 1B: *There is no significant difference in the crime and corruption rate between controls LP and HP.*

Hypothesis 2A: *In treatment LPP, LPW, HPP, and HPW, crime and corruption are fully deterred. Specifically, the agent always chooses not to commit a crime* $c = 0$*, and does not make any offer to the monitor* $\tau = 0$*. The monitor always report truthfully* $m = 0 = c$*.*

Hypothesis 2B: *There is no significant difference in crime or corruption rate across treatments LPP, LPW, HPP, HPW.*

Hypothesis 1B and Hypothesis 2B focus on the comparison between policy designs that have the same DPCC. There should not be any significant differences in crime and corruption deterrence between/across the treatments that have the same DPCC. Furthermore, when $DPCC < 0$, the agent can gain a positive expected payoff by committing a crime, and the monitor won't be worse off by accepting a bribery offer, so crime and corruption are pervasive, which is described by Hypothesis 1A. On the contrary, when $DPCC > 0$, crime and corruption are fully deterred and thus neither crime nor corruption will be observed, which is described by Hypothesis 2A.

The above hypotheses are proposed with the assumption of experimental subjects being rational and risk-neutral. When at least one of the assumptions is violated, the above hypotheses may not be supported any more. For example, if most of the experimental subjects are risk averse (which is likely to be the case), the subjects in controls LP and HP may be reluctant to commit crimes or corruptions, thus, Hypothesis 1A might not be supported. In addition, subjects' perception of probability might vary a lot: some might over-estimate the presented probabilities, while other's might under-estimate the probabilities. This will further generate variations across controls and treatments.

Furthermore, Doob and Webster (2003) review several papers that examine offenders' thought process with surveys and interviews. Those studies consistently show that very few criminals consider legal consequences when they are planning crimes. A majority of offenders do not think that they would be caught, or they refuse to think about the possibility of being caught. If this is also true in our experiment, subjects in our four treatments (LPP, LPW, HPP, and HPW) might also commit crimes as well as corruptions, as a result, the differences predicted by the theory between controls and treatments might be much smaller or even vanish. Therefore, we propose Hypothesis 3 as the null to see whether it will be supported or not.

Hypothesis 3 *Raising the probability of detection p or raising the monitor's expected wage $\mathbb{E}(W_p^*)$ are both significantly effective in deterring crime and corruption, i.e., compared against the control LP (HP, respectively), the crime and corruption rate are significantly lower in treatment LPP, LPW (HPP, HPW respectively).*

Last but not least, a great body of literature supports the hypothesis that an increase in certainty of punishment delivers a larger deterrent effect against crime (or corruption) than an equivalent increase in severity of punishment. Therefore, we also would like to propose Hypothesis 4 to further test this.

Hypothesis 4 *Raising the detection rate p is more effective in deterring crime and corruption than raising the monitor's expected wage $\mathbb{E}(W_p^*)$ in both regimes $\mathcal{HP}$ and $\mathcal{LP}$, i.e., compared against treatment LPW (HPW, respectively), the crime and corruption rate are significantly lower in treatment LPP (HPP respectively).*

4.3 Experimental Procedures

The experiment was conducted at the CATI lab, School of Social Science, Nanyang Technological University using ztree (Fischbacher, 2007). Participants were recruited from a pool of undergraduate volunteers via emails. Upon arrival, the experimenter read the instructions aloud while the subjects were reading their own copies at the same time. Sessions lasted around 60 minutes and participates earned on average S$14 including a show-up fee of S$3. In total, we have 198 undergraduate students participated in our experiment with 102 male students and 96 female students.

Among all the participants, the gender ratio is well balanced, and the average age is around 21∼22 year's old. Half of them has gained some knowledge on game theory before they participate in this experiment, and around 80% of them had participated in other experiments before. The detailed demographics across all the treatments are shown in Table 4.

In our experimental instructions, we describe the game in a neutral way without saying anything about crime or corruption in order to eliminate any potential effects caused by the crime or corruption context, although Abbink and Hennig-Schmidt (2006) show that there is no significant difference in results between neutral-context and in-context presentation of experimental tasks in a bribery game. For details on our instructions, please see Appendix B.

Table 4: Demographics Across All The Treatments

Treatment	No. of Subjects	Male Ratio	Age	% of Singaporean	% Experiment-Exp	% Theory-Exp
HP	28	57.1%	22.4	64.3%	75%	42.9%
HPP	38	47.4%	21.8	34.2%	81.6%	44.7%
HPW	38	63.2%	20.6	15.8%	86.8%	44.7%
LP	22	50.0%	21.8	41.7%	86.1%	38.9%
LPP	38	65.8%	22	42.1%	89.5%	52.6%
LPW	34	47.1%	21.2	41.2%	85.3%	41.2%
Total	198	55.6%	21.6	39.4%	84.3%	42.9%

5 Results

Since the monitor's wage is randomly generated by a computer random device from period to period, it is possible that the realized wage distribution in each treatment is not the same as it should be according to the theoretcial distribution. If this is the case, we might fail to observe some treatment differences that should have been observed, or the observed differences cannot be fully attributed to the design of treatments. Therefore, we compare the mean and standard error between the wages according to the theory and the actual realized wages for each treatment in Table 5, and we also compare the distributions in Figure 2.

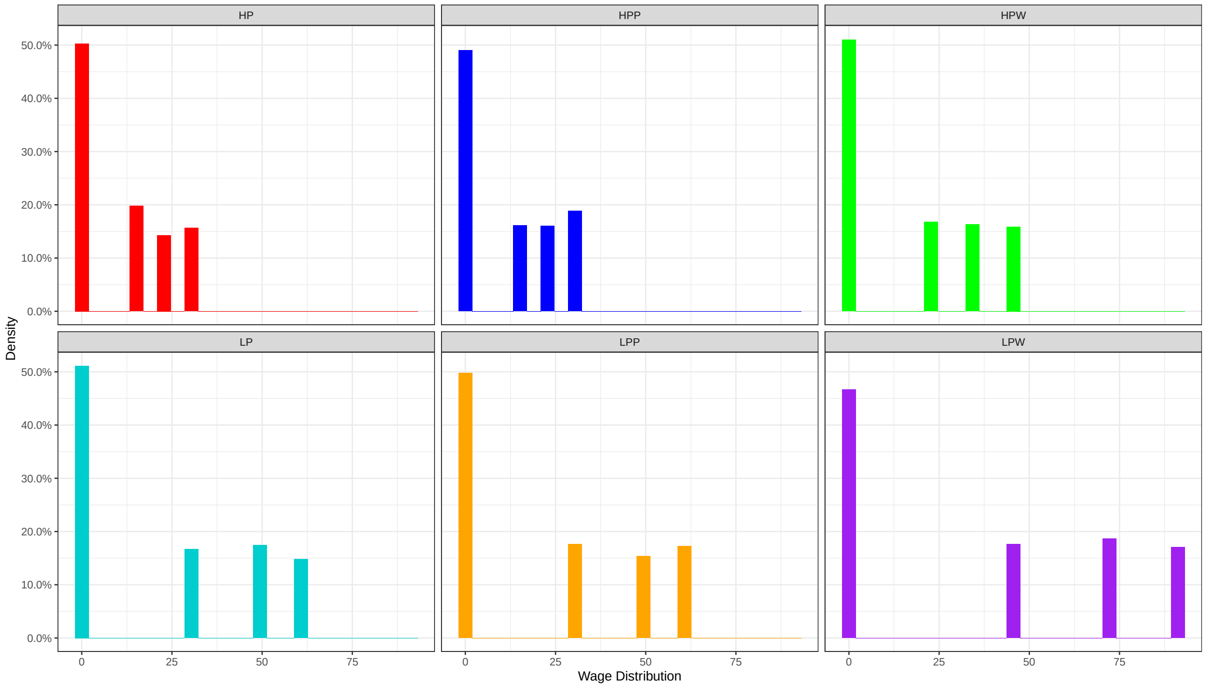
Table 5: Wage Comparison Between Observations and Theory

Treatment	Mean		SD		SE	
	Observed	Theory	Observed	Theory	Observed	Theory
LP	22.2	23	24.3	26.9	1.06	NA*
LPP	23.0	23	24.6	26.9	0.814	NA
LPW	36.7	34.5	36.9	40.4	1.29	NA
HP	11.1	11.5	12.0	13.5	0.464	NA
HPP	11.9	11.5	12.5	13.5	0.414	NA
HPW	16.8	17.2	18.3	20.2	0.607	NA

* Standard error is only available for sampling distributions.

Table 5 shows that the means of randomly realized wage in each treatment is very close to the parameter values in theory. Figure 2 shows that the wage distribution is almost the same as the distribution in theory: 1/2 of the wages are zero and 1/6 for each of the rest wage levels.

Figure 2: Wage Distribution For Each Treatment



5.1 Crime & Corruption Rate in General

Table 6 shows the descriptive statistics of the rate of crime and corruption across all the treatments, and the first thing we can notice is that, in all of the treatments, the crime and corruption rates are positive while far less than unity.

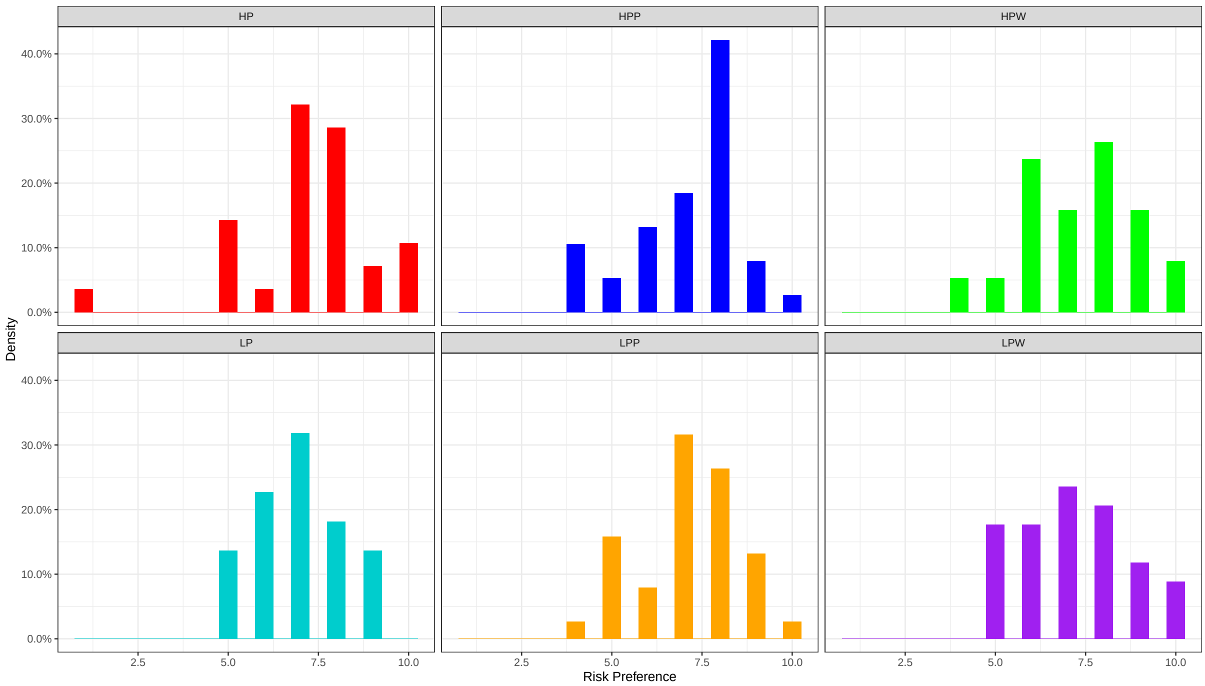
Table 6: Descriptives of Decisions on Illegal Activity and Corruption

Treatment	Crime			Corruption			Theoretical Prediction
	Mean	SD	SE	Mean	SD	SE	
HP	60.0%	0.490	0.0189	38.2%	0.486	0.0188	1
HPP	48.2%	0.500	0.0166	23.1%	0.422	0.0140	0
HPW	54.5%	0.498	0.0165	28.3%	0.451	0.0149	0
LP	53.4%	0.499	0.0217	31.6%	0.465	0.0203	1
LPP	56.9%	0.495	0.0164	31.6%	0.465	0.0154	0
LPW	58.5%	0.493	0.0173	36.5%	0.482	0.0169	0

On one hand, in controls LP and HP, the results violate our Hypothesis 1A which states that, in these two treatments, a rational individual that maximizes his/her expected utility should always choose to commit a crime as an agent and accept the bribery offer as a monitor. However, there are over 50% of them that do not behave as the hypothesis predicted. Two potential explanations might help resolve this issue. First, the theoretical prediction is obtained with the assumption of risk neutral individuals, thus, a risk averse individual would demand some risk premium for him/her to take the risk of committing a crime or accepting an offer. Second, although we describe the game in an neutral context in our experimental instructions, subjects can still infer that the "crime decision" is bad[12]. Therefore, they might restrain from these decisions by their in-built moral constrains.

Though we do not have any measure on subject's moral standards, we do have a measure of risk preferences using a MPL task[13] that is very similar to the one proposed by Holt and Laury (2002). Figure 3 show the distribution of risk preferences across all the treatments. The value of 1∼4 indicates that the subject is risk loving, the value of 5 indicates the subject is either risk-neutral or slightly risk averse, and the value of larger than 5 indicates that the subject is risk averse. Therefore, most of the subjects are risk averse across all the treatments, which partially explains the crime and corruption rates in LP and HP being less than unity.

Figure 3: Distribution of Risk Preference Across Treatments



On the other hand, in treatments LPP, LPW, HPP, HPW, the results violate our Hypothesis 2A which states that, in these four treatments, a rational individual that maximizes his/her expected utility should never commit a crime as an agent nor accept an offer as a monitor. However, the results show that at least 20% of them chooses to commit a crime as an agent or accept an offer as a monitor. Since most of the subjects in our experiment are risk averse, we suggest that myopic preferences contribute to this observation. Subjects might be attracted by the immediate benefits of committing a crime or accepting an offer and overlook the coming negative consequences[14]. A review study by Doob and Webster (2003) also lend strong support on this argument by reporting that the literature consistently show that criminal agents seldom consider the consequences of crimes when they choose to do so.

[12] Although we change the crime context into that of a firm choosing different production methods, the subjects can still infer that production method B is bad since they will be punished if they choose production method B and get caught afterwards. For details, please refer to Appendix B.

[13] For details about the task, please see the experimental instructions in Appendix B.

[14] There are other similar theories that can contribute to this observation, for example, present-bias preference, multi-self models, etc.

Result 1: *There is consistently a positive proportion of subjects that chooses to commit a crime as an agent or accept the bribery offer as a monitor in all the treatments.*

Since Hypothesis 1A and Hypothesis 2A are violated, we would like to examine our Hypothesis 1B and Hypothesis 2B to see whether there is any significant differences between controls or across treatments. In addition, Hypothesis 3 is also subject to be tested to see whether there is any significant treatment effect.

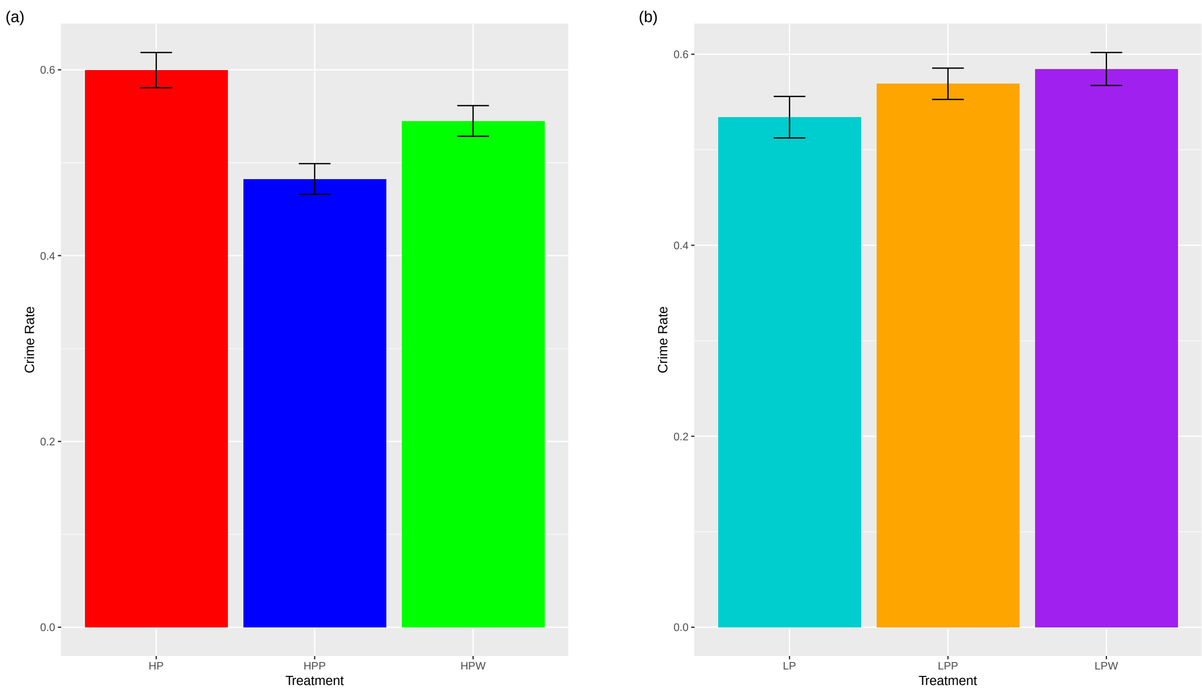
5.2 Nonparametric Results on Crime & Corruption

In this section, we present the results on the mean of crime and corruption rates across all the experimental treatments. We also test the significance of any differences between treatments using the clustered Wilcoxon Rank-Sum test[15] (denoted as C-WRS test hereafter).

5.2.1 Crime Deterrence Across Treatments

Figure 4 shows the rate of crime decisions across all the treatments (as well as the standard error on top of each bar).

Figure 4: Crime Decision Across Treatments



Crime deterrence in regime $\mathcal{HP}$

[15]In our experiment, one subject has to play the game for 24 periods for each role, and the results of the current period is provided at the end of each period. As a result, these 24 observations from one subject are not independent from each other, therefore, each subject is taken as a cluster in our data analyses. We perform the clustered rank-sum test with the D-S method proposed by Datta and Satten (2005). We cannot use the commonly used RGL method since it assumes the observations within one cluster are exchangeable which is not the case in our data generating process.

In regime $\mathcal{HP}$ with high detection probability of false report and low expected wage, Figure 4(a) shows that, compared against the control HP, the crime rate decreases when we raise the detection rate in treatment HPP or when we raise the expected wage in treatment HPW while maintaining DPCC the same between these two treatments, and the deterrent effect is larger in magnitude in treatment HPP than that in treatment HPW.

The rate of choosing to commit a crime in control HP is 60%. It decreases to 48.2% in treatment HPP and the difference is significant ($p = 0.036$, one-sided C-WRS test). By contrast, the crime rate decreases to 54.5% in treatment HPW, while the difference is not significant ($p = 0.209$, one-sided C-WRS test). Therefore, we have our second result.

Result 2: *Compared against control HP with high p and low $\mathbb{E}(w)$, the crime rate decreases significantly in treatment HPP when we increase the p. In contrast, although there is a decease in the crime rate in treatment HPW when we increase the $\mathbb{E}(w)$, the decease is not statistically significant.*

Crime deterrence in regime $\mathcal{LP}$

In regime $\mathcal{LP}$ with low detection probability of false report and high expected wage, Figure 4(b) shows that there is no evident difference of the crime rate between treatment LPP and LP or between treatment LPW and LP, and the clustered rank-sum test shows that these differences are not statistically significant ($p = 0.651$, and $p = 0.713$ respectively, one-sided C-WRS test). Thus, we have our Result 3.

Result 3: *Compared against the control LP with low p and high $\mathbb{E}(w)$, the crime rate does not decrease in either the LPP treatment or the LPW treatment, and the difference is not statistically significant.*

Result 2 only partially support our Hypothesis 3 and Result 3 does not support it at all in the domain of crime deterrence. We only observe a significant decrease (at the 5% level) in crime rate from control HP to treatment HPP when the detection probability of false report increases from 50% to 75% in regime $\mathcal{HP}$.

The above two results partially support our Hypothesis 4. In regime $\mathcal{HP}$ where $p = 50\%$ in control HP, an increase in certainty of punishment significantly deters crime while an increase in severity of punishment does not. However, in regime $\mathcal{LP}$ where $p = 25\%$ in control LP, there is no significant difference in the effectiveness of deterring crime between an increase in certainty and an increase in severity of punishment. Both interventions fail to deter crime significantly.

This suggests that Hypothesis 3 and Hypothesis 4 is dependent on the regime. They are only (partially for Hypothesis 3) supported when the probability of detection is high enough in regime $\mathcal{HP}$ (50% in the control HP). When the probability of detection is low in regime $\mathcal{LP}$ (25% in the control LP), neither increasing the probability of detection nor increasing the expected wage would deliver a significant deterrent effect against crime.

We speculate that, instead of being rational and risk-neutral, the agent forms a subjective belief on the probability of his offer being rejected given the current state of nature. In our experiment, the deterrent effect against crime is indirect. The certainty and severity of punishment that we manipulate in the experiment is the detection probability and the expected wage of the monitor. By manipulating these two factors, the agent formulates his belief on the probability of his offer being rejected. Specifically, if an increase in the probability of detection (or the expected wage) significantly increases the agent's belief on the probability of his offer being rejected, then the crime might be deterred. According to our results, it seems that the responsiveness of the agent's subjective belief on the probability is dependent on the regime. In regime $\mathcal{LP}$ with low probability of detection, the belief is not responsive to variations in either the certainty or severity of punishment on the monitor's side. However, in regime $\mathcal{HP}$ with high probability of detection, it is more responsive to an increase in the certainty of punishment than an increase in the severity of punishment. This actually contributes to resolving the controversy on the relative effectiveness between certainty and severity of punishment in the literature by showing that this issue is regime dependent.

Comparison between regime $\mathcal{HP}$ and $\mathcal{LP}$

In this part, we focus on the comparisons between controls or across treatments that have the same DPCC in both regimes.

In the controls, the crime rate is 60.0% in control HP and 54.5% in LP, however, the difference is not statistically significant ($p = 0.443$, two-sided C-WRS test). This confirms our Hypothesis 1B on crime choices.

In the treatments that share the same DPCC, the difference in crime rate is not statistically significant between any two treatments of these four. The lowest crime rate is observed in treatment HPP which is 48.2%, and the highest is observed in treatment LPW which is 58.5%. However, this difference is not statistically significant ($p = 0.146$, two-sided C-WRS test)[16]. This confirms our Hypothesis 2B on the effectiveness of crime deterrence across these four treatments.

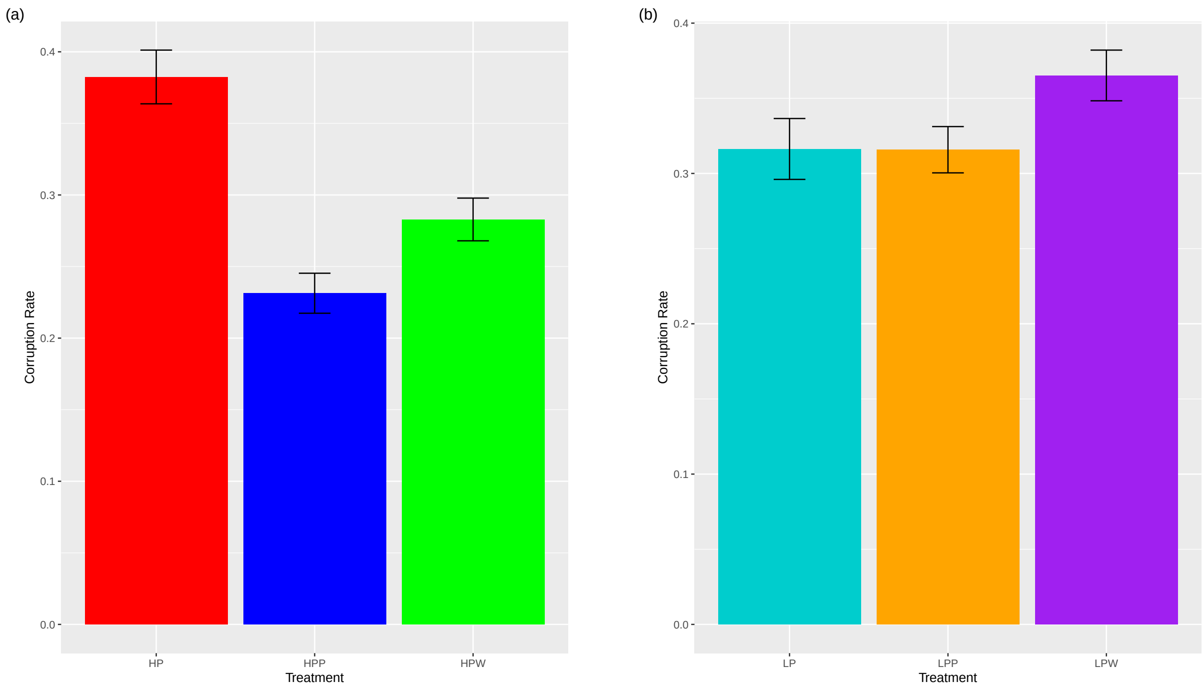
Result 4: *For the effectiveness of crime deterrence between the regimes $\mathcal{HP}$ and $\mathcal{LP}$, there is no statistically significant difference between any two treatments as long as they have the same DPCC.*

5.2.2 Corruption Deterrence Across Treatments

Figure 5 shows the rate of corruption decisions across all the treatments (as well as the standard error on top of each bar).

[16]Here the null hypothesis states no difference between any two treatments, so we use the two-sided test.

Figure 5: Corruption Decision Across Treatments



Corruption deterrence in regime $\mathcal{HP}$

Compared against control HP, Figure 5(a) shows that the corruption rate decreases when we either raise the detection rate in treatment HPP or raise the expected wage in treatment HPW while maintaining the DPCC are the same between these two treatments, and the deterrence effect is larger in magnitude in treatment HPP than that in treatment HPW.

The rate of choosing to accept a bribery offer is 38.2% in control HP. It decreases to 23.1% in treatment HPP which is a huge decrease in magnitude, and the difference is statistically significant ($p = 0.000$, one-sided C-WRS test). Similar result is observed in treatment HPW while the magnitude of the decrease in corruption rate is smaller (from 38.2% to 28.3%), nevertheless, the difference is also statistically significant ($p = 0.004$, one-sided C-WRS test).

This strongly supports both our Hypothesis 3 and Hypothesis 4 in the domain of corruption deterrence. Both policy interventions — increasing the probability of detection and increasing the expected wage — are effective in deterring corruptions which supports the Hypothesis 3. In addition, the deterrent effect when we increase the probability of detection in treatment HPP is larger in magnitude than that when we increase the expected wage in treatment HPW ($p = 0.392$, one-sided C-WRS test), and this supports the Hypothesis 4.

Result 5: *In regime $\mathcal{HP}$, increasing either the probability of detection or the expected wage is significantly effective in deterring corruption. In addition, the deterrent effect is larger when increasing the probability of detection in treatment HPP.*

Corruption deterrence in regime $\mathcal{LP}$

In regime $\mathcal{LP}$, Figure 5(b) shows us that the corruption rate does not decrease when we either raise the probability of detection in treatment LPP or raise the expected wage in treatment LPW while maintaining the DPCC are the same between these two treatments. Neither the difference between treatment LP and LPP nor that between treatment LP and LPW is statistically significant ($p = 0.495$, and $p = 0.881$ respectively, one-sided C-WRS test).

Result 6: *In regime $\mathcal{LP}$, neither increasing the probability of detection nor the expected wage is effective in deterring corruption*

Result 6 shows that regime $\mathcal{LP}$ fails to support our Hypothesis 3 as well as our Hypothesis 4 in the domain of corruption deterrence neither by increasing the detection probability nor by increasing the expected wage.

Result 5 and Result 6 together imply that, Hypothesis 3 and Hypothesis 4 are also regime dependent in the domain of corruption deterrence. In regime $\mathcal{HP}$ where the probability is high (50% in control HP), an increase in either the certainty or severity of punishment deters corruption significantly, and the deterrent effect is larger in magnitude with an increase in certainty of punishment. In contrast, in regime $\mathcal{LP}$, neither an increase in certainty nor severity of punishment deters corruption.

The above results suggest that the monitors are responsive to variations in either certainty or severity of punishment when certainty of punishment is high enough. In contrast, they tend to ignore the variations in either certainty or severity of punishment when certainty of punishment is low, and thus, policy interventions in regime $\mathcal{LP}$ are not expected to deliver an expected deterrent effect.

Comparison between regime $\mathcal{HP}$ and $\mathcal{LP}$

There is a noticeable difference in the corruption rate between the controls HP (38.2%) and LP (31.6%), but this difference is not statistically significant ($p = 0.125$, two sided test). This supports our Hypothesis 1B in the domain of corruption deterrence.

For the other four treatments that share the same DPCC, the corruption rate is lower in treatment HPP (23.1%) than that in treatment LPP (31.6%), and similarly, it is lower in treatment HPW (28.3%) than that in treatment LPW (36.5%). These differences are also statistically significant ($p = 0.011$ and $p = 0.021$ respectively, two sided test).

For the comparison between the treatments within each regime, we only find a slightly significant difference in the corruption rate between treatment HPP (23.1%) and treatment HPW (28.3%, $p = 0.078$, two-sided test), but the difference is not statistically significant between treatment LPP (31.6%) and treatment LPW (36.5%, $p = 0.189$, two-sided test).

Result 7: *Between regime $\mathcal{HP}$ and $\mathcal{LP}$, the corruption rate is not significantly different between controls HP and LP, however, the difference is significant between treatments. Specifically, policy interventions in regime $\mathcal{HP}$ deliver a large deterrent effect against corruption than that in regime $\mathcal{LP}$.*

Result 7 supports our Hypothesis 1B but violates our Hypothesis 2B. This again demonstrates the superiority of regime $\mathcal{HP}$ over regime $\mathcal{LP}$ in corruption deterrence.

5.3 Regression Results on Crime & Corruption

In this section, we present regression results in each regime with several different model specifications to show that our previously obtained results are pretty robust and consistent. We use logit regressions with robust standard errors that are clustered at the individual level. In addition, with a careful look at the regression results, we find a potential Cobra Effect that should draw some attention when the authority intends to deter corruption.

5.3.1 Regression Results on Crime Decisions

Results on crime in regime $\mathcal{HP}$

Table 7 shows the Logit regression results with several different specifications in regime $\mathcal{HP}$. The first two columns share the same model specification with all the demographic covariates as well as the individual risk preferences.

The first column shows that, compare against control HP, there is a significant decrease in the corruption rate when we raise the detection probability in treatment HPP from 50% to 75%. The coefficient of -0.810 translates to an odds ratio of 0.478 which suggests that the odds of committing a crime in treatment HPP is 52.2% lower than that in control HP, fixing other covariates at the same level.

However, in the second column, the regression suggests that, although there is a decrease in the odds of committing a crime in treatment HPW than that in control group HP, the decrease is not statistically significant when we control for other covariates.

When we include more independent variables into the regressions, the above results still hold only with some changes in the magnitude. It is easy to see that the decrease in the odds in treatment HPW is insignificant, and the decrease in the odds in treatment HPP is always significant irrespective of model specifications. Specifically, when we include the variables of whether an agent makes an offer in the previous period as well as of whether the offer has been accepted in the previous period, the odds of committing a crime in treatment HPP is 40.5% lower than that in control HP. If we further include the variable of the accumulated crime choices in all previous periods, the odds in treatment HPP becomes 32.6% lower than that in control HP.

These regression results across different model specifications are pretty consistent with Result 2 that we have obtained earlier via C-WRS tests.

Results on crime in regime $\mathcal{LP}$

Table 8 shows the Logit regression results with several different model specifications in regime $\mathcal{LP}$. The model specifications are the same as that in regime $\mathcal{HP}$. These results from all the regressions further confirms Result 3 that we have obtained earlier via non-parametric methods.

5.3.2 Regression Results on Corruption Decisions

Results on corruption in Regime $\mathcal{HP}$

Table 9 shows the Logit regression results on corruption choices in regime $\mathcal{HP}$. The first two columns share the same model specification with all the demographic covariates as well as the individual risk preferences, and the only difference in whether HPP or HPW treatment is compared against control HP.

The first column shows a significant decrease in the odds of being corrupted in treatment HPP compared against that in control HP. The coefficient of -0.810 translates to an odds ratio of 0.445 which suggests that the odds of being corrupted in treatment HPP is 55.5% lower than that in control HP, fixing other covariates at the same level. This implies that, in regime $\mathcal{HP}$, raising the detection probability of false report significantly lowers the odds of being corrupted.

The second column also shows a significant decrease in the odds of being corrupted in treatment HPW compared against that in control HP. The coefficient of -0.486 translates to an odds ratio of 0.615 which implies that the odds of being corrupted in treatment HPW is 38.5% lower than that in control HP, fixing other covariates the same. Thus, raising the expected wage of the monitor in HPW is also significantly effective in deterring corruption, though the magnitude is smaller compared to HPP treatment.

In other regression model specifications where we include more independent variables, the above results hold in general with some changes in the magnitude. For example, when we include the variables of whether a monitor makes a false report and whether the false report has been detected in the previous period, as well as the accumulated corruption choice in all previous periods, the odds of being corrupted in treatment HPP is 37.6% lower than that in control HP, and that in treatment HPW is 28.3% lower. This further support our Result 5 that we have obtained earlier via non-parametric method.

When we further include the variable of how much the agent has offered to the monitor into the regression

Table 7: Clustered Logit Regression on Crime in Regime $\mathcal{HP}$

	Dependent variable:					
	Crime					
	(HP vs HPP)	(HP vs HPW)	(HP vs HPP)	(HP vs HPW)	(HP vs HPP)	(HP vs HPW)
TreatmentHPP	-0.739^{***}		-0.520^{**}		-0.394^{**}	
	(0.276)		(0.223)		(0.176)	
TreatmentHPW		-0.617		-0.398		-0.358
		(0.393)		(0.288)		(0.235)
Age	0.061	-0.044	0.056	-0.019	0.055	-0.014
	(0.115)	(0.139)	(0.095)	(0.110)	(0.076)	(0.097)
Grade	-0.113	0.224	-0.150	0.190	-0.098	0.139
	(0.206)	(0.203)	(0.167)	(0.157)	(0.127)	(0.138)
Experiment_Exp	-0.138	-0.166	-0.230	-0.080	-0.177	-0.056
	(0.317)	(0.254)	(0.263)	(0.205)	(0.204)	(0.165)
Theory_Exp	0.194	0.764^{**}	0.234	0.661^{***}	0.201	0.513^{***}
	(0.273)	(0.299)	(0.224)	(0.226)	(0.178)	(0.183)
NATL	-0.793^{**}	-1.104^{**}	-0.702^{**}	-0.890^{***}	-0.561^{**}	-0.713^{***}
	(0.359)	(0.444)	(0.289)	(0.344)	(0.225)	(0.274)
GENDER	-0.220	0.084	-0.160	0.091	-0.156	0.058
	(0.271)	(0.260)	(0.219)	(0.201)	(0.168)	(0.163)
RiskPreference	-0.037	0.010	-0.017	0.018	-0.016	0.010
	(0.096)	(0.093)	(0.074)	(0.071)	(0.056)	(0.057)
OfferStatus_lag1			0.624^{***}	0.933^{***}	0.357	0.617^{**}
			(0.239)	(0.271)	(0.227)	(0.265)
OfferAccept_lag1			0.890^{***}	0.738^{***}	0.858^{***}	0.781^{***}
			(0.233)	(0.256)	(0.232)	(0.255)
Cum_Crime					0.116^{***}	0.105^{***}
					(0.019)	(0.018)
OfferStatus_lag1:OfferAccept_lag1						
Constant	0.354	1.075	-0.246	-0.364	-1.213	-0.928
	(2.502)	(2.824)	(2.005)	(2.196)	(1.588)	(1.890)
Observations	1,584	1,512	1,518	1,449	1,518	1,449
Log Likelihood	$-1,060.373$	-969.700	-945.952	-850.750	-901.945	-816.278
Akaike Inf. Crit.	2,138.746	1,957.399	1,913.903	1,723.500	1,827.890	1,656.556

Note: *p<0.1; **p<0.05; ***p<0.01

Table 8: Clustered Logit Regression on Crime in Regime $\mathcal{LP}$

	Dependent variable:					
	Crime					
	(LP vs LPP)	(LP vs LPW)	(LP vs LPP)	(LP vs LPW)	(LP vs LPP)	(LP vs LPW)
TreatmentLPP	0.012		0.031		−0.020	
	(0.411)		(0.329)		(0.243)	
TreatmentLPW		0.314		0.235		0.088
		(0.381)		(0.303)		(0.230)
Age	−0.086	0.086	−0.026	0.056	−0.021	0.045
	(0.123)	(0.099)	(0.104)	(0.076)	(0.070)	(0.058)
Grade	0.067	−0.051	0.019	−0.050	−0.010	−0.052
	(0.215)	(0.180)	(0.173)	(0.141)	(0.117)	(0.103)
Experiment_Exp	0.391	−0.388	0.290	−0.383	0.177	−0.248
	(0.619)	(0.536)	(0.484)	(0.406)	(0.382)	(0.324)
Theory_Exp	0.390	−0.142	0.275	−0.084	0.234	−0.035
	(0.406)	(0.411)	(0.342)	(0.318)	(0.248)	(0.244)
NATL	0.456	−0.160	0.342	−0.084	0.294	−0.044
	(0.449)	(0.376)	(0.352)	(0.299)	(0.252)	(0.227)
GENDER	0.401	0.453	0.272	0.203	0.204	0.095
	(0.405)	(0.344)	(0.346)	(0.281)	(0.245)	(0.220)
RiskPreference	-0.241^{*}	-0.211^{*}	−0.152	−0.131	−0.070	−0.062
	(0.138)	(0.120)	(0.124)	(0.092)	(0.092)	(0.070)
OfferStatus_lag1			1.437^{***}	1.243^{***}	0.822^{***}	0.702^{***}
			(0.293)	(0.301)	(0.260)	(0.272)
OfferAccept_lag1			0.398^{*}	1.127^{***}	0.486^{**}	1.279^{***}
			(0.237)	(0.271)	(0.228)	(0.285)
Cum_Crime					0.181^{***}	0.141^{***}
					(0.023)	(0.021)
OfferStatus_lag1:OfferAccept_lag1						
Constant	2.653	0.108	0.233	−0.683	−1.081	−1.707
	(2.590)	(2.156)	(2.221)	(1.727)	(1.604)	(1.391)
Observations	1,416	1,344	1,357	1,288	1,357	1,288
Log Likelihood	−942.669	−890.723	−801.163	−718.599	−703.142	−659.962
Akaike Inf. Crit.	1,903.338	1,799.445	1,624.327	1,459.199	1,430.283	1,343.925

Note: *p<0.1; **p<0.05; ***p<0.01

model, the regression results are still consistent with other model specifications in treatment HPP.

However, the results in column six in Table 9 suggest that there is not a significant decrease in the odds of being corrupted in treatment HPW compared against control HP holding other covariates the same. Recall that Figure 5 shows an abrupt decrease in the corruption rate in treatment HPW (28.3%) compared against control HP (38.2%), and Result 5 states that this decrease is statistically significant. The reason for this exceptional result in column six in Table 9 will be discussed with another seemingly contradictory result observed in regime $\mathcal{LP}$ together in Section 5.4.

Results on corruption in regime $\mathcal{LP}$

Table 10 shows the Logit regression results with several different model specifications in regime $\mathcal{LP}$, and the model specifications are the same as that in regime $\mathcal{HP}$. Except for the regression result in column six, all the regressions fail to show any significant treatment effects which further confirms Result 6 that we have obtained earlier via non-parametric methods.

The regression in column six suggests a significant decrease in the odds of being corrupted in treatment LPW than that in control LP. In column six, the coefficient of -0.341 translates to an odds ratio of 0.711 which implies that the odds of being corrupted in treatment LPW is 28.9% lower than that in control LP holding other covariates the same. However, Figure 5 actually shows a noticeable increase in the corruption rate in treatment LPW (36.5%) than that in control LP (31.6%), although Result 6 tells us that this increase is not statistically significant ($p = 0.119$, one-sided C-WRS test). This seemingly contradictory result is actually a representation of the Cobra Effect, which will be discussed in detail in Section 5.4.

The general results obtained above suggests a strong superiority of regime $\mathcal{HP}$ over regime $\mathcal{LP}$ when policy makers are pondering between these two. Even if one can only choose between two policy designs that resembles the two controls HP and LP due to resources constraint at the moment, he/she should still choose HP over LP. Although there is no significant difference between them in deterring crime and corruption, regime $\mathcal{HP}$ gives the authority the power to effectively deter crime and corruption in the future when he/she has the required resources to operate on the corresponding policy instruments, while regime $\mathcal{LP}$ does not.

5.3.3 Path Dependence of Crime and Corruptions

For crime decisions in both regime $\mathcal{HP}$ and $\mathcal{LP}$, the results on treatment effect is not consistent. However, the regression results from column three to six in both Table 7 and Table 8 consistently show that the offer status in the previous period is a significant predictor of the crime decision in the current period.

Table 9: Clustered Logit Regression on Corruption in Regime $\mathcal{HP}$

	Dependent variable:					
	Corruption					
	(HP vs HPP)	(HP vs HPW)	(HP vs HPP)	(HP vs HPW)	(HP vs HPP)	(HP vs HPW)
TreatmentHPP	-0.810^{***}		-0.472^{***}		-0.468^{**}	
	(0.154)		(0.110)		(0.218)	
TreatmentHPW		-0.486^{***}		-0.332^{***}		0.015
		(0.164)		(0.116)		(0.177)
Age	-0.003	-0.009	-0.007	-0.011	-0.065	-0.099
	(0.058)	(0.078)	(0.039)	(0.060)	(0.083)	(0.116)
Grade	-0.065	0.050	0.001	0.070	-0.128	0.227
	(0.094)	(0.141)	(0.060)	(0.101)	(0.132)	(0.142)
Experiment_Exp	0.048	-0.059	-0.050	-0.142	-0.319	-0.445^{**}
	(0.210)	(0.191)	(0.118)	(0.122)	(0.254)	(0.212)
Theory_Exp	0.258	0.510^{***}	0.164	0.363^{***}	0.172	0.583^{**}
	(0.159)	(0.169)	(0.102)	(0.120)	(0.191)	(0.230)
NATL	-0.186	-0.077	-0.145	-0.055	-0.013	-0.237
	(0.198)	(0.196)	(0.130)	(0.147)	(0.251)	(0.278)
GENDER	-0.083	0.046	-0.037	0.045	-0.009	0.107
	(0.169)	(0.163)	(0.105)	(0.124)	(0.214)	(0.182)
RiskPreference	0.027	0.071	0.006	0.032	0.133^{***}	0.125^{**}
	(0.044)	(0.044)	(0.028)	(0.031)	(0.050)	(0.054)
Offer					0.332^{***}	0.349^{***}
					(0.020)	(0.020)
Cum_Corruption			0.163^{***}	0.139^{***}	0.182^{***}	0.172^{***}
			(0.021)	(0.021)	(0.027)	(0.026)
FalseReport_lag1			0.122	0.100	-0.083	0.044
			(0.203)	(0.166)	(0.312)	(0.255)
Detection_lag1			0.025	0.071	-0.141	-0.022
			(0.185)	(0.142)	(0.279)	(0.233)
FalseReport_lag1:Detection_lag1			-0.090	0.014	-0.077	-0.094
			(0.245)	(0.211)	(0.387)	(0.342)
Constant	-0.366	-1.123	-1.149	-1.516	-2.331	-2.884
	(1.213)	(1.459)	(0.857)	(1.100)	(1.815)	(2.276)
Observations	1,584	1,512	1,518	1,449	1,518	1,449
Log Likelihood	-933.631	-931.086	-864.485	-862.050	-510.268	-522.636
Akaike Inf. Crit.	1,885.263	1,880.172	1,754.969	1,750.100	1,048.537	1,073.271

Note: *p<0.1; **p<0.05; ***p<0.01

Table 10: Clustered Logit Regression on Corruption in Regime $\mathcal{LP}$

	Dependent variable:					
	Corruption					
	(LP vs LPP)	(LP vs LPW)	(LP vs LPP)	(LP vs LPW)	(LP vs LPP)	(LP vs LPW)
TreatmentLPP	−0.118		−0.114		−0.092	
	(0.173)		(0.117)		(0.154)	
TreatmentLPW		0.099		−0.019		-0.341^{**}
		(0.186)		(0.116)		(0.166)
Age	−0.092	−0.012	−0.052	−0.029	−0.057	-0.099^{*}
	(0.076)	(0.052)	(0.043)	(0.035)	(0.051)	(0.054)
Grade	0.191	0.177^{*}	0.113	0.161^{***}	0.094	0.337^{***}
	(0.128)	(0.096)	(0.076)	(0.060)	(0.073)	(0.094)
Experiment_Exp	−0.326	−0.130	−0.165	−0.049	−0.111	−0.135
	(0.248)	(0.278)	(0.185)	(0.195)	(0.255)	(0.279)
Theory_Exp	0.278	0.406^{*}	0.105	0.239^{*}	−0.116	0.320^{*}
	(0.206)	(0.208)	(0.138)	(0.126)	(0.150)	(0.178)
NATL	−0.103	−0.068	−0.111	−0.075	0.122	0.093
	(0.252)	(0.181)	(0.158)	(0.106)	(0.205)	(0.173)
GENDER	−0.092	−0.091	−0.116	−0.064	−0.119	−0.124
	(0.188)	(0.182)	(0.119)	(0.110)	(0.182)	(0.175)
RiskPreference	-0.178^{**}	−0.0001	-0.087^{*}	0.014	0.020	0.026
	(0.072)	(0.061)	(0.048)	(0.038)	(0.065)	(0.065)
Offer					0.269^{***}	0.280^{***}
					(0.023)	(0.022)
Cum_Corruption			0.206^{***}	0.140^{***}	0.183^{***}	0.143^{***}
			(0.024)	(0.019)	(0.029)	(0.030)
FalseReport_lag1			−0.151	0.064	−0.206	0.008
			(0.141)	(0.136)	(0.185)	(0.147)
Detection_lag1			−0.020	−0.160	0.049	0.078
			(0.180)	(0.193)	(0.215)	(0.247)
FalseReport_lag1:Detection_lag1			0.139	0.407	−0.053	0.248
			(0.233)	(0.300)	(0.271)	(0.374)
Constant	2.277^{*}	−0.834	0.079	-1.167^{*}	-1.988^{**}	−1.767
	(1.240)	(1.006)	(0.815)	(0.702)	(0.882)	(1.082)
Observations	1,416	1,344	1,357	1,288	1,357	1,288
Log Likelihood	−860.214	−855.263	−771.286	−786.624	−570.532	−545.455
Akaike Inf. Crit.	1,738.428	1,728.526	1,568.572	1,599.248	1,169.064	1,118.910

Note: *p<0.1; **p<0.05; ***p<0.01

Since an agent only makes an offer when he chooses the crime decision[17], this implies that the agent's crime decision in the current period largely depends on his crime decision in the previous period, and it seems to be a strong evidence of the path dependence of individual's crime decision. In addition, the cumulative frequency of crime choices in the past is also a significant predictor of the crime choice in the current period.

However, in those regressions, the main purpose is to test the significance of the treatment effect when other covariates are controlled. Since the variable *TreatmentHPP/W* is very likely to be correlated with the variable OfferStatus_lag1, the presence of *TreatmentHPP/W* in the regression might affect the magnitude as well as the significance of the variable *OfferStatus_lag1*. Therefore, we run regressions within each treatment to see whether we still observe a significant result that demonstrates the presence of path dependence.

Table 11 shows that the offer status in the previous period is not consistently a significant predictor of the crime decision in the current period across the six treatment groups. This is mainly due to the fact that the offer might be rejected which discourages the agent from committing a crime again. Table 11 also shows that whether the bribery offer has been accepted or not in the previous period is consistently a significant predictor of the agent's crime decision in the current period (except that in treatment LPP), which is very intuitive. When the bribery offer was accepted in the previous period, the odds of committing a crime in the current period is 182% more likely on average than that when the offer was rejected in the previous period.

This result implies that criminal activities can be deterred if the monitor is hard to be corrupted, i.e., if corruption is deterred effectively. The *Operation Ampscam* quoted by Ortner and Chassang (2018) also serves as an example here. Undercover police inspectors in this operation are the hard-to-be-corrupted monitors, so they reject the bribery offers and arrest the contractors trying to get approval for low-quality work.

Furthermore, the accumulated frequency of crime decisions in all the previous periods turns out to be a very significant predictor of the likelihood of committing a crime in the current period. On average, the odds of committing a crime in the current period increases by 14.0% if the agent commits one more crime in previous periods. This result is also a demonstration of path dependence of the crime decisions.

Table 12 shows the regression results on corruption decisions within each treatment groups. The false report that the monitor made in the previous period indicates that she had accepted the bribery offer and thus was corrupted. Neither the corruption choice nor the detection of the false report in the previous period is a significant predictor of the corruption choice in the current period. However, both

[17] This is not always the case in our data. If we consider the choice of making a positive offer in despite of a non-crime decision as irrational choice, there are 3% of the observations that are irrational, and the percentage of irrational choices is pretty persistent over time. Specifically, there are always 2.9% $\sim$ 3% of irrational choices even though we exclude the observations of the first 5 periods, the first 10 periods, or even the first 15 periods (there are 24 periods in total).

Table 11: Clustered Logit Regression on Crime Within Each Treatment

	Dependent variable:					
	Crime					
	HP	HPP	HPW	LP	LPP	LPW
Cum_Crime	0.089***	0.120***	0.109***	0.164***	0.182***	0.119***
	(0.024)	(0.029)	(0.027)	(0.034)	(0.029)	(0.027)
OfferStatus_lag1	0.197	0.382	0.864**	1.090**	0.573*	0.322
	(0.312)	(0.315)	(0.402)	(0.452)	(0.343)	(0.340)
OfferAccept_lag1	0.760**	0.986***	0.929**	0.668*	0.451	1.802***
	(0.367)	(0.289)	(0.389)	(0.347)	(0.298)	(0.412)
Age	0.154	−0.069	−0.240**	0.054	−0.038	0.102
	(0.137)	(0.088)	(0.116)	(0.186)	(0.074)	(0.074)
Grade	0.074	−0.167	0.359*	−0.225	0.090	0.174
	(0.243)	(0.163)	(0.213)	(0.216)	(0.158)	(0.149)
Experiment_Exp	−0.133	−0.399	0.145	−0.499	0.577	−0.195
	(0.292)	(0.306)	(0.295)	(0.711)	(0.531)	(0.278)
Theory_Exp	0.508*	−0.019	0.627**	−0.438	0.305	0.255
	(0.279)	(0.284)	(0.300)	(0.539)	(0.301)	(0.270)
NATL	−0.820**	−0.204	−0.564	−0.050	0.544	−0.482
	(0.333)	(0.330)	(0.372)	(0.357)	(0.391)	(0.294)
GENDER	0.094	−0.199	−0.215	0.261	0.035	−0.142
	(0.336)	(0.207)	(0.225)	(0.494)	(0.299)	(0.260)
RiskPreference	0.095	−0.189*	−0.011	−0.097	−0.079	0.036
	(0.080)	(0.103)	(0.124)	(0.142)	(0.114)	(0.075)
Constant	−4.633*	2.655	2.758	−1.212	−1.198	−3.759**
	(2.365)	(2.382)	(2.388)	(2.839)	(1.910)	(1.810)
Observations	644	874	805	506	851	782
Log Likelihood	−382.506	−504.491	−417.915	−248.704	−446.160	−396.423
Akaike Inf. Crit.	787.013	1,030.982	857.830	519.407	914.320	814.846

Note: *p<0.1; **p<0.05; ***p<0.01

the accumulated corruption choice and the amount of offer in the current period become significant predictors of the corruption choice in the current period.

On average across all the treatments, the odds of being corrupted in the current period significantly increases by 18.1% (ranges from 14.2% to 21.2%) if the monitor accepts the bribery offer once more in all previous periods. This suggests that corruption is also greatly path dependent. The more one is corrupted in the past, the more likely she is going to be corrupted this time.

Result 8: *Both crime and corruption display a strong feature of path dependence. The more one has committed a crime (corruption) in the past, the more likely he is going to commit a crime (corruption) this time. According to our experimental data, the odds of committing a crime (corruption) increases by 14.0% (18.1%) on average per extra crime (corruption) incidents in the past.*

5.4 The Cobra Effect in Regime LP and HP

The famous Cobra Effect is about the misuse of incentives where some unintended consequences lead to the opposite of the "should have been induced" outcomes by the provided incentive. This is the thing that we should be very careful and try the best to avoid when we use incentives to achieve some purpose because it does not only waste the resources that we have, but also makes our situations even worse.

It is possible that Cobra Effect is present in our experiment. We have two policy instruments at hand to modify in order to deter crime and corruption, and one of them is the expected wage of the monitor. We presuppose that an increase in the expected wage of the monitor would make the expected (opportunity) cost of accepting a bribery offer higher than before, and thus the likelihood of the monitor accepting a bribery offer becomes lower. Anticipating this, the agent would be less likely to make an offer and thus less likely to commit a crime. Therefore, increasing the expected wage deters crime and corruption effectively. However, the agent might respond to an increase in the expected wage by increasing his bribery offer such that the effect dominates that of the increase in the expected wage, and thus the monitor takes the offer more likely. And this is what actually happens in treatment LPW. Figure 6 illustrates this potential Cobra Effect.

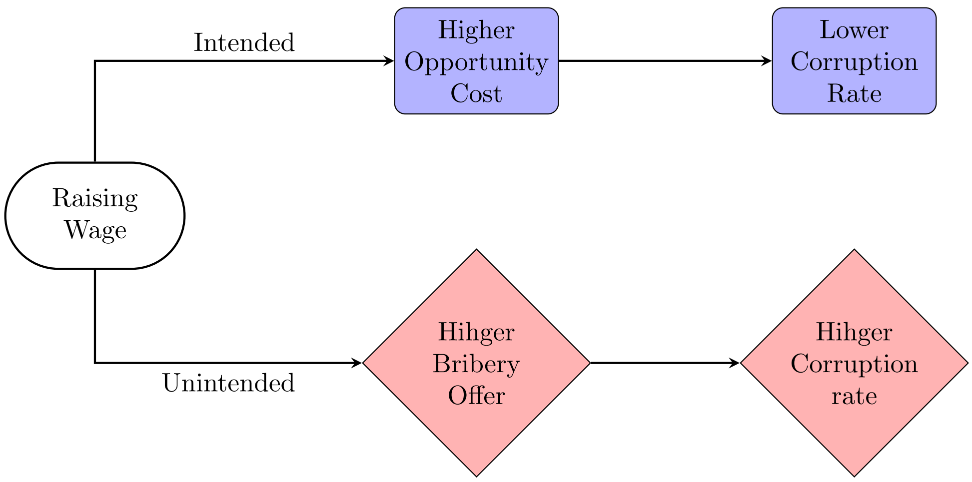
Now let's turn to the experimental data in treatment LPW and control LP. On one hand, Figure 5 shows a noticeable increase in the corruption rate in treatment LPW (36.5%) compared against that in control LP (31.6%), although such an increase is not statistically significant ($p = 0.119$, one-sided C-WRS test).

On the other hand, column six in Table 10 shows a significant decrease in the odds of being corrupted in treatment LPW compared against that in control LP. We should notice that this treatment effect in LPW only becomes significant when we include *how much offer that a monitor receives in the current period* on the R.H.S. of the regression in column six, which implies that, as long as we control the amount

Table 12: Clustered Logit Regression on Corruption Within Each Treatment

	Dependent variable:					
	Corruption					
	HP	HPP	HPW	LP	LPP	LPW
Offer	0.370***	0.321***	0.357***	0.316***	0.251***	0.268***
	(0.030)	(0.025)	(0.030)	(0.045)	(0.024)	(0.025)
Cum_Corruption	0.164***	0.192***	0.168***	0.154***	0.188***	0.133***
	(0.036)	(0.061)	(0.036)	(0.060)	(0.029)	(0.035)
FalseReport_lag1	0.175	−0.510	−0.082	−0.200	−0.230	0.090
	(0.409)	(0.579)	(0.341)	(0.202)	(0.285)	(0.201)
Detection_lag1	0.137	−0.504	−0.199	0.265	−0.046	−0.007
	(0.424)	(0.358)	(0.277)	(0.395)	(0.257)	(0.302)
Age	−0.404**	−0.0002	0.153	−0.092	−0.083*	−0.079
	(0.160)	(0.097)	(0.106)	(0.101)	(0.048)	(0.064)
Grade	0.425*	−0.232	−0.065	0.311**	−0.024	0.410***
	(0.240)	(0.163)	(0.188)	(0.132)	(0.100)	(0.130)
Experiment_Exp	−0.882**	−0.247	−0.210	0.388	−0.250	−0.317
	(0.400)	(0.333)	(0.275)	(0.341)	(0.325)	(0.305)
Theory_Exp	0.340	0.005	0.631**	0.632**	−0.224	0.354*
	(0.413)	(0.222)	(0.281)	(0.313)	(0.162)	(0.201)
NATL	0.433	0.011	−0.600*	0.436	0.169	−0.233
	(0.342)	(0.352)	(0.349)	(0.288)	(0.269)	(0.192)
GENDER	0.685**	−0.102	−0.060	−0.017	−0.099	−0.099
	(0.284)	(0.258)	(0.236)	(0.273)	(0.189)	(0.210)
RiskPreference	0.266***	0.019	0.020	−0.080	0.055	0.085
	(0.055)	(0.065)	(0.086)	(0.114)	(0.069)	(0.073)
FalseReport_lag1:Detection_lag1	−0.218	0.196	0.062	−0.058	−0.032	0.336
	(0.608)	(0.634)	(0.419)	(0.553)	(0.330)	(0.481)
Constant	1.742	−2.638	−6.556***	−2.089	−1.143	−2.784*
	(3.030)	(2.259)	(2.070)	(1.400)	(1.085)	(1.492)
Observations	644	874	805	506	851	782
Log Likelihood	−226.573	−274.417	−287.655	−201.336	−362.736	−339.049
Akaike Inf. Crit.	479.147	574.833	601.311	428.671	751.472	704.098
Note:						*p<0.1; **p<0.05; ***p<0.01

Figure 6: Illustration of Potential Cobra Effect



of offer that a monitor receives at the same level between treatment LPW and LP, the odds of being corrupted in LPW is going to be significantly lower than that in LP. Therefore, our previously observed result in Figure 5 that there is an increase in the corruption rate in LPW is mainly due to a higher level of offer in LPW than that in LP.

Figure 7: Mean Offer Across Treatments

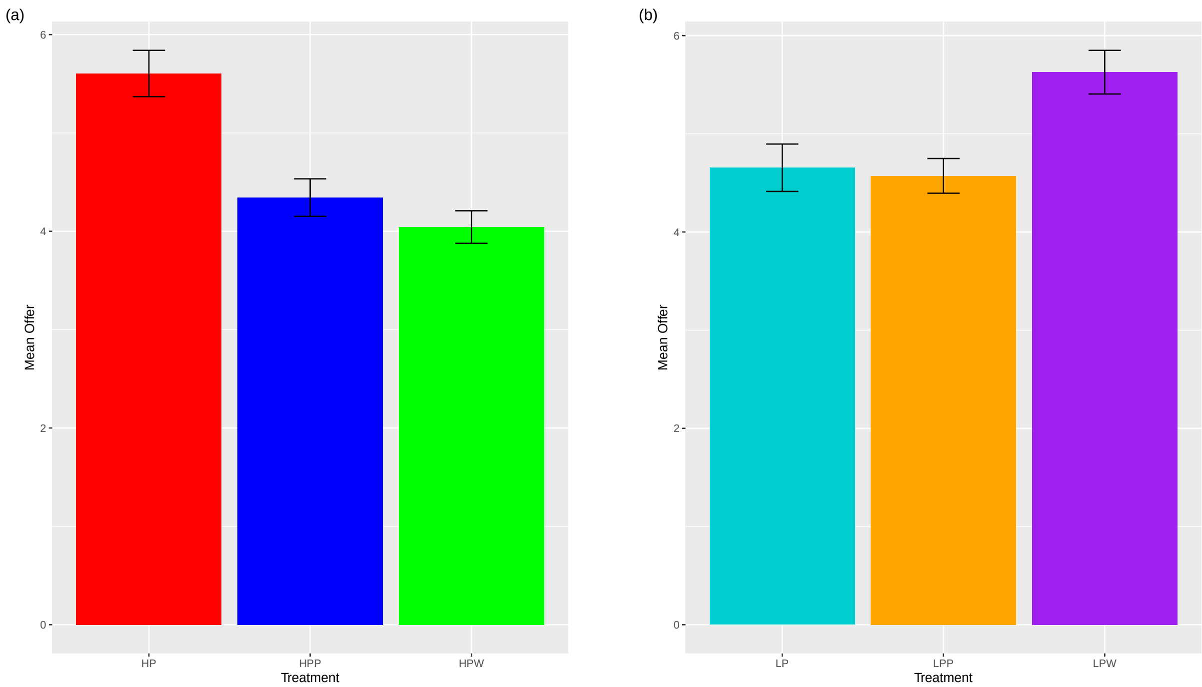


Figure 7 shows the mean offer across all the treatments which confirms the above conjecture. It is obvious that the mean offer in treatment LPW (5.63) is (21.1%) higher than that in control LP (4.65), and this suggests the underlying mechanism of the observed Cobra Effect in treatment LPW. Increasing the expected wage of the monitor in treatment LPW is intended to increase the deterrence power against crime and corruption by increasing the opportunity cost of corruption and thus decrease the corruption rate as well as the crime rate, however, this also induces the agent to provide a higher level of bribery offers, which in turn leads to a higher corruption level.

Is there a similar Cobra Effect when we raise the expected wage of the monitor in regime $\mathcal{HP}$? The answer is no. Figure 7 shows that there is an evident decrease in the mean offer in treatment HPW compared against that in control HP. This rules out the underlying mechanism of the Cobra Effect observed in regime $\mathcal{LP}$. Furthermore, Figure 5 and Result 5 tell us that the corruption rate in treatment HPW is significantly lower than that in control HP, which completely negates the presence of Cobra Effect in regime $\mathcal{HP}$.

In addition, the regression results in column six in Table 9 shows that there is not a significant difference in the corruption rate between treatment HPW and HP, which seems to contradict the results shown in Figure 5 and Result 5. This is also due to a significant decrease in the average offer level in treatment HPW than that in control HP, as a result, if we control for the offer level in the regression, the observed significant decrease in corruption rate in treatment HPW will disappear.

Given the above results, it seems that the presence of the Cobra Effect in our study is regime dependent. In regime $\mathcal{LP}$ with low probability of detection(25% in LP), the increase in the expected wage of the monitor only increases the DPCC on paper, and in effect it leads to an increase in corruption rate since it induces the agent to make a higher bribery offer. In contrast, in regime $\mathcal{HP}$ with high probability of detection (50% in HP), the increase in the expected wage of the monitor results in an increase in the DPCC in effect as intended, and the agents are discouraged to make bribery offers.

Why do the agents respond so differently to an increase in the expected wage of the monitor between regime $\mathcal{HP}$ and $\mathcal{LP}$? We suggest that this is because of the difference in the perception of detection probabilities. In regime $\mathcal{LP}$ where the detection probability is low (25% in LP), the agents are not responsive to changes in the probability of detection (up to a 50% increase). In addition, the agents would consider the monitor to be easily corruptible, and she will take his offer as long as the offer is attractive. As a result, a higher expected wage level of the monitor induces a higher level of bribery offers and thus a higher corruption rate. On the contrary, in regime $\mathcal{HP}$ where the detection probability is high (50% in HP), the agents would take it seriously and consider the monitor to be hardly corruptible. As a result, a higher expected wage level of the monitor in treatment HPW together with the high detection probability induce the agents to believe that the monitor is even harder to be corrupted. As a result, the average level of bribery offers decreases and the corruption rate decreases consequently[18].

Result 9: *As a policy instrument against crime and corruption, raising the (expected) wage of the monitor induces the Cobra Effect in regime $\mathcal{LP}$ where the probability of detection is low (25% in LP). However, such a perversive incentive effect is not present in regime $\mathcal{HP}$ where the detection probability*

[18]The decrease in the mean offer level in treatment HPW shown in Figure 7 is mainly driven by the fact that fewer subjects in treatment HPW make an bribery offer than in HP treatment. Conditional on an offer is being made, the mean offer level is roughly the same. This is shown clearly in Section 5.5.

is high (50% in HP).

This result further lends a strong support on regime $\mathcal{HP}$ over regime $\mathcal{LP}$ since the former would not suffer from the Cobra Effect whereas the latter would. Not only does regime $\mathcal{LP}$ fail to deter crime and corruption effectively, but it might also induce a higher corruption rate if raising the expected wage is taken as the policy intervention.

The regression results in Table 12 show that, when the amount of the offer increases by 1 unit (Experimental Currency), the odds of the monitor being corrupted in the current period increases by 37.0% on average (ranges from 28.5% to 44.8% across all the treatments). This clearly demonstrates the detrimental effect of the presence of Cobra Effect, which further invalidate the use of higher (expected) wages as a policy intervention to deter corruption when the detection probability is low.

The literature also suggests the existence of such a Cobra Effect in theory. Kugler, Verdier, and Zenou (2005) show in a theoretical model that, when bribing costs are low and rents are sufficiently high, increasing policing as well as sanctions can generate higher crime rates. Increases in intended expected punishment leads to an extended corruption rings, which further results in a fall of actual expected punishment and thus yields more crime. Basu, Basu, and Cordella (2016) develop an theoretical model and they notice that, in many cases, a rise in the expected punishment (either by increasing the probability of detection or magnitude of the punishment) will lead to an adjustment of the bribery size to compensate the increased expected punishment which is merely an reallocation of surplus. Our study shows empirically that this is true when the probability of detection is low and thus induces the Cobra Effect, however, when the probability of detection is high, an increase in the expected punishment delivers real deterrence power against crime and corruption in effect.

5.5 Analyses on Intensive-Extensive Margins

Previously we have shown that there are some significant changes in the means of crime rate, corruption rate, and bribery offers across all the treatments. For example, Figure 4 shows that there's a noticeable decrease in the crime rate when we raise the detection probability in treatment HPP compared against that in control HP. However, we do not know where this decrease should be attributed to. Is it a result of fewer agents committing a crime? Or is it a result of each agent committing crimes less frequently? The same concern remains for the significant decease in the corruption rate in both treatment HPP and HPW compared against that in control HP. To uncover the veil, it is necessary for us to perform an analysis on the changes of extensive and intensive margins of crime and corruption.

In addition, the presence of Cobra Effect in treatment LPW requires us to perform such an analysis to further validate its underlying mechanism. Specifically, Figure 7 shows that there is a visibly increase in the mean bribery offer in treatment LPW compared against that in control LP. The question is where this increase should be attributed to. Is it due to an increase in the number of agents that makes a positive bribery offer in treatment LPW? Or is it due to an increase in the bribery offer made by each agent? The argument for the underlying mechanism of Cobra Effect requires that the increase in the mean bribery offer is caused by an increase in the intensive margin of the bribery offer, rather than the other possibility.

For crime choices, we define the extensive margin as the percentage of individuals that commits a crime in each period, and the intensive margin as the mean intensity of crime choices over the 24 periods conditional on an agent does commit a crime. The definitions of the ex- and intensive margins of corruption and offer choices are pretty much the same.

5.5.1 Ex- & Intensive Margin of Crime Decisions

In order to measure the intensive margin of crime choices, we calculate the frequency of committing a crime choice over the entire 24 periods for each individual as his crime intensity, and then take the mean within each treatment conditional on he does commit a crime, i.e., his intensive margin is positive.

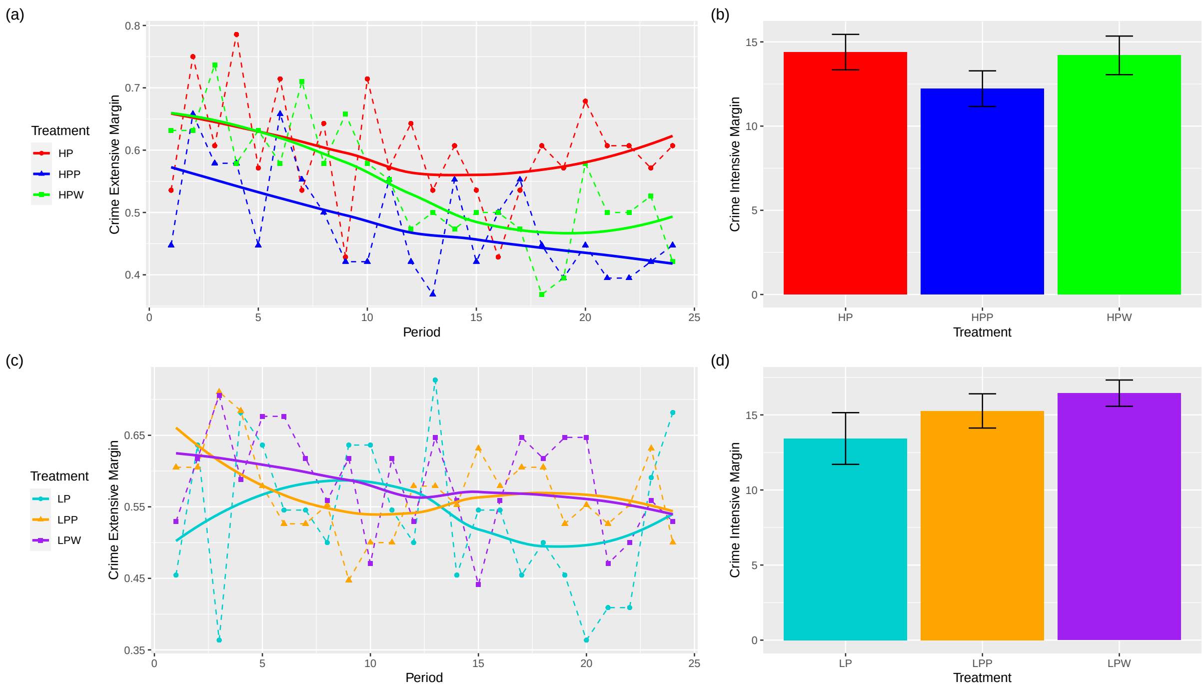
For the measure of extensive margins, we first calculate the percentage of agents that commits a crime in each period as the extensive margin in that period, and then show the dynamic changes of the extensive margin over the entire 24 periods.

Figure 8(a) shows the dynamics of the extensive margin of crime over time in regime $\mathcal{HP}$. It is obvious that the extensive margin in treatment HPP is consistently lower than that in control HP[19]. The interesting part is the dynamics of the extensive margins in treatment HPW. It is almost the same as that in control HP in early periods. However, the agent gradually realizes that it is hard to corrupt the monitor, therefore, the number of the agents that chooses to commit a crime decreases over time, and the extensive margin becomes closer to that in treatment HPP in the end (recall that by design the DPCC is the same in treatment HPP and HPW).

This result suggests that the extensive margins are more sensitive to a change in the probability of detection rather than an expected wage change. As long as the detection probability increases in regime $\mathcal{HP}$, they would take a prompt response to decrease their likelihood of committing a crime, which leads to a lower extensive margin immediately from the very beginning in treatment HPP. In contrast, the agents are slow in response to an expected wage increase when the detection probability remains the same in treatment HPW. Actually, the agents in general cannot differentiate between treatment HPW and control HP in the beginning periods. The deterrent effect that due to an decrease in the extensive

[19] The fitted line is drawn with LOESS (locally estimated scatterplot smoothing) method.

Figure 8: Ex- & Intensive Margin of Crime Across Treatments



margin of crime requires some time to take effect in treatment HPW when we use a higher expected wage as a policy intervention.

This suggests a difference in the celerity of deterrent effect against crime between an increase in certainty and an increase in severity of punishment: increasing the probability of detection deters crime immediately by decreasing the extensive margin immediately, while increasing the expected wage takes some time to produce a decrease in the extensive margin. As a result, if we exclude the first 10 periods of our data, increasing the expected wage should also be effective in deterring crime, and the regression results excluding the first 10 periods in Table A.1 do show that this is true (For details, see Table A.1 in Appendix A).

Figure 8(b) shows the intensive margin of these treatments in the regime $\mathcal{HP}$. Compared against control HP, there is a moderate significant decrease in the intensive margin in treatment HPP ($p = 0.066$, one-sided C-WRS test), while there is not a noticeable change in treatment HPW.

Result 10a: *In the regime $\mathcal{HP}$, there is a significant decrease in the extensive margin and a slightly significant decrease in the intensive margin in treatment HPP when the detection probability is raised (from 50% to 75%). There is not a significant decrease in the intensive margin in treatment HPW when the expected wage is raised, and the extensive margin lies between that in treatment HPP and HP.*

Result 10b: *The agents are more sensitive to a probability change in detection than a change in the expected wage.*

Result 10a and Result 10b further supports our Hypothesis 4 that an increase in certainty of punishment delivers a larger deterrent effect against crime than an increase in the severity of punishment (on the monitor's side). More importantly, it shows where the difference should be attributed to.

Figure 8(c) displays the dynamics of the extensive margin of crime over time in the regime $\mathcal{LP}$. The extensive margins in treatments LP, LPP, and LPW are intertwined with each other and it is hard to say which one is higher or lower than the other. This suggests that, when the detection probability is low (25% in LP), neither increasing the detection probability (from 25% to 37.5%) nor raising the expected wage would have an impact on the extensive margins of crime choices, i.e., both policy interventions won't lower the population of criminals in the regime $\mathcal{LP}$.

In addition, Figure 8(d) shows the intensive margins in these three treatments. Although there are some differences in the mean values, none of them is close to statistical significance.

Result 11: *In the regime $\mathcal{LP}$, there is no significant difference in either the extensive margin or the intensive margin by increasing either the probability of detection or the expected wage.*

This result stands in contrast to the significant differences in the extensive margins among treatments in the regime $\mathcal{HP}$, and it fails to support both Hypothesis 3 and Hypothesis 4. It further validates the idea that both Hypothesis 3 and Hypothesis 4 are regime dependent. The detection probability has to be high enough (as it is in the regime $\mathcal{HP}$) so that both increasing the detection probability and increasing the expected wage would have impacts on the extensive margins of crime, and increasing the probability of detection produces a more prompt decrease in the extensive margin. In addition, the increasing the probability of detection in the regime $\mathcal{HP}$ also produces a significant decrease in the intensive margin of crime.

5.5.2 Ex- & Intensive Margin of Offer & Corruption Decisions

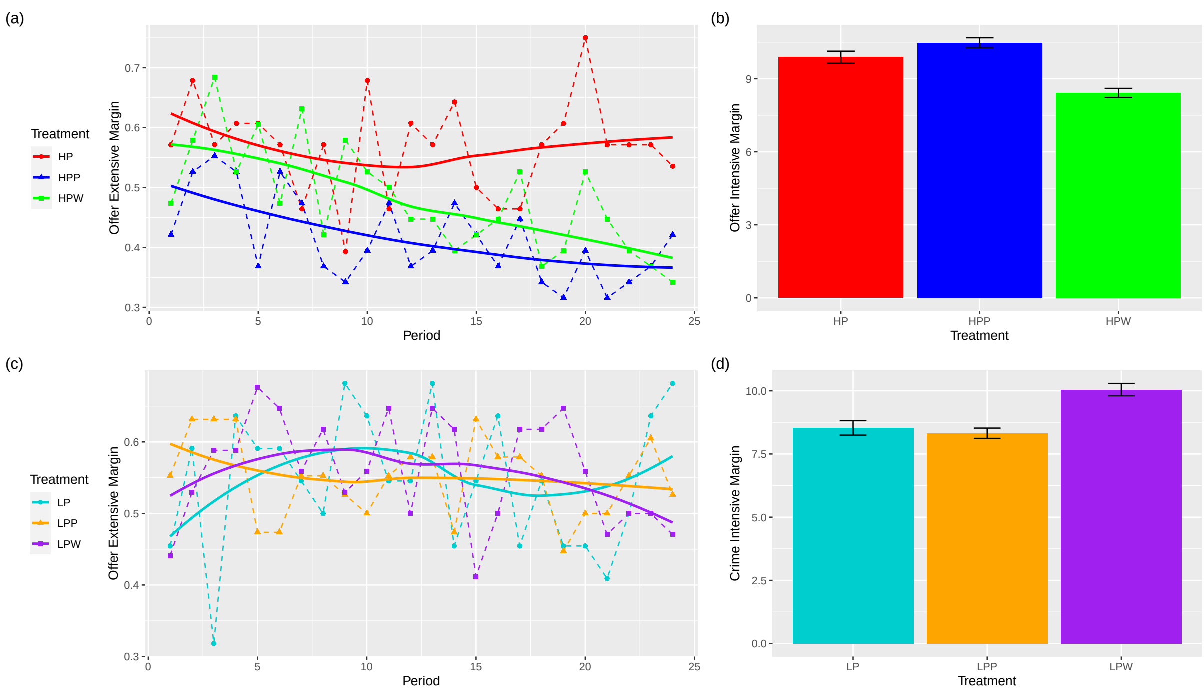
The ex- and intensive margins of offer and corruption across all the treatments are displayed in Figure 9 and Figure 10 respectively.

For the dynamics of the extensive margins of offer and corruption over time across all the treatments, the characteristics are very similar to that of the extensive margins of crime. On one hand, the extensive margins of both offer and corruption in the $\mathcal{HP}$ regime shows a significant difference between treatment HPP and control HP, and the extensive margins in treatment HPW lies in between of them. On the other hand, the extensive margins of both offer and corruption in $\mathcal{LP}$ regime are intertwined with each other which implies insignificant differences among the treatments in $\mathcal{LP}$ regime.

It is quite intuitive and easy to understand the resemblance among the extensive margins of crime, offer, and corruption. As the agent chooses to commit a crime, he will make an offer to the monitor, and more offer implies more opportunity to be corrupted for the monitor, therefore, the extensive margins of crime, offer, and corruption goes hand by hand with each other. For example in control HP, the extensive margin of crime and offer is around 15 at the beginning, and the extensive margin of corruption is around 10 at the beginning which is 5 less than the bribery offer. The pattern of the dynamics of extensive margin over time are very similar among crime, offer, and corruption.

Similar to the results of the extensive margin of crime in regime $\mathcal{HP}$, this suggests a difference in the celerity of deterrent effect against corruption between an increase in certainty and an increase in severity of punishment, which is probably the reason that we observe a larger deterrent effect against corruption in treatment HPP than that in treatment HPW.

Figure 9: Ex- & Intensive Margin of Offer Across Treatments

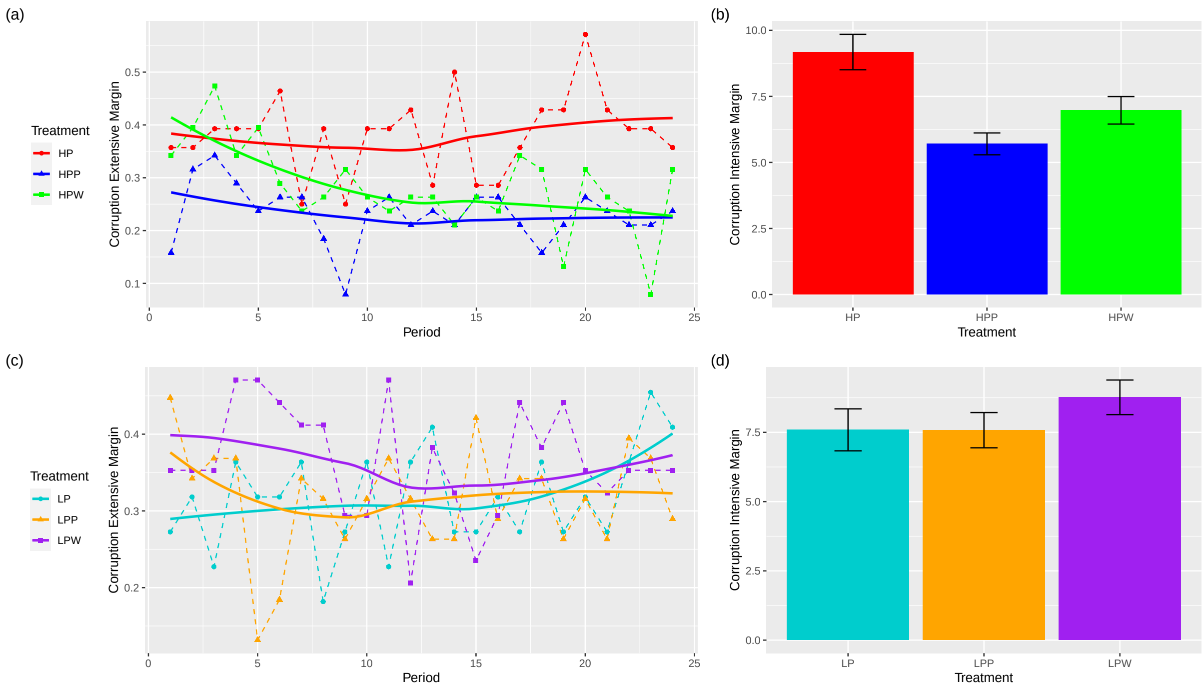


It is a different story for the intensive margins of offer and corruption. Figure 9(b) shows the intensive margins of bribery offers in the $\mathcal{HP}$ regime. Compared against control HP, there is a minor increase in treatment HPP and a slight decrease in treatment HPW, but none of them is statistically significant.

Nonetheless, Figure 10(b) shows an evident decrease in the intensive margins of corruption in both treatment HPP and HPW compared against that in control HP, and these differences are both statistically significant ($p = 0.000$ and $p = 0.007$ respectively, one-sided C-WRS test). This sharp contrast with Figure 9(b) implies a real deterrence power against corruption in both treatment HPP and HPW, especially for the case of treatment HPP - Despite a slightly higher intensive margin of the bribery offer, the intensive margin of corruption in treatment HPP is significantly lower than that in control HP.

However, in the $\mathcal{LP}$ regime, Figure 9(c) and Figure 10(c) show that the extensive margins of both bribery offer and corruption are very close among LP, LPP, and LPW. In addition, the intensive margins

Figure 10: Ex- & Intensive Margin of Corruption Across Treatments



of bribery offers and corruption respectively shown in Figure 9(d) and Figure 10(d) closely resembles each other. When the intensive margin of the bribery offer is higher in treatment LPW, the intensive margin of corruption is also higher. Together with the pattern of extensive margins being roughly the same across treatments, we know that the increase in the mean bribery offer in treatment LPW (compared against that in LP) is mainly due to an increase in the intensive margin of the bribery offers, namely, the agents on average make higher offers to the monitor which is driven by a higher expected wage of the monitor. This completes the demonstration of the underlying mechanism of the observed Cobra Effect in LPW treatment.

All the above results further suggest that the high detection probability in the regime $\mathcal{HP}$ (50% in HP) plays a significant role in rendering not only the theoretical deterrence power on paper but also the real deterrence power in effect when the authority takes one of the policy interventions: increasing the detection probability or raising the expected wage.

Result 12: *The probability of detection plays a vital role in deterring crime and corruption. Specifically, it has to be high enough in order to:*

1. *Render real deterrence power in effect from either increasing the detection probability or raising the expected wage as policy interventions;*
2. *Exclude the potential Cobra Effect when the authority chooses to raise the expected wage of the monitors;*
3. *Generate a larger deterrent effect by increasing the probability of detection than by increasing the expected wage.*

6 Conclusion and Discussion

Follow the long debate on the effectiveness of crime and corruption deterrence between certainty and severity of punishment, this study investigates that, grafted onto the existing policy design concerning crime and corruption deterrence, whether there is a real deterrent effect by increasing the certainty or severity of punishment, and which one delivers a greater impact if there is any.

We experimentally examine the deterrence against crime and corruption simutaneously in a game developed by Ortner and Chassang (2018). The agent in the game decides whether to engage in a crime and receive a benefit of π_A, and he can also bribe the monitor in exchange for an innocent report. The monitor fully observes the agent's decision and report to the authority. However, any false report will be detected with a probability $p > 0$ and the wage will be deprived if a false report is detected. The two main parameters that we manipulate among treatments are the probability of detection p (equivalent to the probability of punishment) and the expected wage of the monitor $\mathbb{E}(W)$ (the higher the $\mathbb{E}(W)$, the severe the punishment), so that we can see the changes in deterrent effect against crime and corruption when there is a change in p or $\mathbb{E}(W)$ on the monitor's side. We further divide our treatments into two regimes: The $\mathcal{HP}$ regime featuring a control with a high p and low $\mathbb{E}(W)$ and the $\mathcal{LP}$ regime featuring a control with a low p and a high $\mathbb{E}(W)$.

In this study we document a novel take-off effect in crime and corruption deterrence when one compares the effectiveness between certainty and severity of punishment. In the $\mathcal{HP}$ regime where the probability of detection is high enough (50% in control HP), both increasing the certainty and increasing the severity of punishment have a significant noticeable deterrent effect against crime and corruption, and the magnitude of deterrence is greater when the effect is due to an increase in the probability of detection. This suggests that both policy interventions would be able to deliver significant corruption and crime deterrence as long as the certainty of punishment is high enough. However, in the $\mathcal{LP}$ regime where the probability of detection is low (25% in control LP), neither increasing the certainty nor increasing the severity of punishment would deter crime or corruption. This might provide an explanation for many empirical studies that fail to find any significant deterrent effect of increased severity of punishment against crime and/or corruption.

We suggest that the above result relates to the difference in risk perception in different regimes. In regime $\mathcal{LP}$, the agents do not take the possibility of the public officials being detected seriously and believe that the officials are easy to be corrupted as long as the bribery offer is high enough. However, when the probability of detection is high, the agents consider the public officials to be hard to be corrupted and thus either policy intervention can produce significant deterrent effect. Nagin (1998, 2013) repeatedly states a research gap on the relationship between risk perceptions and policy regimes regarding crime and corruption deterrence. Our result might shed some light on this issue.

The sharp contrast between the $\mathcal{HP}$ and $\mathcal{LP}$ regimes still persists when we talk about the potential Cobra Effect. When the authority increases the expected wage of the monitor expecting a lower corruption rate due to a higher opportunity cost, the criminal agent can anticipate this and might compensate the monitors with a higher bribery offer which lead to a higher corruption rate instead. This is indeed what happened in LPW treatment. However, in the $\mathcal{HP}$ regime, since the probability of detection is high enough, an increase in the expected wage in HPW treatment yields a significant deterrent effect against corruption.

Furthermore, we perform analyses on changes of extensive and intensive margins of crime and corruption across treatments within each regime, therefore, we are able to identify where the deterrent effect against crime or corruption is attributed to if there is any. Specifically, we investigate whether the policy intervention deters crime (corruption) by decreasing the population of those who commit a crime (corruption) or by decreasing the intensity of those who have committed a crime (corruption). Our results show that the changes of extensive margins of crime and corruption are pretty consistent within each regime. In the $\mathcal{LP}$ regime, there is no significant differences among the three treatments. However, in the $\mathcal{HP}$ regime, the extensive margin of crime (corruption) in HPP treatment is significantly lower than that in control HP from the very beginning. In contrast, the extensive margin of crime (corruption) in HPW treatment is roughly the same as that in control HP in the beginning periods but it gradually decreases over time. This implies that, an increase in the certainty of punishment in the $\mathcal{HP}$ regime can decrease the crime (corruption) population immediately while an increase in the severity will only have a similar impact after a certain period of time. One major difference in the relative effectiveness between certainty and severity of punishment lies in the celerity of their deterrent effect.

The intensive margins of crime and corruption do not differ significantly across treatments within the $\mathcal{LP}$ regime. The story goes differently for crime and corruption in the $\mathcal{HP}$ regime. The intensive margin of crime is not significantly different across treatments within the $\mathcal{HP}$ regime, while the intensive margin of corruption significantly decreases when we increase the probability of detection. We consider that the difference in the pattern of intensive margins between crime and corruption is due to the fact that the policy interventions are on the monitor's side, so they are more prompt and sensitive to the changes.

Future studies can aim to identify the true take-off threshold beyond which certainty and severity are both effective deterrent instruments. It might be different for different types of crime or corruption, and it requires a suitable data set to deliver credible results. It might also be affected by culture, religious, etc., so, the take-off threshold might also be an interval. Another direction is to allow the monitor to demand bribery offers from the agent when the agent is innocent. The harassment bribery game might be the ideal stage game to address this issue, and investigate how corruption would be deterred and how the extensive and intensive margins would change accordingly.

References

Abbink, K., & Hennig-Schmidt, H. (2006). Neutral versus loaded instructions in a bribery experiment. *Experimental Economics*, *9*(2), 103–121.

Ang, Y. Y. (2020). *China's Gilded Age: The Paradox of Economic Boom and Vast Corruption*. Cambridge University Press.

Antunes, G., & Hunt, A. L. (1973, aug). The Impact of Certainty and Severity of Punishment on Levels of Crime in American States: An Extended Analysis. *The Journal of Criminal Law and Criminology (1973-)*, *64*(4), 486–493.

Armantier, O., & Boly, A. (2011). A controlled field experiment on corruption. *European Economic Review*, *55*(8), 1072–1082.

Bajo-Buenestado, R., & Borrella-Mas, M. Á. (2019). Passing-through taxes beyond borders with a cobra effect. *Journal of Public Economics*, *177*, 104040.

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

Banerjee, R., & Mitra, A. (2018). On monetary and non-monetary interventions to combat corruption. *Journal of Economic Behavior & Organization*, *149*, 332–355.

Barr, A., Lindelow, M., & Serneels, P. (2009). Corruption in public service delivery: An experimental analysis. *Journal of Economic Behavior & Organization*, *72*(1), 225–239.

Basu, K., Basu, K., & Cordella, T. (2016). Asymmetric Punishment as an Instrument of Corruption Control. *Journal of Public Economic Theory*, *18*(6), 831–856.

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

Becker, G. S., & Stigler, G. J. (1974, aug). Law Enforcement, Malfeasance, and Compensation of Enforcers. *The Journal of Legal Studies*, *3*(1), 1–18.

Borcan, O., Lindahl, M., & Mitrut, A. (2014). The impact of an unexpected wage cut on corruption: Evidence from a "Xeroxed" exam. *Journal of Public Economics*, *120*, 32–47.

Chalfin, A., & McCrary, J. (2017). Criminal Deterrence: A Review of the Literature. *Journal of Economic Literature*, *55*(1), 5–48.

Chauncey, R. (1975, feb). Certainty, Severity, and Skyjacking. *Criminology*, *12*(4), 447–473.

Datta, S., & Satten, G. A. (2005, aug). Rank-Sum Tests for Clustered Data. *Journal of the American Statistical Association*, *100*(471), 908–915.

Doob, A. N., & Webster, C. M. (2003, aug). Sentence Severity and Crime: Accepting the Null Hypothesis. *Crime and Justice*, *30*, 143–195.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*(2), 171–178.

Friesen, L. (2012, oct). Certainty of Punishment versus Severity of Punishment: An Experimental Investigation. *Southern Economic Journal*, *79*(2), 399–421.

Grogger, J. (1991, apr). CERTAINTY VS. SEVERITY OF PUNISHMENT. *Economic Inquiry*, *29*(2), 297–309.

Holt, C. A., & Laury, S. K. (2002). Risk Aversion and Incentive Effects. *American Economic Review*, *92*(5), 1644–1655.

Kessler, D., & Levitt, S. D. (1999, apr). Using Sentence Enhancements to Distinguish between Deterrence and Incapacitation. *The Journal of Law and Economics*, *42*(S1), 343–364.

Kugler, M., Verdier, T., & Zenou, Y. (2005). Organized crime, corruption and punishment. *Journal of Public Economics*, *89*(9), 1639–1663.

Lucas, D. S., & Fuller, C. S. (2018, aug). Bounties, Grants, and Market-Making Entrepreneurship. *The Independent Review*, *22*(4), 507–528.

Lueck, D., & Michael, J. A. (2003, apr). Preemptive Habitat Destruction under the Endangered Species

Act. *The Journal of Law and Economics*, *46*(1), 27–60.

Mungan, M. C. (2017). The certainty versus the severity of punishment, repeat offenders, and stigmatization. *Economics Letters*, *150*, 126–129.

Nagin, D. S. (1998, jan). Criminal Deterrence Research at the Outset of the Twenty-First Century. *Crime and Justice*, *23*, 1–42.

Nagin, D. S. (2013, aug). Deterrence in the Twenty-First Century. *Crime and Justice*, *42*(1), 199–263.

Niehaus, P., & Sukhtankar, S. (2013). Corruption Dynamics: The Golden Goose Effect. *American Economic Journal: Economic Policy*, *5*(4), 230–269.

Ortner, J., & Chassang, S. (2018). Making corruption harder: Asymmetric information, collusion, and crime. *Journal of Political Economy*, *126*(5), 2108–2133.

Witte, A. D. (1980, aug). Estimating the Economic Model of Crime with Individual Data. *The Quarterly Journal of Economics*, *94*(1), 57–84.

Appendix A Extra Tables

Table A.1: Logit Regression on Crime Excluding The First Ten Periods

	Dependent variable:					
	Crime					
	(HP vs HPP)	(HP vs HPW)	(HP vs HPP)	(HP vs HPW)	(HP vs HPP)	(HP vs HPW)
TreatmentHPP	−0.797**		−0.504*		−0.182	
	(0.323)		(0.260)		(0.221)	
TreatmentHPW		−0.844*		−0.583*		−0.633**
		(0.448)		(0.324)		(0.304)
Age	0.124	−0.035	0.100	−0.005	0.079	0.020
	(0.135)	(0.182)	(0.110)	(0.138)	(0.086)	(0.129)
Grade	−0.078	0.194	−0.177	0.111	0.001	−0.020
	(0.225)	(0.252)	(0.189)	(0.191)	(0.138)	(0.178)
Experiment_Exp	−0.120	−0.242	−0.185	−0.081	0.012	0.059
	(0.377)	(0.293)	(0.301)	(0.231)	(0.226)	(0.164)
Theory_Exp	0.326	0.737**	0.321	0.546**	0.282	0.234
	(0.315)	(0.341)	(0.256)	(0.250)	(0.208)	(0.207)
NATL	−0.944**	−1.239**	−0.774**	−1.041***	−0.316	−0.684***
	(0.422)	(0.488)	(0.341)	(0.375)	(0.244)	(0.260)
GENDER	−0.307	0.039	−0.188	0.024	−0.169	−0.143
	(0.319)	(0.302)	(0.259)	(0.243)	(0.211)	(0.204)
RiskPreference	−0.038	0.032	−0.021	0.032	−0.010	0.013
	(0.104)	(0.102)	(0.081)	(0.076)	(0.052)	(0.064)
OfferStatus_lag1			0.704**	1.163***	−0.160	0.173
			(0.308)	(0.299)	(0.294)	(0.287)
OfferAccept_lag1			1.008***	0.562*	1.055***	0.653**
			(0.310)	(0.295)	(0.327)	(0.310)
Cum_Crime					0.243***	0.247***
					(0.030)	(0.029)
OfferStatus_lag1:OfferAccept_lag1						
Constant	−1.174	0.905	−1.325	−0.515	−3.943**	−2.579
	(2.935)	(3.587)	(2.290)	(2.659)	(1.825)	(2.392)
Observations	924	882	924	882	924	882
Log Likelihood	−617.337	−565.942	−564.597	−512.434	−482.951	−439.714
Akaike Inf. Crit.	1,252.673	1,149.884	1,151.195	1,046.867	989.903	903.429
Note:	*p<0.1; **p<0.05; ***p<0.01					

Appendix B Experimental Instruction

General Instruction

You are now taking part in an interactive study on decision making. **Please pay attention to the information provided here and make your decisions carefully. If at any time you have questions to ask, please raise your hand and we will attend to you in private.**

Please note that unauthorized communication is prohibited. Failure to adhere to this rule would force us to stop the experiment and you may be held liable for the cost incurred in this experiment. You have the right to withdraw from the experiment at any point, and if you decide to do so your payoff earned during this study will be forfeited.

Your **anonymity will be preserved** for the study. You will **never be aware of** the personal identities of other players **during or after** the study. Similarly, other players will also **never be aware of** your personal identities **during or after** the study. You will only be identified by your subject ID in our data collection. All information collected will **strictly be kept confidential** for the sole purpose of this study.

By participating in this study, you will be able to earn a considerable amount of money. The amount depends on the decisions you and others make. Your earnings in the experiment are denominated by "**Experimental Currency Unit(s)**" or "**ECU(s)**". At the end of the experiment, they will be converted into Singapore Dollars at the rate of

1 ECU = 0.05 SGD.

The real-dollar equivalent of your final earnings will be added to your **show-up fee** as your final payoff and paid to you privately in cash at the end of the experiment. It would be contained in an envelope indicated with your unique subject ID. You will need to sign a receipt form to acknowledge that you have been given the correct amount.

Specific Instructions

You will participate in **three** parts of our experiment, the specific instructions will be given to you at the beginning of each part. The following is the specific instruction for part one.

Part One

In this part of the experiment, you will play a game repeatedly for several periods. There are two roles in this game, **Monitor** and **Employee**. At the beginning of Part One, you will get **50 ECUs** as your initial wealth, and your role will be **randomly determined** which will **remain the same** throughout Part One of the experiment.

At the beginning of each period, each Monitor and each Employee will **be randomly paired**, so the **group formation changes from period to period**. The Employee can **choose a production method, either A or B**.

The Monitor is hired by an authority to **watch over the employee and report the production method** chosen by the Employee. The wage the Monitor is going to receive in each period **is determined randomly by a computer program**. Specifically, with 1/2 of the chance, the wage is going to be 0 ECUs; With 1/6 of the chance, the wage is 30 ECUs; With 1/6 of the chance, the wage is 48 ECUs; With 1/6 of the chance, the wage is 60 ECUs. The wage structure can be summarized as follows:

Once the wage is determined at the beginning of each period, the **Monitor will get to know his wage**, while the **Employee will not know the Monitor's wage**. The sequence of decisions for the Employee and the Monitor is described in detail as follows:

Table B.1: Monitor's Wage Structure

Chance	1/2	1/6	1/6	1/6
Wage (ECUs)	0	30	48	60

Step 1 Employee's production decision: The Employee can **choose a production method, either A or B**. Method A yields **10** ECUs, while method B yields **30** ECUs which is 20 ECUs more than method A. Once the Employee has made the production method choice, the Monitor will observe the choice.

If the Employee chooses production method A, **he will get 10** ECUs regardless of what report the Monitor makes.

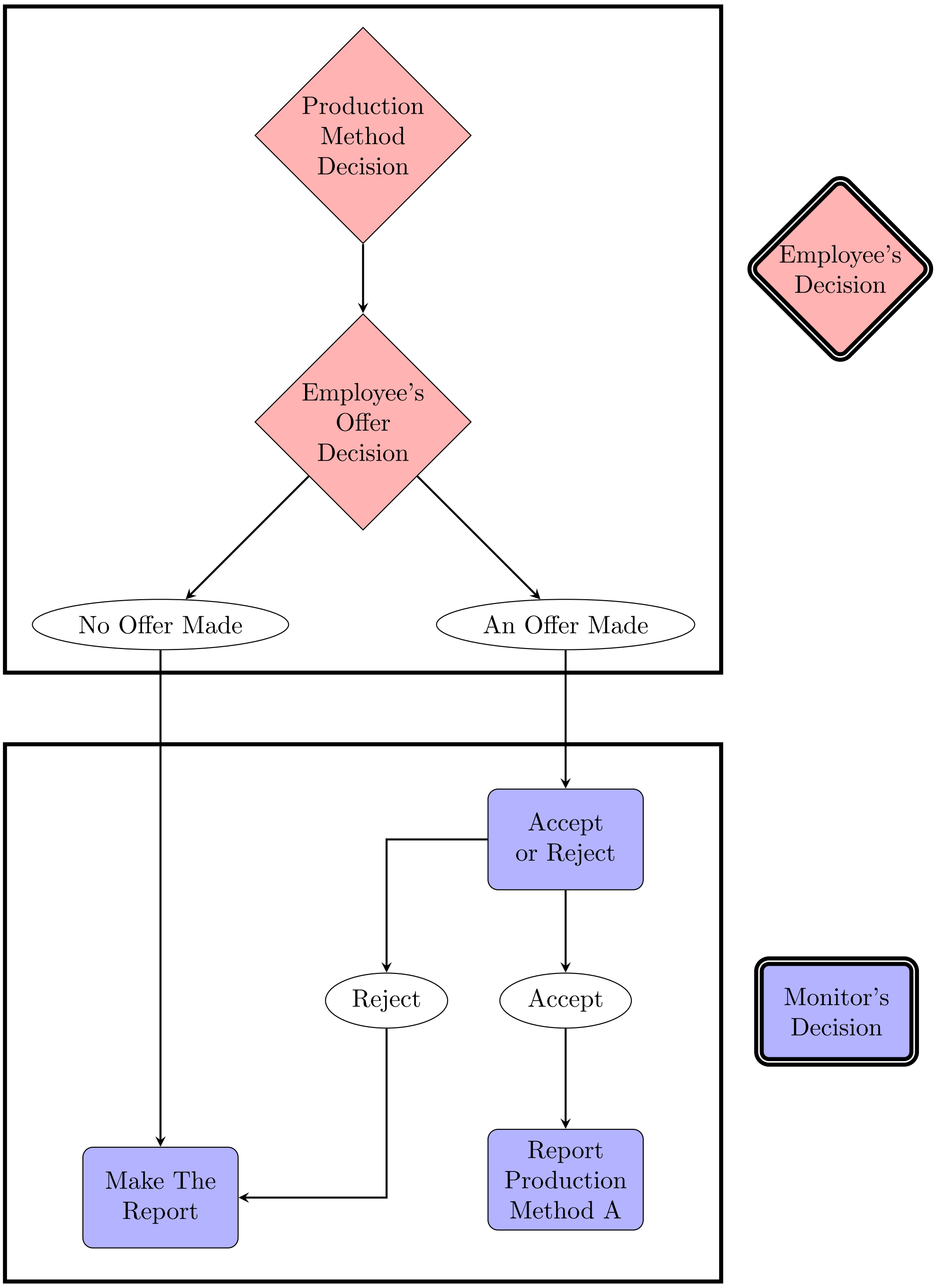
If the Employee chooses production method B, there might be a fine of 40 ECUs imposed on the Employee which depends on the Monitor's report. If the Monitor reports to the corresponding authority that the chosen method is B, the Employee will be fined for 40 ECUs. However, **if the Monitor reports that the chosen method is A**, **no fine will be imposed** even if the authority later on finds out that the Employee's chosen production method is B.

Step 2 Employee's offer decision: The Employee can **make an offer to the Monitor**, so that, upon acceptance of the offer, the Monitor will report that the chosen production method is A regardless of the Employee's actual choice of production method. If the offer is rejected, the Monitor can make the report that s/he would like to make. The offer can be any amount of ECUs that the Employee thinks it is worth to make.

Step 3 Monitor's report decision: No matter whether the Employee makes an offer or not, the Monitor needs to decide what report to make about the Employee's choice of production method. If the reported choice of production method is different from the Employee's actual choice, **the authority will detect it with a 25% chance** and **the monitor's wage is going to be deducted** in this period.

When there is an offer from the Employee, the Monitor first decides whether to accept the offer or not. **If s/he accepts the offer, s/he automatically agree to report that the Employee chosen production method is A regardless of the actual choice**. In addition, if the reported choice of production method is different from the Employee's actual choice and the authority detects it (with a 25% chance), **the monitor's wage is going to be deducted but the monitor can keep the offer received**.

The following chart shows the structure of the game.



As a Monitor, there are three different scenarios that you should consider:

(M_1) The Employee did not make an offer. Suppose the Employee chose production method A (B), you keep your wage for sure if you report A (B), while you keep your wage W with 75% and lose it with 25% if you report B (A).

(M_2) The Employee did make an offer, for example 50 ECUs, and you choose to accept the offer. Thus, you agree to report A. Suppose the Employee chose production method A, you keep your wage W for sure as well as the offer, and you get $W + 50$ in this period. Suppose the Employee chose production method B, you keep your wage W with 75% and lose it with 25%. Together with the offer, you get $W + 50$ ECUs with 75% and 50 ECUs with 25%.

(M_3) The Employee did make an offer, for example 50 ECUs, and you choose to reject the offer. It becomes the same as scenario (M_1) from now on.

As an employee, there are four different scenarios that you should consider:

(E_1) You choose production method A, and you choose do not make any offer. Then, you will get 10 ECUs in this period no matter whether the Monitor reports A or B.

(E_2) You choose production method A, and you choose to make an offer, say 50 ECUs. Then, you will get $10-50=-40$ ECUs if the Monitor accepts the offer, and 10 ECUs if the Monitor rejects the offfer.

(E_3) You choose production method B and get 30 ECUs from it. You further choose not to make any offer. Suppose the Monitor reports A, no fine will be imposed on you and thus you get 30 ECUs in this period (Keep in mind that, by reporting A which is different from your actual choice B, the Monitor suffers a risk of losing his/her wage with a 25% chance). Suppose the monitor reports B, you will be imposed a fine of 40 ECUs and get $30-40=-10$ ECUs.

(E_4) You choose production method B and get 30 ECUs from it. You further choose to make an offer, for example 50 ECUs. Suppose the Monitor accepts the offer, s/he agrees to report A and thus you are not imposed of any fine, and you receive $30-50=-20$ ECUs. Suppose the Monitor rejects the offer, it becomes the same as scenario (E_3) from now on.

There are a few test questions before Part One actually starts. You have to answer all of them correctly in order to proceed. Please raise your hand if you have any questions.

At the end of each period, **your decisions as well as your earnings in this period will be displayed on your screen.** The game will be played repeatedly for 24 periods.

At the end of the experiment, **two out of 24 periods** in this part will be randomly selected and **the sum of your earnings in these two periods as well as the initial wealth (50 ECUs) will be your total earnings in Part One**, which will be added to your final earnings from this experiment when the experiment is completed. Then you will be shown on your screen **your earnings** in each period in Part One as well as the two periods that are selected. Since you do not know which periods are going to be selected, **the best strategy is to take each period equally important.**

Part Two

In Part Two, you are going to play the same game as you've played in Part One for another 24 periods with only one change: the roles are exchanged in this part. **The monitor in Part One becomes the Employee in Part Two, and the Employee in Part One becomes the Monitor in Part Two.** Everything else remains the same, and you can refer to Part One instruction for details.

At the end of each period, **your decisions as well as your earnings in this period will be displayed on your screen.** The game will be played repeatedly for 24 periods.

At the end of the experiment, **two out of 24 periods** in this part will be randomly selected and **the sum of your earnings in these periods as well as the initial wealth (50 ECUs) will be your total earnings in Part Two**, which will be added to your final earnings from this experiment when the experiment is completed. Then you will be shown on your screen **your earnings** in each period in Part Two as well as the two periods that are selected. Since you do not know which periods are going to be selected, **the best strategy is to take each period equally important.**

Part Three

In Part Three, you will be asked to make a series of choices. How much you receive will depend partly on chance and partly on your own choices. The decision problems are not designed to test you. What we want to know is **what choices you would make** in them. The only right answer is what you really would choose.

For each of the ten lines in the table on the computer screen, please state whether you prefer **Option L** or **Option R**. Table 1 below is an example of what you will see on your computer screen later on. Both **Option L** and **Option R** give you either a high amount of ECUs or a low amount of ECUs with different chances in different lines. In **Option L**, the difference between the high amount and the low amount is relatively small, which is $40 - 32 = 8$ (ECUs). By contrast, in **Option R**, the difference between the high amount and the low amount is relatively large, which is $77 - 2 = 75$ (ECUs).

Let's first look at Line 1. **Option L** gives you 40 ECUs with a 10% chance and 32 ECUs with a 90% chance. In other words, you will get 40 ECUs in 1 out of 10 cases and 32 ECUs in 9 out of 10 cases. By contrast, **Option R** gives you 77 ECUs with a 10% chance and 2 ECUs with a 90% chance. Therefore, if you want to stay on the safe side and get either 40 ECUs or 32 ECUs, you can choose **Option L** in Line 1. However, if you want to take the risk and try to get the 77 ECUs with a 10% chance, you would choose **Option R** in Line 1.

When you move from Line 1 to Line 2, the high amount and low amount are the same for each option, but the chances are different. In Line 2, compared against Line 1, the chance of getting the high amount increases by 10%, and the chance of getting the low amount decreases by 10%. Similarly, whenever you go down the table by one line, the chance of getting the high amount increases by 10%, and the chance of getting the low amount decreases by 10%, with Line 10 gives you the high amount with 100% chance and the low amount with 0% chance.

Notice that there are a total of ten lines in the table but **just one line** will be randomly selected for your earning. Since you do not know which line will be paid when you make your choices, **you should pay attention to the choice you make in every line.** After you have completed all your choices, the computer will randomly choose a line to be paid with equal chance of 1/10 for each line.

Your earning for the selected line depends on which option you chose: If you chose **Option L** in that line, you will receive **either 40 ECUs or 32 ECUs** with the **chances stated in Option L in that line**, which will be executed by a computer program. If you chose **Option R** in that line, you will receive **either 77 ECUs or 2 ECUs** with the **chances stated in Option R**, which will also be executed by a computer program.

Your earning in Part Three will be added to your final earnings from this experiment when the experiment

Table B.2: Option Task

Line	Option L (ECUs)	Option R (ECUs)	Your Choice
1	(40 with 10% chance, 32 with 90% chance)	(77 with 10% chance, 2 with 90% chance)	
2	(40 with 20% chance, 32 with 80% chance)	(77 with 20% chance, 2 with 80% chance)	
3	(40 with 30% chance, 32 with 70% chance)	(77 with 30% chance, 2 with 70% chance)	
4	(40 with 40% chance, 32 with 60% chance)	(77 with 40% chance, 2 with 60% chance)	
5	(40 with 50% chance, 32 with 50% chance)	(77 with 50% chance, 2 with 50% chance)	
6	(40 with 60% chance, 32 with 40% chance)	(77 with 60% chance, 2 with 40% chance)	
7	(40 with 70% chance, 32 with 30% chance)	(77 with 70% chance, 2 with 30% chance)	
8	(40 with 80% chance, 32 with 20% chance)	(77 with 80% chance, 2 with 20% chance)	
9	(40 with 90% chance, 32 with 10% chance)	(77 with 90% chance, 2 with 10% chance)	
10	(40 with 100% chance, 32 with 0% chance)	(77 with 100% chance, 2 with 0% chance)	

is completed.

This is the end of the specific instructions for each part.

Final Payoff

For your reference, your **total earnings** in this experiment would be the sum of the following parts:

1. Total earnings of **Two randomly chosen binding periods** and the **initial wealth** in Part One.
2. Total earnings of **Two randomly chosen binding periods initial wealth** in Part Two.
3. Earning in Part Three.

Your total earnings will then be converted to S$ and added to your **show-up fee as your final payoff** in this experiment. You will be paid privately according to your unique subject ID.

Thank you again for your participation!