EUI WORKING PAPERS

MWP 2019/07
Max Weber Programme

Social Norms in Corruption: A Bribery Experiment

Alice Guerra and Tatyana Zhuravleva

European University Institute
Max Weber Programme

Social Norms in Corruption: A Bribery Experiment

Alice Guerra and Tatyana Zhuravleva

EUI Working Paper **MWP** 2019/07

ISSN 1830-7728

Printed in Italy
European University Institute
Badia Fiesolana
I – 50014 San Domenico di Fiesole (FI)
Italy
www.eui.eu
cadmus.eui.eu

Abstract
This paper analyzes social norms in corruption by exploring whether engagement in bribery induces costly third-party punishment. By introducing third-party punishment in a bribery experiment we disentangle social norms from other non-normative motives − such as retaliation, negative reciprocity − maintained merely by parties involved in corruption. We manipulate two main characteristics of bribery: the private benefits gained by corrupt actors, and the negative externality generated by bribery on passive members of society. We find that third parties punish bribers more often than the bribed, but to a lesser extent. Greater private benefits induce third parties to punish more, whereas greater negative externalities have no impact on punishment choices. We unveil the role of emphatic anger as a micro-determinant of third-party punishment. We find gender differences in punishment behavior, with females being more willing to punish the bribed than males, but to a lesser amount.

Keywords
Corruption, Third-party punishment, Experiment, Emphatic anger

JEL
D73, C91, D81, Z13

Acknowledgements
We are grateful to Maria Bigoni, Emanuela Carbonara, Marco Casari, Joep Sonnemans, Werner Raub, as well as seminar participants at the University of Bologna and the participants of the Interdisciplinary Experimental Working Group at the European University Institute for their valuable comments.

Alice Guerra
Copenhagen Business School, Dpt.of International Economics, Government and Business
Tel.: +45 38153537 Fax: +123-45-678910
email: agu.egb@cbs.dk

Tatyana Zhuravleva
Max Weber Fellow, 2017-2018 and 2018-2019
European University Institute, Dpt. of Economics
emai: Tatyana.Zhuravleva@eui.eu

1 Introduction

Social norms play an important role in sustaining corrupt practices (Kubbe and Engelbert, 2017). Since complying with social norms often comes with personal costs, in the absence of any punishment individuals might decide to violate social norms in the pursuit of private gains. The presence of punishment is indeed often invoked to analyze norm enforcement (Fehr and Fischbacher, 2004).

Third-party punishment is a powerful tool to assess whether a certain behavioral pattern qualifies as a social norm (Ohtsubo et al., 2010). Among others, Bendor and Swistak (2001) defined social norms as behavioral rules involving punishment implemented by third parties. Indeed, introducing third parties allows us to disentangle social norms from other behavioral rules – such as retaliation, negative reciprocity, conditional cooperation – sustained by parties directly involved in corruption (Ohtsubo et al., 2010).

While extensively explored in other contexts (e.g., Fehr and Fischbacher, 2004; Charness et al., 2008; Nelissen and Zeelenberg, 2009; Nikiforakis and Mitchell, 2014; Balafoutas et al., 2014), third-party punishment as norm enforcement has never been studied in a corruption framework. Previous bribery experiments have looked at the deterrent effect of punishment by introducing exogenous punishment (e.g., Abbink et al., 2002; Schulze and Franck, 2003), exogenous penalties with the endogenous risk of being punished (e.g., Berninghaus et al., 2013; Serra, 2011), and punishment implemented by second parties, i.e., by victims of the corrupt acts (Cameron et al., 2009).[1] Differently from previous experiments in corruption, this paper introduces punishment not to analyze its deterrent effect on bribery, but to detect social norms, i.e., whether engagement in bribery induces costly third-party punishment.

We design a novel bribery game with the aim of disentangling social norms from any engagement in corruption, as well as from other non-normative behavioral rules such as retaliation, negative reciprocity and conditional cooperation. For this purpose, we introduce third-party punishment à la Fehr and Fischbacher (2004) in a bribery game à la Barr and Serra (2009).

The basic idea of the experiment is the following. Each participant is randomly assigned to one of the following roles: private citizens, public officials, other members of society, monitors. As the simplest petty corruption situation, each citizen can offer a bribe to a public official, who can accept or reject it. Accepting the bribe implies delivering a corrupt service to the citizen at the expense of all the other members of society. Similarly to Cameron et al. (2009) and Barr and Serra (2009), the other members are passive victims: they cannot do anything during the experiment, but they suffer a monetary loss for each corrupt transaction.

The key original feature of our design is represented by the role of monitors, who act as impartial third-party observers. They cannot engage in corruption and their payoff is not affected by corrupt acts, but they can react to corrupt behaviors by expending their own resources to impose a greater cost on the corrupt actor/s.

Monitoring by humans has been introduced by Azfar and Nelson (2007) – whose

[1] For a comprehensive review of the literature on endogenous and exogenous punishment in corruption experiments, see Abbink and Serra (2012).

design has been later used by Barr et al. (2009) – to analyze corruption in public service delivery. Azfar and Nelson (2007) and Barr et al. (2009) designed embezzlement games with eight different roles including attorney generals, who act as external monitors either appointed by principals (e.g., governmental executives or service providers) or selected by agents (e.g., service recipients). Differently from Azfar and Nelson (2007) and Barr et al. (2009) our monitors are observers with neither an institutional role, nor *a duty* of monitoring. They are neither appointed nor elected, but are randomly selected among participants and provided with *a possibility* to report corrupt behaviors at their own expense. In our experiment, the role of monitors is not meant to mimic any professional, institutional role, but to cleanly measure social norms in corruption. Our monitors can be regarded to as impartial witnesses of corrupt transactions, who are not directly involved in corrupt acts but can spend time and effort to report them. Consider the example of a hospital patient who requires a change of linen and offers a bribe to a nurse in exchange of having it done immediately, at the expense of other patients who will have to wait more time (Barr and Serra, 2009). Imagine that an external, impartial person observes this situation and can report it at his own personal cost (e.g., time and effort). Our paper investigates precisely this: whether and to what extent someone who observes an illicit act as bribery is willing to spend his own resources to punish it.

To better understand the determinants of social norms, we manipulate two main elements of corruption transactions: the private benefits to the bribers and the bribed, and the negative externality imposed on the other members of society. Some previous contributions have already analyzed the effects of negative externalities on bribing behaviors (Abbink et al., 2002; Cameron et al., 2009; Barr and Serra, 2009; Chen et al., 2016b). The results are mixed: while Abbink et al., 2002 and Büchner et al. (2008) found no evidence of externality effects on individuals' decisions to engage in bribery, Barr and Serra (2009) did find an effect, i.e., that greater negative externalities reduce individuals' propensity to engage in bribery.[2]

Finally, in an attempt to understand the micro-determinants of punishment behaviors, we asked participants to answer some questions about their emotions towards corrupt actors. Feelings of anger have been suggested to drive costly punishment (Frank, 1988; Elster, 1989), and some contributions provided empirical support to this idea. While numerous studies have tested the association between moral emotions and second-party punishment (Bosman and Van Winden, 2002; Ben-Shakhar et al., 2007; Pillutla and Murnighan, 1996; Sanfey et al., 2003), very little is known about the role of emotions in third-party punishment. With the exception of Fehr and Gächter (2002), Fehr and Fischbacher (2004) and Nelissen and Zeelenberg (2009), which show that negative emotions can drive third-party punishment in standard cooperation games, no previous contributions to our knowledge explored the role of anger in a bribery game. Emphatic anger, i.e., third parties' anger at witnessing injustice or negative externalities to someone else, has been defined as a distinct emotion from direct anger, i.e., victims' anger about personal harm (Batson et al., 2007). We contribute to the literature on moral emotions in two ways. First, we analyze feelings of anger as a micro-determinant of third-party punishment. Secondly, we test whether third-parties'

[2]For an extensive review and discussion on these effects, see Barr and Serra (2009) and Abbink and Serra (2012).

emphatic anger is different from victims' direct anger.

The rest of the paper is organized as follows. Section 2 presents our experimental design and hypotheses. Section 3 describes the experimental procedure and introduces our participants. Section 4 reports the results. Section 5 presents additional analysis on the role of emotions and robustness checks. Section 6 discusses our findings and draws conclusions.

2 Experimental design, treatments and hypotheses

We first present our experimental design, treatments and parameters, and secondly we develop theoretical predictions. Finally, we describe the post-experiment questionnaire.

2.1 The bribery game

Building on Barr and Serra (2009), our experimental design consists of a four-player, sequential-move game which mimics a petty corruption situation. At the beginning of the experiment, participants are randomly and anonymously matched in N groups of four players. In each group, the following roles are randomly assigned: a citizen, a public official, other member of society, and a monitor.

All players receive an initial endowment E. The first player – who acts as a citizen – can offer to pay a bribe, B, to the public official, in exchange for a corrupt service, the value of which to him is αB, with $\alpha > 1$. If the citizen does not offer a bribe, his final payoff equals E. If the citizen does offer a bribe, his final payoff depends upon the public official's and the monitor's decisions.

The second player – who acts as a public official – is informed about the citizen's decision and, if a bribe is offered he can either accept or reject it. If the public official accepts the bribe, he automatically has to supply the corrupt service to the citizen at the expense of the other members of society. In this case, the citizen's and the public official's payoffs increase each by the value of the corrupt service, that is αB, with $\alpha > 1$. The payoff increase that the citizen experiences may represent, for example, the provision of a better and/or quicker service, a work promotion, an exemption from a regulation or from certain job tasks. The payoff of the public official increases as well by αB, in addition to the bribe offered by the citizen, B. The rationale for this design choice is to ensure mutual gains from corruption for both the citizen and the public official, while avoiding excessively unequal payoffs.[3] If the public official rejects the bribe, he gives back the amount B to the citizen and does not supply any corrupt service.

The third player – who acts as another member of society – is a passive victim: he cannot take any actions, but incurs a cost γB, with $\gamma \geq 1$, for every bribe offered

[3] In Abbink et al. (2002) and Abbink and Hennig-Schmidt (2006), the bribe is tripled before passing on to the public official. The rationale for this multiplication is that the marginal utility for any given bribe is likely to be greater for a public official than for a business person, given that the latter are generally richer. This rationale may also be applied to petty corruption contexts, but there is only little evidence that public server providers are poorer than service recipients (Barr et al., 2009; Abbink and Serra, 2012).

and accepted.[4]. Each other member of society's final payoff is $E - n\gamma B$, where $n \in \{0, \ldots, N\}$ is the number of citizen-official pairs who offer and accept bribes.[5]

The key original feature of our design is represented by the fourth player, who acts as a monitor. The monitor receives information about the choices of the citizen and the public official of his group and he can punish, at his own personal cost, either the citizen, or the public official, or both. Specifically, if a bribe has been offered, the monitor can choose either not to punish, or to spend an amount $p_C > 0$ to punish the citizen, and $p_O > 0$ to punish the public official. The punishment technology is linear and imposes a monetary sanction on the citizen and public official by reducing their payoff by mp_C and mp_O, with $m > 1$, respectively.

The final payoffs are computed as following. The citizen's final payoff equals: E if he does not offer a bribe; $E - B + \alpha B$ if he offers a bribe that is accepted, minus mp_C if he is punished; E if he offers a bribe that is rejected, minus mp_C if he is punished. The public official's final payoff is: E if he is not offered a bribe; E if he is offered a bribe but rejects it, minus mp_O if he is punished;[6] and $E + B + \alpha B$ if he accepts a bribe, minus mp_O if he is punished. Each other member of society's final payoff is $E - n\gamma B$. The monitor's final payoff equals E, minus p_C and/or p_O if he punishes the citizens and/or the public official.

Figure 1 represents the structure of the game. In equilibrium, with standard preferences where all players behave as selfish money-maximizers, the game has the following Nash equilibrium: the citizen offers the bribe, the public official accepts the bribe, the monitor does not to punish.

The game consists of ten identical and independent rounds, whereby the equilibrium does not vary with the number of rounds the game is played. The roles remain fixed throughout the experiment, whereas in each round, participants are randomly rematched to different groups.

To cancel out any learning effects from punishment, after each round neither the citizen, nor the public official, nor the other member of society were informed about the monitors' choices. Moreover, to rule out any peer effect and any bias from knowing the results of the previous round, monitors received no feedback about the other monitors' choices.

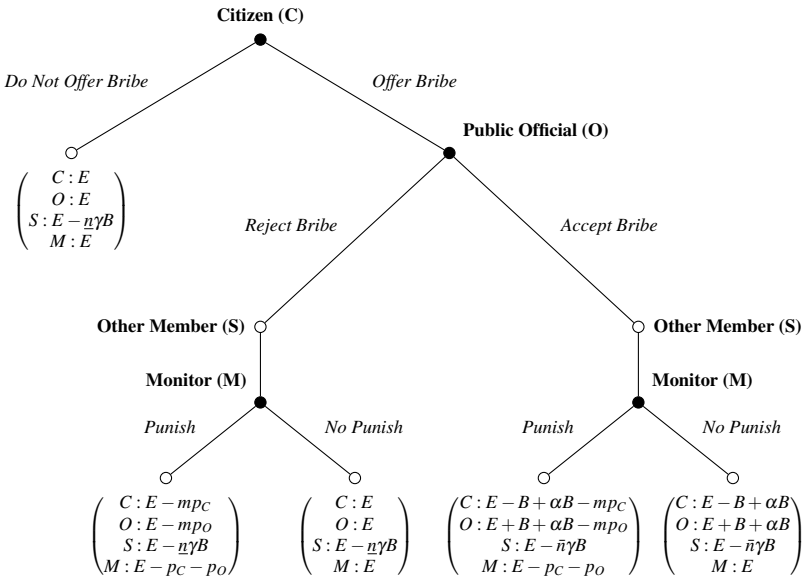
The repetition of rounds allows us to accurately evaluate social norms in corruption under different circumstances. At the same time, the combination of strangers random-matching, the independent nature of rounds, and the absence of any feedback to participants allow us to avoid issues related to signaling, conditional cooperation,

[4]While waiting for the other players' decisions, the other members of society are asked to answer some questions with no consequences on their final payoff. This is meant not only to keep them paying attention to the instructions, but also to capture their emotions towards corrupt actors. See Section 5.

[5]This design choice is similar to Cameron et al. (2009) and Barr and Serra (2009). As pointed out by Barr and Serra (2009), the presence of a group of passive victims could introduce interdependence between individuals' choices within sessions. Some individuals may – or may not – engage in corruption conditional on their expectations concerning the number of other participants choosing to act corruptly. Cameron et al. (2009) avoided that interdependence issue by letting each act of bribery affect only one of the other members of society. We decided to follow Barr and Serra (2009), since we aim at mimicking real-world cases where bribery harms many individuals, and not only one.

[6]We deliberately chose this design feature to check the consistency of monitors' behavior.

Figure (1) The bribery game



Notes: E is the initial endowment; B is the size of the bribe; p_C and p_O are the amount chosen by the monitor to punish the citizen and the public official respectively; $m > 1$ is the punishment multiplier; $n \in [0,\ldots,N]$ is the total number of bribes accepted and offered, with $\underline{n} \in [0,\ldots,N-1]$ and $\bar{n} \in [1,\ldots,N]$; $\alpha > 1$ and $\gamma \geq 1$ reflect the treatment conditions.

reputation formation and negative reciprocity.[7]

Since the aim of our experiment is to mimic a real-life petty corruption context, similarly to Barr and Serra (2009) and Cameron et al. (2009) we deliberatively chose to use loaded terms in the instructions, such as "bribe" and "punishment."[8]

2.2 Treatments and parametrization

We used a 2X2 between-subject design where we varied the magnitude of private benefits, α – setting it either low, $\alpha_L = 3$, or high, $\alpha_H = 6$ – and the magnitude of negative externalities, γ – setting it either low, $\gamma_L = 1$, or high, $\gamma_H = 2$. The four treatments are presented in Table 1.

In our experiment we used a fictitious currency called token (1 token = 0.20 Euro), and set $E = 50$ tokens, $B = 3$ tokens, $p_C \in (0,10]$, $p_O \in (0,10]$, and $m = 2$. We chose

[7]For a discussion on this design choice, see Abbink and Serra (2012). See also Duffy and Ochs (2009), who conducted an experiment with a sequence of indefinitely repeated 2-player prisoner's dilemma games. Their results show that, contrary to the theoretical predictions (Kandori, 1992), a cooperative norm does not emerge in the treatments where players are matched randomly, whereas it does emerge in the fixed-pairing treatments.

[8]Abbink and Hennig-Schmidt (2006) and Serra (2011) show that framing the game as bribery did not make any difference in individuals' behavior with respect to using neutral wordings.

Table (1) Experimental design: treatments

	Low benefit	High benefit
Low externality	*Treatment 1* If B is accepted, the citizen and the public official gain $\alpha_L B$ each. Per each B accepted, each other member of society suffers $\gamma_L B$.	*Treatment 2* If B is accepted, the citizen and the public official gain $\alpha_H B$ each. Per each B accepted, each other member of society suffers $\gamma_L B$.
High externality	*Treatment 3* If B is accepted, the citizen and the public official gain $\alpha_L B$. each. Per each B accepted, each other member of society suffers $\gamma_H B$.	*Treatment 4* If B is accepted, the citizen and the public official gain $\alpha_H B$. each. Per each B accepted, each other member of society suffers $\gamma_H B$.

these parameters to avoid negative payoffs and excessively unequal payoffs between players.

2.3 Hypotheses with reference to moral costs and beliefs

In Section 2.1 we posted the Nash equilibrium of the bribery game assuming that subjects are motivated solely by their own payoffs – that is, offer and accept bribes, and no punishment.

However, several contributions have already suggested that people may suffer moral costs when they observe behavior that breaks social norms (e.g., Abbink et al., 2002; Barr and Serra, 2009; Gneezy et al., 2018). Moral costs may discourage people from engaging in dishonest behaviors, thereby reducing their tendency to pursue their own private interests at the expense of someone else or society as a whole (e.g., Gneezy et al., 2018). Intrinsic motivations (e.g., feelings of shame and guilt) generally result from the internalization of social norms (for a literature review, see Abbink and Serra, 2012).

In this section, we build on Barr and Serra's (2009) theoretical framework to illustrate how the inclusion of moral costs in our bribery game may affect the Nash equilibrium.

We assume that a monitor may suffer intrinsic moral costs when observing bribery (T_M). These may consist in a moral cost when observing a citizen offering a bribe (T_M^C), and/or a moral cost when observing a public official accepting the bribe (T_M^O). If bribing is *not* a social norm for this monitor, $T_M^C > 0$ and $T_M^O > 0$.

We also assume that a monitor's intrinsic moral costs are affected by the parameters

of our design, i.e., private benefit (α) and negative externality (γ). Formally, $T_M^C = T_M^C(\alpha,\gamma)$ and $T_M^O = T_M^O(\alpha,\gamma)$.

For instance, $dT_M^C/d\alpha > 0$ means that if a monitor observes a greater private benefit of a corrupt citizen he experiences a higher level of intrinsic moral cost; $dT_M^O/d\alpha > 0$ represents the analogous situation vis-à-vis public official. $dT_M^C/d\alpha = 0$ means that the monitor's moral cost is not affected by the level of private benefit; $dT_M^O/d\alpha = 0$ represents the analogous situation upon observing the public official. Similarly, $dT_M^C/d\gamma > 0$ means that if a monitor observes a greater negative externality as a result of a citizen's behavior he experiences a higher level of intrinsic moral cost; $dT_M^O/d\gamma > 0$ represents the analogous situation vis-à-vis the public official. $dT_M^C/d\gamma = 0$, $dT_M^O/d\gamma = 0$ means that the monitor's moral costs are not affected by negative externalities.

For the sake of simplicity we assume a linear relationship between punishment and intrinsic moral costs, i.e., $p_C = T_M^C$, $p_O = T_M^O$. The monitor decides to punish *only* if he experiences intrinsic moral costs when observing bribing behavior. This implies that, if $dT_M^C/d\alpha > 0$, $dT_M^O/d\alpha > 0$ and $dT_M^C/d\gamma > 0$, $dT_M^O/d\gamma > 0$, an increase in either α or γ leads to an increase in the monitors' punishment.

A public official who causes a loss to others by accepting a bribe may suffer a similar intrinsic moral cost, $T_O = T_O(\alpha,\gamma)$. Having $dT_O/d\alpha > 0$ and $dT_O/d\gamma > 0$ means that the public official's moral cost increases with greater private benefit and greater negative externality. His expected utility becomes E if he does not accept the bribe and to $E + (1+\alpha)B - T_O - \hat{T}_M^O$ if he does accept the bribe, where $\hat{T}_M^O$ is the public official's belief s/he will be punished by the monitor.

A public official accepts the bribe if, omitting arguments,

$$T_O + \hat{T}_M^O < (\alpha+1)B$$

For a given α, if the official's intrinsic costs and/or his beliefs about the monitor's intrinsic costs are sensitive to negative externalities, an increase in γ reduces the public officials' incentive to accept the bribe. Similarly, for a given γ, if $dT_O/d\alpha > 0$ and $d\hat{T}_M^O/d\alpha > 0$, the public officials accept the bribe only if the marginal benefit from an increase in α (i.e., B) is greater than the related increase in intrinsic costs (i.e., $T_O/d\alpha$) and in the belief s/he will be punished by the monitor ($d\hat{T}_M^O/d\alpha$).

If a public official does not experience intrinsic moral cost to accept a bribe $T_O = 0$ and believes that a monitor does not experience it either when observing his behavior ($\hat{T}_M^O = 0$), he will always accept a bribe.

Lastly, a private citizen who causes a loss to others by offering a bribe may suffer $T_C = T_C(\alpha,\gamma)$, with $dT_C/d\alpha > 0$, $dT_C/d\gamma > 0$ if he cares about private benefit and negative externality. The citizen's expected utility function is equal to E if he does not offer any bribe and to $E + (\alpha-1)\hat{B} - T_C - \hat{T}_M^C$ if he does offer a bribe, where $\hat{B} = \hat{B}(\hat{T}_O)$ is the citizen's belief whether the bribe will be accepted by the public official, and $\hat{T}_M^C$ is the citizen's belief s/he will be punished by the monitor.

The citizens offer a bribe if, omitting arguments,

$$T_C + \hat{T}_M^C < (\alpha-1)\hat{B}$$

For a given α, if $dT_C/d\gamma > 0$ and/or $d\hat{T}_M^C/d\gamma > 0$ and/or $d\hat{T}_O/d\gamma > 0$, an increase in γ reduces the citizens' incentive to offer a bribe. For a given γ, if $dT_C/d\alpha > 0$ and/or

$d\hat{T}_M^C/d\alpha > 0$ and/or $d\hat{T}_O d\alpha > 0$, the citizens offer a bribe only if the marginal benefit from an increase in α (i.e., $\hat{B}$) is greater than the marginal decrease in the believes that the public official accepts the bribe (i.e., $(\alpha - 1)d\hat{B}/d\alpha$), the marginal increase in intrinsic costs (i.e., $dT_C/d\alpha$), and in the belief s/he will be punished by the monitor (i.e., $d\hat{T}_M^C/d\alpha$).

If a citizen does not experience intrinsic moral cost to offer a bribe ($T_C = 0$) and believes that both public official and monitor do not have intrinsic moral cost for bribing behavior either ($\hat{T}_O = 0$ and $\hat{T}_M^C = 0$), he will always offer a bribe.

Based on this setup and assuming that agents' beliefs are correct, we formulate the following hypotheses.

Hypothesis 1. Intrinsic moral costs and beliefs are heterogeneous across citizens and public officials.

Hypothesis 2. $dT_C/d\gamma > 0$ and $dT_O/d\gamma > 0$, i.e., in the high negative externality treatment fewer bribes will be offered and fewer bribes accepted.

Hypothesis 3. $dT_C/d\alpha > 0$ and $dT_O/d\alpha > 0$, but the marginal benefit from the increase of private benefit outweighs the increase in the intrinsic moral costs, i.e., in the high private benefit treatment more bribes will be offered and more bribes accepted.

Hypothesis 4. Intrinsic moral costs are heterogeneous across monitors.

Hypothesis 5. Observing greater private benefit and greater negative externality, monitors will experience a higher level of intrinsic moral costs towards both citizens ($dT_M^C/d\alpha > 0$; $dT_M^C/d\gamma > 0$) and public officials ($dT_M^O/d\alpha > 0$; $dT_M^O/d\gamma > 0$).

Hypothesis 6. Monitors do not differentiate between citizens and public officials in terms of intrinsic moral costs, i.e., they experience the same costs observing citizens who offer a bribe and officials who accept it: $T_M^C = T_M^O$ and $dT_M^C/d\alpha = dT_M^O/d\alpha$, $dT_M^C/d\gamma = dT_M^O/d\gamma$.

Hypothesis 7. Self-reported level of anger is a proxy of intrinsic moral costs. The results obtained from hypotheses 1-6 hold when considering the relationship between anger, punishment and treatment conditions.

2.4 Survey Questions

Our experiment includes an extensive post-experiment survey that allows us to better understand participants' choices in the bribery game, as well as to capture their feelings towards corrupt actors.[9]

Our survey is composed of two parts. In the first part – which was conducted after each round of the bribery game – we asked participants in the role of monitors and other members of society to express their feelings towards participants in the role of citizens and public officials. Specifically, we asked them the following question: "Imagine

[9]The survey questions are available upon request.

you casually meet the citizen / public official of this round on the street. What are your feelings towards this person? Please indicate the intensity of your anger and annoyance towards this citizen / public official on a scale of 1 to 10, where 1 = not at all angry and annoyed, and 10 = very angry and annoyed." We analyze the role of emotions in punishment behavior in Section 5.1.

In the second part of the survey – which was conducted after all participants completed the ten rounds of the bribery game – we asked participants to answer a set of questions about demographic information (gender, occupation, place of birth, education, field of study, risk preferences, religion). This information will be used as control variables in the econometric analysis in Section 4.

3 Procedure and participants

We conducted eight sessions (two sessions per treatment) with 32 subjects each. All sessions were run at the Bologna Laboratory for Experiments in Social Science (BLESS) at the University of Bologna on May 2-3 and May 6-7, 2019. Each session lasted approximatively 70 minutes in total. Each subject participated in only one session, whereby the order of sessions was randomized to control for session effects (cf: Fréchette, 2012). Immediately after each session, subjects were paid anonymously in cash. The exchange rate was set to 0.20 Euros per token. Subjects' earnings included a show-up fee of 5 Euros and ranged from 11 Euros to 18.40 Euros, with an average of 14.56 Euros.

The experiment was computerized using oTree (Chen et al., 2016a), and participants performed all the tasks via computer. Computer stations at BLESS are partitioned to ensure confidentially and avoid any communication between participants. Upon arrival at the laboratory, all the subjects were welcomed and randomly directed to separate workstations. The written instructions were handed out to the subjects, the experimenter read the instructions out loud and answered any questions on an individual basis. In order to ensure against bias, the same experimenter and laboratory staff conducted all sessions.

To make sure that all the subjects understood the instructions, a computer-based quiz with eight comprehension questions was conducted before starting the experiment, with direct feedback and explanations in case of an incorrect answer. Most of the participants answered correctly. Translations of the instructions to participants are provided in Appendix B.[10]

Participants were recruited using ORSEE (Greiner, 2015) among graduate and undergraduate students which met the following criteria: being an Italian citizen, born of Italian parents in the North of Italy.[11] These recruitment filters were applied to rule out any cultural and socio-economic differences in corrupt behaviors that previous con-

[10]The comprehension questions, the post-experiment questionnaire, and the original texts in Italian are available on request.

[11]By following the definition of the Italian geographical areas used by the Italian National Institute of Statistics, the North if Italy includes the following regions: Piemonte, Valle d'Aosta, Liguria, Lombardia, Trentino-Alto Adige, Veneto, Friuli-Venezia Giulia, Emilia-Romagna. See http://dwcis.istat.it/cis/docs/4-8.htm; last access: July 2019.

tributions proved to exist between Italy and other countries (e.g., Treisman, 2000), as well as between Italian regions (e.g., Del Monte and Papagni, 2007).

Most participants were undergraduate students (48.83% with a high school diploma, 33.98% with a Bachelor's degree, 14.06% with a Master's degree),[12] from different fields of study (21.48% arts and humanities studies, 17.58% engineering sciences, 12.89% economics, 9.77% political science, 10.55% physics and mathematics, 8.98% medical sciences, 1.56% sociology). The ages of the participants ranged from 19 to 50 years old, with the average age being 24 years and most participants (90%) being younger than 28 years old. Regarding gender, 51.56% of the students were males. Most subjects declared to be atheists (61.33%), while 36.33% declared to be Catholic.[13]

According to Chi-squared tests, none of the individual characteristics varied significantly at the 5% significance level neither across assigned roles, nor across treatments.

4 Results

The presentation of the results is divided into two parts. In Section 4.1 we analyze bribing behaviors, i.e., citizens' decisions to offer the bribe and public officials' decisions to accept it. In Section 4.2 we investigate monitors' decisions to punish either the citizen, or the public official, or both, and the amount of punishment.

4.1 Bribing behaviors

In this section we report the effects of changing the magnitude of the private benefits and negative externalities on the behaviors of the citizens and the public officials. Table 2 reports the descriptive statistics of bribing behavior. Overall, pooling across treatments, 52% of citizens (333 over 640) offered a bribe, and 70% of public officials (234 over 333) accepted one if it was offered.

Result 1 intrinsic moral costs and beliefs are heterogeneous across citizens and public officials. For 48% of citizens intrinsic moral costs and beliefs are high enough to refrain from offering a bribe. For 30% of public officials intrinsic moral costs and beliefs are high enough to reject the bribe.

Table 2 reveals that subjects are sensitive to treatment effects. For instance, public officials do significantly react to greater benefits by increasing bribe acceptance when negative externalities are low (62.1% in the 'low externality, low benefit' treatment *versus* 79.6% in the 'low externality, high benefit' treatment; Chi-squared test p-value = 0.008).

To further analyze treatment effects, in Figure 2 we show the frequency of bribes offered by citizens (black bars) and bribes accepted by public officials (white bars) with observations pooled across the low and high private benefit treatments (T1&T3 *versus* T2&T4, Panel (a)) and across the low and high negative externality treatments (T1&T2 *versus* T3&T4, Panel (b)).

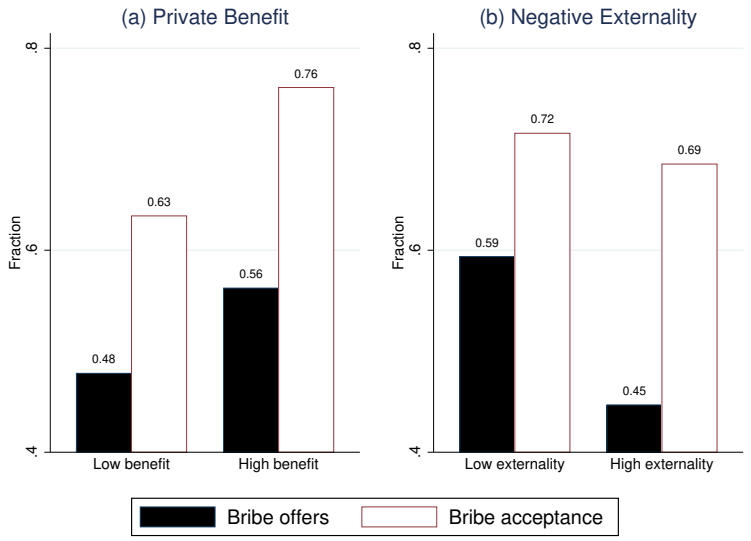
[12] Six participants (2.34%) had a second level master and one participant (0.39%) had a PhD degree.

[13] One participant declared to be Muslim (0.39%), another Buddhist (0.39%), and four participants declared to have another religious denomination (1.56%).

Table (2) Bribing behavior: descriptive statistics

Treatment	Bribe offer			Bribe acceptance		
	Freq.	Perc.	Obs.	Freq.	Perc.	Obs.
Low externality, low benefit (T1)	87	54.4%	160	54.0	62.1%	87
Low externality, high benefit (T2)	103	64.4%	160	82.0	79.6%	103
High externality, low benefit (T3)	66	41.3%	160	43.0	65.2%	66
High externality, high benefit (T4)	77	48.1%	160	55.0	71.4%	77
Obs.	333	52%	640	234	70%	333
Tests	Chi-squared test			Chi-squared test		
	Coeff.	*p*-value		Coeff.	*p*-value	
T1 = T2	3.316	0.069		7.135	0.008	
T3 = T4	1.529	0.216		0.649	0.420	
T1 = T3	5.523	0.019		0.154	0.695	
T2 = T4	8.584	0.003		1.623	0.203	
T1&T2 = T3&T4	13.829	0.000		0.362	0.547	
T1&T3 = T2&T4	4.563	0.033		6.397	0.011	

Figure (2) Bribing behavior: treatment differences



Panel (a) of Figure 2 shows that when the private benefits are greater, more citizens offer bribes (47.81% in the low benefit treatments *versus* 56.25% in the high benefit treatments; chi-squared test *p*-value = 0.033) and more public officials accept bribes (63.40% in the low benefit treatments *versus* 76.11% in the high benefit treatments; chi-squared test *p*-value = 0.011). Panel (b) of Figure 2 shows that when the externalities are greater, fewer citizens offer bribes (59.38% in the low externality treatments *versus* 44.69% in the high externality treatments; chi-squared test *p*-value = 0.000) and

fewer public officials accept bribes. This latter effect is, however, not statistically significant (71.58% in the low externality treatments *versus* 68.53% in the high externality treatments; chi-squared test p-value = 0.547).

The regression analysis confirms these findings. In Table 3 we report the marginal effects of probit regressions for the citizens' decision to offer a bribe (columns 1, 2) and for the public officials' decision to accept a bribe (columns 3, 4). In addition to treatment dummies (columns 1 and 3), the econometric analyses account for individual characteristics (columns 2 and 4). In particular, we control for gender, age, main field of study, risk aversion, religiosity and religious denomination. More specifically, we include the dummy variable *Economists* for subjects whose main field of study is Economics. Risk aversion was measured in the post-experiment survey where subjects were supposed to estimate whether they take risks or try to avoid them. Following Armantier and Boly (2011), religiosity was measured in the post-experiment survey by asking subjects how frequently they attend religious services, apart from weddings, funerals or christenings. The variable *Religiosity* has six categories, ranging from one (more than once a week) to seven (never). We also include the dummy variable *Catholic* for those subjects who declared themselves to be Catholic (36.33%, against 61.33% who declared themselves to be atheist or agnostic). Table A1 in Appendix A reports the marginal effects of each control variable.

Table (3) Bribing behavior: probit regressions

	Bribe offers		Bribe acceptance	
	(1)	(2)	(3)	(4)
'High benefit' treatment	0.0841**	0.0604**	0.125***	0.113***
	(0.0346)	(0.0296)	(0.0316)	(0.0287)
'High externality' treatment	-0.145***	-0.174***	-0.0313	-0.0883**
	(0.0312)	(0.0322)	(0.0376)	(0.0425)
Controls	NO	YES	NO	YES
Observations	640	640	333	333

Standard errors in parentheses, clustered by rounds.

Marginal effects are reported.

Control variables: Female, Age, Economists, Risk aversion, Religiosity, Catholic.

See Table A for the marginal effects of each control variable.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Consistent with our statistical tests, Table 3 shows that increasing private benefits significantly increases both the probability of offering a bribe by 8.4% (6% if considering Column 2) and the probability of its acceptance by 12.5% (11.3% if considering Column 4). In line with Barr and Serra (2009), our estimates also show that a greater negative externality significantly reduces the probability of offering a bribe by 14.5 (17.4% if considering Column 2), and has a negative, but smaller effect on the probability of accepting one (8.8%).

This leads to the following results.

Result 2 In the 'high negative externality' treatment, fewer bribes are offered and fewer bribes are accepted.

Result 3 In the 'high private benefit' treatment, more bribes are offered and more bribes are accepted.

We can interpret these results under the lens of moral costs, as described in Section 2.3. Result 2 reveals that the increase in negative externality increases the intrinsic moral costs of both citizens and public officials, which in turn reduces citizens' and officials' incentives for bribing behavior. The increase in private benefits generates two opposite incentives: increasing personal monetary payoffs *versus* mitigating intrinsic moral costs. Result 3 reveals that the incentives to increase personal monetary payoffs outweigh the corresponding increase in intrinsic moral costs.

Table A1 in Appendix A shows some micro-determinants of bribing behavior. In line with Barr and Serra (2009), and in contrast with Frank and Schulze (2000) and Rivas (2013), we find no robust gender differences in bribing behaviors: women are not less corruptible than men. In addition, the probability of bribe offers decreases approximatively by 2.5% for less religious subjects (i.e., those who declared that they do not attend religious services frequently). When considering the probability of bribe acceptance, religiosity remains statistically significant but with an opposite sign: in line with Barr and Serra (2009), less religious subjects are 4.7% more likely to accept a bribe. Being Catholic has no significant impact on bribing behaviors. In line with Frank et al. (1996) and Frank and Schulze (2000), Economists are 12.5% more likely to offer bribes and 26.1% more likely to accept them.

4.2 Third-party punishment

In this section we report the results about monitors' choices. Recall that in our design, monitors – who act as third parties – were given the opportunity to punish either the citizen or the public official or both whenever a bribe was offered, which in our sample corresponds to 333 times over 640.[14]

We analyze third-party punishment mainly by two variables: the *frequency of punishment* – which measures monitors' likelihood to punish – and the *size of punishment* – which measures the level of monitors' expenditure in punishment, conditional upon participation in punishment.

Table 4 reports the frequency of punishment in Panel (A), and the size of punishment in Panel (B) – i.e., the average tokens spent by monitors – if the punishment took place. Overall, pooling across treatments, slightly more than a half of the monitors decided to punish citizens if they offered a bribe and to punish public officials if they accepted the bribe. Those monitors who punished spent on average 5.09 tokens to punish citizens and 5.68 tokens to punish public officials, which in total corresponds to 25% of the maximum amount they could spend to punish each role (i.e., maximum 20 tokens).

[14]This holds also when a bribe was offered but not accepted (see Section 2.1). In our data, only one monitor punished a public official by 1 token when a bribe was offered and not accepted.

Table (4) Frequency and size of punishment: descriptive statistics

PANEL (A): FREQUENCY OF PUNISHMENT

Treatment	Punish citizens			Punish officials		
	Freq.	Perc.	Obs.	Freq.	Perc.	Obs.
Low ext., low ben. (T1)	37	42.5%	87	27	50.0%	54
Low ext., high ben. (T2)	68	66.0%	103	55	67.1%	82
High ext., low ben. (T3)	26	39.4%	66	14	32.6%	43
High ext., high ben. (T4)	47	61.0%	77	32	58.2%	55
Pooled sample	178	53.5%	333	128	54.7%	234
Tests	Chi-squared test			Chi-squared test		
	Coeff.	p-value		Coeff.	p-value	
T1 = T2	10.526	0.001		9.615	0.002	
T3 = T4	6.6629	0.010		6.742	0.009	
T1 = T3	0.152	0.696		1.849	0.174	
T2 = T4	0.474	0.491		2.473	0.116	
T1&T2 = T3&T4	0.582	0.445		4.164	0.041	
T1&T3 = T2&T4	17.146	0.000		16.209	0.000	

PANEL (B): SIZE OF PUNISHMENT

Treatment	Punish citizens			Punish officials		
	Mean	Std. Dev.	Obs.	Mean	Std. Dev.	Obs.
Low ext., low ben. (T1)	3.973	2.061	37	4.815	2.961	27
Low ext., high ben. (T2)	5.441	3.533	68	5.967	3.415	55
High ext., low ben. (T3)	4.346	2.134	26	5.428	2.593	14
High ext., high ben. (T4)	5.872	2.763	47	6.031	3.431	32
Pooled sample	5.09	2.961	178	5.68	3.248	128
Tests	t-test (Welch's formula)			t-test (Welch's formula)		
	Diff.	Std. Err.	p-value	Diff.	Std. Err.	p-value
T1 = T2	-1.468	0.546	0.008	-1.149	0.732	0.122
T3 = T4	-1.526	0.581	0.011	-0.602	0.921	0.517
T1 = T3	-0.373	0.538	0.491	-0.614	0.897	0.499
T2 = T4	-0.431	0.588	0.465	-0.067	0.761	0.929
T1&T2 = T3&T4	0.405	0.437	0.356	0.262	0.594	0.659
T1&T3 = T2&T4	1.490	0.399	0.000	0.964	0.572	0.095

In the post-experiment questionnaire we asked citizens and public officials their expectations about monitors' decisions. Our data reveals that almost 65% of them believed that monitors decided on average not to punish them at all, while actually in ten bribery game 53.5% of citizens and 54.7% of public officials has been punished.

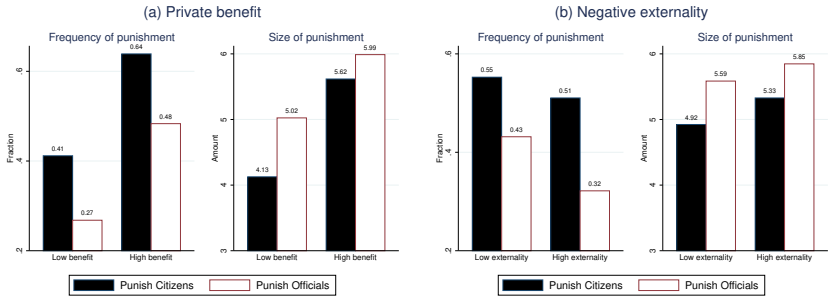
Result 4 Intrinsic moral costs are heterogeneous across monitors. Pooling across treatments, approximatively half of monitors decided to punish citizens if they offered a bribe and to punish public officials if they accepted the bribe. Those monitors who punished spent on average 25% of the maximum amount they could spend

to punish each role.

The chi-squared tests and t-tests presented in Table 4 compare treatments in pairs and clearly demonstrate that both frequency and size of punishment differ among treatments. Specifically, greater private benefit induces monitors to punish more, while the effect of negative externality does not seem to be significant.

To investigate further treatment effects, in Figure 3 we report the frequency and size of punishment for citizens (black bars) and public officials (white bars), with observations pooled across the low and high private benefits treatments (T1&T3 *versus* T2&T4; Panel (a)) and the low and high negative externality treatments (T1&T2 *versus* T3&T4;Panel (b)).

Figure (3) Third-party punishment: treatment differences



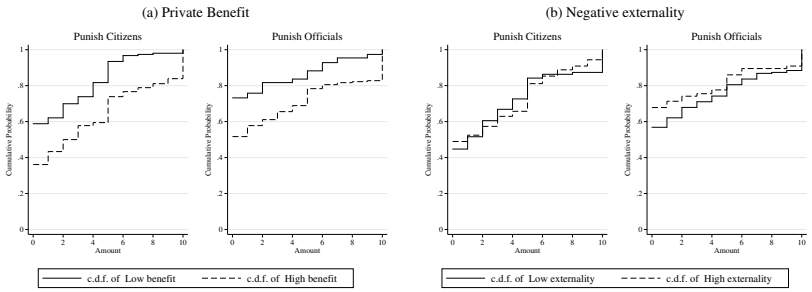
Panel (a) shows that when the private benefits are greater, monitors punish citizens more often (63.89% in the high benefits treatments *versus* 41.18% in the low benefits treatments; chi-squared test p-value = 0.000) and to a greater extent (5.617 tokens in the high benefits treatments *versus* 4.127 in the low benefits treatments; t-test with Welchs formula p-value = 0.000). Similar results hold for public officials: when private benefits are greater, monitors punish them more often (48.33% in the high benefits treatments *versus* 26.80% in the low benefits treatments; chi-squared test p-value = 0.000) and to a greater extent (5.988 tokens in the high benefits treatments *versus* 5.024 in the low benefits treatments; t-test with Welchs formula p-value = 0.095).

Panel (b) shows that when externalities are greater, monitors punish citizens and public officials less often, but only this latter effect is statistically significant (32.17% in the high externality treatments *versus* 43.16% in the low externality treatments; chi-squared test p-value = 0.041). The size of punishment of both citizens and public officials increases with greater externalities, but this effect is not statistically significant.

The treatment effects in monitors' behaviors can be seen clearly in Figure 4, which plots the cumulative distribution functions (CDFs) of the amount of punishment for citizens and public officials, pooling across the private benefit treatments (Panel (a)) and negative externality treatments (Panel (b)).

That the CDFs of the high benefit treatments (dashed lines in Panel (a)) are significantly lower than the cumulative distribution functions of the low benefit treatments (solid lines in Panel (a)) is consistent with the fact that the average frequency and size

Figure (4) Third-party punishment: cumulative distribution functions by treatment



of punishment of both citizens and public officials are greater in the high benefit treatments than in the low benefit treatment. The CDFs in Panel (a) are almost parallel to each other, indicating that the difference in punishment between low and high benefit treatments holds true at all levels of punishment (i.e., both at the extensive and intensive margins). Regarding externality treatments, the 'Punish Citizen' graph in Panel (b) shows that the CDFs of the high externality treatments (dashed line) almost coincide with the CDFs of the low externality treatments (solid line). Instead, the 'Punish Officials' graph in Panel (b) shows that the CDFs of the high externality treatments (dashed line) are slightly higher than the CDFs of the low externality treatments (solid line), but the two CDFs tend to converge at the higher ends of the distribution. These paths are consistent with the fact that greater externalities – while having no significant impact on monitors' behavior towards citizens – lead monitors to punish public officials less often.

the regression results presented in Table 5 corroborate descriptive statistics. Greater private benefits significantly increase both the probability of punishment (by 21.6% – 22% to punish citizens and 20% – 21% to punish public officials) and the size of punishment (by 1.8 tokens – 2.6 tokens to punish citizens and by 1.7 tokens to punish public officials). These results are robust to controls on individual characteristics.

The regression models also confirm that greater negative externalities have no significant effect on either the size of punishment, or the frequency of punishment of citizens and – contrary to the expected – has a small negative effect on the frequency of punishment of public officials (approximately 11%).

These observations lead to our core result.

Result 5 Greater private benefit has a positive, statistically significant effect on both the frequency and the size of punishment towards both citizens and public officials. Greater negative externality has a negative, statistically significant effect on the frequency of punishment towards public officials. Contrary to expectations, greater negative externality does not affect either frequency or size of punishment of citizens.

By interpreting these results through the lens of moral costs as described in Section 2.3, observing greater private benefit leads monitors to experience greater intrinsic

moral costs towards both citizen and public official. By contrast, observing greater negative externalities leads monitors to experience smaller intrinsic moral costs towards public official, and no variations in moral costs towards citizens.

Finally, by exploring some micro-determinants of punishment behavior, we find an interesting effect of gender. Female subjects were approximatively 14% more likely to punish citizens and 9% more likely to punish public officials. Interestingly, conditional upon participation in punishment, female subjects spent significantly less (almost 3 tokens less) than male subjects to punish citizens and public officials (see Tables A2 and A3). Subjects who study economics demonstrate a similar effect: they are more prone to punish but if they do, they punish to a smaller extent. Other control variables do not demonstrate robust effects.

Another relevant question is to understand whether monitors differentiate between citizens and public officials, i.e., whether they tend to punish more one or the other for a corrupt transaction. It should be recalled that we deliberately chose to assign a greater responsibility to public officials: they receive a greater monetary benefit and have the last word. Overall, pooling across treatments, the rates of punishment are very close: 53.5% of monitors decide to punish a citizen who offers a bribe and 54.7% of monitors punish a public official who accepts one (test of proportions p-value=0.577). Conditional upon participation in punishment and pooling across treatments, monitors spent 5.09 tokens to punish citizens and 5.68 tokens to punish public officials; this difference is not statistically significant (t-test with Welchs formula p-value=0.24). There is no significant difference if we compare the frequency and amount of punishment by treatments.

Let us now analyze whether treatment conditions impose any change on the behavior of monitors in terms of punishment differentiation between citizens and officials. To this aim, we consider observations where a bribe was offered and accepted. We then estimate regressions where the dependent variable is the difference of the size of punishment between the citizen and the public official. Table A4 reports the regressions estimates.

Contrary to expectations, we find that both high private benefit and high negative externality lead monitors to punish citizens to a greater extent than they do to public officials. The difference is around a half of an experimental token.

These findings lead to the following result.

Result 6 . Greater externality or greater private benefits lead monitors to punish citizens more than public officials.

This result is in line with *Result 5*, where we find that greater negative externality decreases the frequency of punishment towards public officials, whereas it does not affect the punishment of citizens. By interpreting Result 6 through the lens of moral costs as described in Section 2.3, when observing greater negative externality or greater private benefit, monitors experience greater intrinsic moral costs towards citizens than towards public officials.

Table (5) Frequency and size of punishment: regressions estimates

	Punish citizens		Punish officials	
	(1)	(2)	(3)	(4)
PANEL (A): FREQUENCY OF PUNISHMENT (PROBIT REGRESSIONS)				
'High benefit' treatment	0.220***	0.206***	0.211***	0.201***
	(0.0194)	(0.0240)	(0.0355)	(0.0364)
'High externality' treatment	-0.0413	-0.0177	-0.109***	-0.115***
	(0.0458)	(0.0433)	(0.0421)	(0.0360)
Controls	NO	YES	NO	YES
Observations	333	333	333	333
PANEL (B): SIZE OF PUNISHMENT (TOBIT REGRESSIONS)				
'High benefit' treatment	1.872***	2.644***	1.730**	1.764**
	(0.306)	(0.316)	(0.730)	(1.008)
'High externality' treatment	0.295	-0.172	0.348	-0.294
	(0.313)	(0.388)	(0.955)	(0.954)
Controls	NO	YES	NO	YES
Observations	178	178	128	128

Standard errors in parentheses, clustered by rounds.
For probit regressions marginal effects are reported.
Control variables: Female, Age, Economists, Risk aversion, Religiosity, Catholic.
See Table A2 for the marginal effects of each control variable.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5 Additional Analyses

In this section we explore additional micro-determinants of individuals' tolerance of corruption (Section 5.1), and conduct robustness checks (Section **??**).

5.1 Emotions as determinants of third-party punishment

In the attempt to better understand the motives driving third-party punishment, in this section we investigate whether and to what extent feelings of anger are associated with punishment choices. We further analyze whether *empathic* anger at witnessing bribery acts felt by the monitors (third parties) is a distinct emotion from *direct* anger about personal harm as felt by the other members of society (passive victims, or second parties).

Feelings of anger were measured at the end of each round of the bribery game, where we asked the other members of the society and the monitors the following question: "Imagine to you casually meet the citizen / public official of this round on the street. What would your feelings towards this person be? Please indicate the intensity of your anger and annoyance towards this citizen / public official on a scale of 1 to 10,

where 1 = not at all angry and annoyed, and 10 = very angry and annoyed." For the sake of clarity, the scale of anger ranging from 1 to 10 has been recoded to a scale ranging from -1 to 1.

Overall, we find a high level of heterogeneity of direct and empathic anger. Table 6 presents the descriptive statistics pooled across treatments.

Table (6) Emphatic and direct anger: descriptive statistics

Anger:	Towards citizens:		Towards public officials:	
	offered a bribe	did not offer	accepted a bribe	did not accept
Direct anger :	0.47	-0.92	0.65	-0.86
Empathic anger of monitors:				
All monitors	0.36	-0.97	0.5	-0.94
Who punished	0.49	-	0.66	-
Who did not punish	0.19	-0.97	0.29	-0.94

Notes: self reported emotions are scaled from 1 to +1: -1 indicates the absence of anger; 0 - a neutral judgment; +1 -the absolute anger.

First, we observe a strong and positive level of direct and empathic anger towards citizens who offered a bribe (0.47 and 0.36) and officials who accepted an offered bribe (0.65 and 0.5); on the other hand there is almost no anger at all, either direct or empathic, towards citizens who did not offer a bribe (-0.92 and -0.97) or towards officials who did not accept one if offered (-0.86 and -0.94). *Second*, in line with our expectations, direct anger is higher than empathic anger towards both citizens ($0.47 > 0.36$, t-test p-value =0.034) and public officials ($0.65 > 0.5$, t-test p-value =0.051). *Third*, and most interestingly, monitors who punished experienced higher levels of anger than monitors who did not not (0.49 *versus* 0.19 anger towards citizens and 0.66 *versus* 0.29 anger *versus* public officials). This is consistent with the view that emotions drive sanctioning decisions. Table 6 shows that negative emotions induce third parties to punish. This is exactly what we observe in . *Finally*, we find that both second and third parties experience greater levels of anger towards public officials than towards citizens ($0.65 > 0.47$, t-test p-value =0.000; $0.5 > 0.36$, t-test p-value =0.000).

To further investigate whether levels of direct and empathic anger varies between treatment conditions, we perform regression analysis. Results are presented in *Table 7*, where the dependent variables are the self-reported level of direct anger (Columns 2 and 3) and of empathic anger (Columns 4 and 5).

The results presented in Table 7 clearly demonstrate that corrupt behavior significantly increases the level of both direct and empathic anger towards both citizens and public officials. We observe that greater private benefit significantly decreases the level of direct anger and does not have a significant impact on the level of empathic anger.

Table (7) Self reported anger of second and third parties: regression results

	Direct anger towards:		Empathic anger towards:	
	citizen	official	citizen	official
Bribe is offered	1.383***		1.344***	
	(0.0469)		(0.0447)	
Bribe is offered * High private benefit	-0.147*		0.0776	
	(0.0686)		(0.0499)	
Bribe is offered * High negative externality	0.184***		-0.136**	
	(0.0377)		(0.0506)	
Bribe is accepted		1.542***		1.531***
		(0.0522)		(0.0708)
Bribe is accepted * High private benefit		-0.190**		-0.0103
		(0.0667)		(0.0686)
Bribe is accepted * High negative externality		0.191***		-0.206**
		(0.0358)		(0.0646)
Observations	640	333	640	333

Standard errors in parentheses, clustered by rounds

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In contrast, greater negative externality significantly increases the level of direct anger, whereas it significantly decreases the level of empathic anger. This result, along with others, will be discussed in section 6. In line with *Result 6*, Table 7 indicates that in the presence of greater negative externality, third parties feel greater anger towards citizens than public officials (empathic anger decreases from 0.206 to 0.136).

The main results of this section can be summarized as follow.

Result 7 . Corrupt behavior significantly increases the levels of both direct and empathic anger towards both citizens and public officials. In line with Result 5, greater negative externality leads monitors to feel lower levels of anger towards both citizens and officials. In line with Result 6, greater negative externality leads monitors to feel higher levels of anger towards citizens than towards public officials.

5.2 Robustness checks

To exclude bias due to start-game and end-game effects (e.g., Gonzalez et al., 2005), we dropped the observations of the first and last round of the bribery game and estimated again the main regressions presented in Tables 5. The estimates presented in Tables A5 and A6 confirm our results, both in statistical significance and in magnitude.

6 Discussion and conclusion

Recent research has revealed that corruption is not only a function of economic and political factors, such as GDP per capita, income distribution, democracy and political instability, but also a function of social and cultural norms that shape individuals' propensity to engage in and punish corruption (e.g., Alatas et al., 2009; Cameron et al., 2009; Banuri and Eckel, 2012) Studying corruption through the paradigm of social norms is a relatively new line of inquiry in the experimental literature (Abbink 2016, Julien Senn 2019). We significantly contribute to the development of this line by applying a third-party punishment game which is considered to be a "powerful tool for studying the characteristics and the content of social norms" (Fehr and Gächter, 2002).

We recruited 256 individuals living in the north of Italy to play a bribery game with a third-party punishment component. Our findings indicate that more than half of the subjects offer and accept bribes, and that more than a half of monitors, whose economic payoff is unaffected by the behavior of bribers and the bribed, are willing to give away on average 25% of their endowment to impose sanctions on either the briber or the bribed or both.

In our design we varied two basic properties of corruption: private benefits and negative externalities. These treatment conditions made it possible to obtain deeper insights into monitors' behavior and to understand the source of the intrinsic moral costs that monitors experience when they observe corruption.

We obtain two novel sets of results. Our main set of results is that, first, greater private benefit leads monitors to increase both the frequency and the size of punishment towards both citizens and public officials. Secondly, greater negative externality does

not have a significant effect on the punishment towards citizens, whereas it does have a negative effect on the punishment towards public officials. We provide an interpretation of our results though the lens of inequality aversion: monitors punish corruption because they care about their rank on the scale of welfare distribution. That is, they increase (decrease) their punishment if their comparative advantage with respect to other players is higher (lower).

On the one hand, holding negative externality constant, greater private benefits make monitors worse off with respect to citizens and officials. This leads monitors to punish bribers and the bribed to a greater extent. On the other hand, holding private benefits constant, greater negative externality makes monitors worse off with respect to citizens and officials, but better off with respect to the other members of society. This leads them to punish bribers and the bribed to a lesser extent. This reasoning is confirmed by emotions analysis: monitors feel significantly lower levels of anger towards both citizens and public officials in the high negative externality treatment.

Our second set of results concerns the comparison of punishment of citizens vis-à-vis public officials. Pooling across treatments, the difference in the frequency of punishment towards citizens and towards public officials is respectively 53.5% and 54.7%. Conditionally on participation in punishment, the difference in the size of punishment towards citizens and towards public officials is respectively 5.09 tokens and 5.68 tokens. Despite this difference not being statistically significant, we find that monitors experience a significantly higher level of anger towards public officials as opposed to citizens, and this result holds for each treatments. Moreover, interesting results emerge from the analysis of treatment conditions: greater externality or greater private benefit leads monitors to punish citizens more than public officials, as well as to feel higher levels of anger towards citizens than towards public officials. It should be recalled that in the low private benefit treatment, the advantage obtained by citizens for bribing is 6 tokens while the advantage of officials is 9 tokens, i.e., the advantage of officials is 50% higher; in high private benefit treatment, it is 15 and 21 tokens correspondingly, i.e., the advantage of officials with respect to citizens is only 40% higher. The fact that in the low private benefit condition monitors punish officials, who have a greater advantage vis-à-vis citizens, relatively more than in the high private benefit condition, is consistent with our interpretation of the results, that inequality aversion mainly drives the behavior of monitors.

Inequality aversion may also explain the reason why greater negative externality decreases the punishment – and the level of anger – towards officials more than punishment with respect to citizens. Holding private benefit low, in the low externality treatment citizens gain 6 tokens, monitors gain 12 tokens, each other members of the society lose 3 tokens. The difference in gains between citizens and other members of the society is 9 tokens, while the difference in gains between officials and other members of the society is 12 tokens. The comparative advantage of officials vis-à-vis other members is 33% higher than the advantage of citizens. In the high externality treatment, where each other member of the society loses 6 tokens for each bribe, the comparative advantage of officials is 25% higher than the advantage of citizens. This might explain the reason why a greater negative externality leads monitors to punish officials less than citizens – and to feel less anger towards the former with respect to the latter.

To conclude, our research adds the important contribution to the literature, demonstrating that one's own relative standing on the scale of welfare distribution is the main driving force for the enforcement of no-corruption norm, and that people are ready to punish for corruption if they understand that the violation of the no-corruption norm puts them in a less favorable position on the scale of welfare distribution and to increase their punishment if their comparative advantage is scaled down. The willingness of people to punish those who reduce their personal welfare is a well-known fact in the literature; we find that this tendency holds also in relative terms. Our results also indicate that monitors do not increase their punishment if social loss expands.

Our results open up various new avenues for research. The most intuitive extension would be to fix the parameters of the game so that the welfare of monitors is always higher than the welfare of all other players in all the treatments. This extension would make it possible to separate willingness to punish for the inequality observed from the decrease in the observer's own comparative advantage. On the other hand, this extension changes the position of monitors in society: in our design monitors are representatives of the middle class, since their welfare is always lower than the welfare of bribers and the bribed, but higher than the welfare of other members of the society. It would be interesting to analyze the behavior of members of the upper class, which is endowed with higher bargaining power in many societies.

It would be also interesting to study punishment *versus* assistance. According to pure altruism theory (Andreoni, 1990), altruistic people will never use costly options to reduce other subjects' payoffs. In our design we apply a Third-party punishment paradigm which allows monitors to punish bribers and the bribed at their own cost. An alternative treatment would be to allow them to redirect a part of their endowment to assist the victims of corrupt acts. Is a society more willing to punish norm violators, or rather to redirect resources to those who have suffered norm violations? For example, while in Protestant communities it is a norm to report those who violate norms, in Catholic communities assistance and forgiveness are prescribed.

Another fascinating question would be to study the effect of information. In our design, monitors do not know how many bribes were offered and accepted in the room; they only know whether a bribe was offered and accepted in their own group. Analyzing how the level of public disclosure of information about corruption differs affects tolerance to corruption may yield important policy implications.

Finally,carrying out similar research in different countries has potential to provide additional insights and to open further avenues for explanation of corruptive behavior across cultures.

References

Abbink, Klaus and Danila Serra, "Chapter 4 anticorruption policies: Lessons from the lab," in "New advances in experimental research on corruption," Emerald Group Publishing Limited, 2012, pp. 77–115.

_ **and Heike Hennig-Schmidt**, "Neutral versus loaded instructions in a bribery experiment," *Experimental Economics*, 2006, *9* (2), 103–121.

_ , **Bernd Irlenbusch, and Elke Renner**, "An Experimental Bribery Game," *Journal of Law, Economics, and Organization*, 2002, *18* (2), 428–454.

Armantier, O. and A. Boly, "A Controlled Field Experiment on Corruption," *European Economic Review*, 2011, (55(8)), 1072–1082.

Azfar, Omar and William Robert Nelson, "Transparency, wages, and the separation of powers: An experimental analysis of corruption," *Public Choice*, 2007, *130* (3-4), 471–493.

Balafoutas, Loukas, Kristoffel Grechenig, and Nikos Nikiforakis, "Third-party punishment and counter-punishment in one-shot interactions," *Economics letters*, 2014, *122* (2), 308–310.

Barr, Abigail and Danila Serra, "The Effects of Externalities and Framing on Bribery in a Petty Corruption Experiment," *Experimental Economics*, 2009, *12* (4), 488–503.

_ , **Magnus Lindelow, and Pieter Serneels**, "Corruption in public service delivery: An experimental analysis," *Journal of Economic Behavior & Organization*, 2009, *72* (1), 225–239.

Batson, C Daniel, Christopher L Kennedy, Lesley-Anne Nord, EL Stocks, D'Yani A Fleming, Christian M Marzette, David A Lishner, Robin E Hayes, Leah M Kolchinsky, and Tricia Zerger, "Anger at unfairness: Is it moral outrage?," *European Journal of Social Psychology*, 2007, *37* (6), 1272–1285.

Ben-Shakhar, Gershon, Gary Bornstein, Astrid Hopfensitz, and Frans van Winden, "Reciprocity and emotions in bargaining using physiological and self-report measures," *Journal of economic psychology*, 2007, *28* (3), 314–323.

Bendor, Jonathan and Piotr Swistak, "The evolution of norms," *American Journal of Sociology*, 2001, *106* (6), 1493–1545.

Berninghaus, Siegfried K, Sven Haller, Tyll Krüger, Thomas Neumann, Stephan Schosser, and Bodo Vogt, "Risk attitude, beliefs, and information in a corruption game–An experimental analysis," *Journal of economic psychology*, 2013, *34*, 46–60.

Bosman, Ronald and Frans Van Winden, "Emotional hazard in a power-to-take experiment," *The Economic Journal*, 2002, *112* (476), 147–169.

Büchner, Susanne, Andreas Freytag, Luis G González, and Werner Güth, "Bribery and public procurement: an experimental study," *Public Choice*, 2008, *137* (1-2), 103–117.

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

Charness, Gary, Ramón Cobo-Reyes, and Natalia Jiménez, "An investment game with third-party intervention," *Journal of Economic Behavior & Organization*, 2008, *68* (1), 18–28.

Chen, Daniel L, Martin Schonger, and Chris Wickens, "oTreeAn open-source platform for laboratory, online, and field experiments," *Journal of Behavioral and Experimental Finance*, 2016, *9*, 88–97.

Chen, Yefeng, Shuguang Jiang, and Marie Claire Villeval, "The Tragedy of Corruption," IZA Discussion Papers 10175 2016.

Duffy, John and Jack Ochs, "Cooperative behavior and the frequency of social interaction," *Games and Economic Behavior*, 2009, *66* (2), 785–812.

Elster, Jon, *The cement of society: A survey of social order*, Cambridge University Press, 1989.

Fehr, Ernst and Simon Gächter, "Altruistic punishment in humans," *Nature*, 2002, *415* (6868), 137.

_ **and Urs Fischbacher**, "Third-Party Punishment and Social Norms," *Evolution and Human Behavior*, 2004, *25(2)*, 63–88.

Frank, Björn and Günther G Schulze, "Does economics make citizens corrupt?," *Journal of economic behavior & organization*, 2000, *43* (1), 101–113.

Frank, Robert H, *Passions within reason: The strategic role of the emotions*, WW Norton & Co, 1988.

_ **, Thomas D Gilovich, and Dennis T Regan**, "Do economists make bad citizens?," *Journal of Economic Perspectives*, 1996, *10* (1), 187–192.

Fréchette, Guillaume R, "Session-effects in the laboratory," *Experimental Economics*, 2012, *15* (3), 485–498.

Gneezy, Uri, Silvia Saccardo, and Roel van Veldhuizen, "Bribery: Behavioral drivers of distorted decisions," *Journal of the European Economic Association*, 2018.

Gonzalez, Luis G, Werner Güth, and M Vittoria Levati, "When does the game end? Public goods experiments with non-definite and non-commonly known time horizons," *Economics Letters*, 2005, *88* (2), 221–226.

Greiner, Ben, "Subject pool recruitment procedures: organizing experiments with ORSEE," *Journal of the Economic Science Association*, 2015, *1* (1), 114–125.

Kandori, Michihiro, "Repeated games played by overlapping generations of players," *The Review of Economic Studies*, 1992, *59* (1), 81–92.

Kubbe, Ina and Annika Engelbert, *Corruption and Norms: Why Informal Rules Matter*, Springer, 2017.

Monte, Alfredo Del and Erasmo Papagni, "The determinants of corruption in Italy: Regional panel data analysis," *European Journal of Political Economy*, 2007, *23* (2), 379–396.

Nelissen, Rob MA and Marcel Zeelenberg, "Moral emotions as determinants of third-party punishment: Anger, guilt and the functions of altruistic sanctions," *Judgment and Decision making*, 2009, *4* (7), 543.

Nikiforakis, Nikos and Helen Mitchell, "Mixing the carrots with the sticks: Third party punishment and reward," *Experimental Economics*, 2014, *17* (1), 1–23.

Ohtsubo, Yohsuke, Fumiko Masuda, Esuka Watanabe, and Ayumi Masuchi, "Dishonesty invites costly third-party punishment," *Evolution and Human Behavior*, 2010, *31* (4), 259–264.

Pillutla, Madan M and J Keith Murnighan, "Unfairness, anger, and spite: Emotional rejections of ultimatum offers," *Organizational behavior and human decision processes*, 1996, *68* (3), 208–224.

Rivas, M Fernanda, "An experiment on corruption and gender," *Bulletin of Economic Research*, 2013, *65* (1), 10–42.

Sanfey, Alan G, James K Rilling, Jessica A Aronson, Leigh E Nystrom, and Jonathan D Cohen, "The neural basis of economic decision-making in the ultimatum game," *Science*, 2003, *300* (5626), 1755–1758.

Schulze, G. and B. Franck, "Deterrence versus Intrinsic Motivation: Experimental Evidence on the Determinants of Corruptility.," *Economics of Governance*, 2003, (4(2)), 143–160.

Serra, D., "Combining Top-down and Bottom-up Accountability: Evidence from a Bribery Experiment.," *Journal of Law, Economics, and Organization*, 2011, (28(3)), 569–587.

Treisman, Daniel, "The causes of corruption: a cross-national study," *Journal of public economics*, 2000, *76* (3), 399–457.

Appendix A Additional Tables

Table (A1) Bribing behavior: Probit regressions (with all coefficients)

	Bribe offers		Bribe acceptance	
	(1)	(2)	(3)	(4)
'High benefit' treatment	0.0841**	0.0604**	0.125***	0.113***
	(0.0346)	(0.0296)	(0.0316)	(0.0287)
'High externality' treatment	-0.145***	-0.174***	-0.0313	-0.0883**
	(0.0312)	(0.0322)	(0.0376)	(0.0425)
Female		0.000423		0.0553
		(0.0415)		(0.0562)
Age		0.0309		0.0668
		(0.0195)		(0.0695)
Age squared		-0.000784***		-0.00125
		(0.000276)		(0.00136)
Economists		0.124**		0.261**
		(0.0713)		(0.108)
Risk aversion		-0.00465		0.00501
		(0.00631)		(0.00883)
Religiosity		-0.0254**		0.0474***
		(0.0123)		(0.0152)
Catholic		0.0462		0.0510
		(0.0377)		(0.0527)
Observations	640	640	333	333

Standard errors in parentheses, clustered by rounds.

Marginal effects are reported.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table (A2) Frequency of punishment:probit regressions (with all coefficients)

	Punish citizens		Punish officials	
	(1)	(2)	(3)	(4)
'High benefit' treatment	0.220***	0.206***	0.211***	0.201***
	(0.0194)	(0.0240)	(0.0355)	(0.0364)
'High externality' treatment	-0.0413	-0.0177	-0.109***	-0.115***
	(0.0458)	(0.0433)	(0.0421)	(0.0360)
Female		0.143***		0.0866**
		(0.0178)		(0.0349)
Age		0.130***		0.102***
		(0.0312)		(0.0282)
Age squared		-0.00229***		-0.00198***
		(0.000529)		(0.000454)
Economists		0.153**		0.201***
		(0.0718)		(0.0647)
Risk		-0.0456***		-0.0418***
		(0.0116)		(0.0152)
Religiosity		0.0509***		0.0341*
		(0.0165)		(0.0177)
Catholic		0.126		0.0405
		(0.104)		(0.0590)
Observations	333	333	333	333

Standard errors in parentheses, clustered by rounds

Marginal effects are reported.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table (A3) Size of punishment: tobit regressions (with all coefficients)

	Punish citizen		Punish official	
	(1)	(2)	(3)	(4)
'High benefit' treatment	1.872***	2.644***	1.730**	1.764**
	(0.306)	(0.316)	(0.730)	(1.008)
'High externality' treatment	0.295	-0.172	0.348	-0.294
	(0.313)	(0.388)	(0.955)	(0.954)
Female		-2.972***		-2.365***
		(0.672)		(0.689)
Age		-3.918***		0.386
		(1.242)		(1.513)
Age squared		0.0763***		-0.0133
		(0.0255)		(0.0291)
Economists		-1.662**		-2.532**
		(0.733)		(0.979)
Risk		0.152		-0.926***
		(0.202)		(0.345)
Religiosity		-0.182		0.677**
		(0.229)		(0.286)
Catholic		-0.325		-0.0430
		(0.730)		(0.850)
Constant	3.910***	55.28***	4.882***	6.407
	(0.227)	(14.66)	(0.435)	(20.93)
sigma				
Constant	3.830***	3.443***	4.952***	4.401***
	(0.249)	(0.286)	(0.531)	(0.494)
Observations	178	178	128	128

Standard errors in parentheses, clustered by rounds

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table (A4) Size of punishment: citizens *versus* public officials

	(1)	(2)
'High benefit' treatment	0.514*	0.477*
	(0.260)	(0.223)
'High externality' treatment	0.571***	0.603**
	(0.155)	(0.195)
Female		-0.490*
		(0.265)
Age		-0.397
		(0.236)
Age squared		0.00790**
		(0.00345)
Economists		0.236
		(0.468)
Risk		0.341***
		(0.103)
Religiosity		-0.264*
		(0.125)
Catholic		0.469
		(0.317)
Constant	-0.655**	3.903
	(0.217)	(3.510)
Observations	234	234

The sample is restricted to cases when the bribe was sent and accepted.
DV: difference in size of punishment between citizens and public officials.
The dependent variable = punishment of a citizen - punishment of an official.
Standard errors in parentheses, clustered by rounds
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table (A5) Robustness check on frequency of punishment: probit regressions

	Punish citizen		Punish official	
	(1)	(2)	(3)	(4)
'High benefit' treatment	0.217***	0.187***	0.204***	0.189***
	(0.0239)	(0.0289)	(0.0394)	(0.0378)
'High externality' treatment	-0.0575	-0.0484	-0.131***	-0.155***
	(0.0561)	(0.0497)	(0.0431)	(0.0306)
Female		0.140***		0.0716*
		(0.0239)		(0.0403)
Age		0.131***		0.108***
		(0.0369)		(0.0376)
Age squared		-0.00236***		-0.00214***
		(0.000609)		(0.000649)
Economists		0.165**		0.217***
		(0.0775)		(0.0715)
Risk		-0.0440***		-0.0489***
		(0.0147)		(0.0181)
Religiosity		0.0643***		0.0476**
		(0.0167)		(0.0202)
Catholic		0.218**		0.0943**
		(0.0980)		(0.0466)
Observations	264	264	264	264

Standard errors in parentheses, clustered by rounds

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table (A6) Robustness check on size of punishment: tobit regressions

	Punish citizen		Punish official	
	(1)	(2)	(3)	(4)
'High benefit' treatment	2.071***	2.911***	2.451***	2.640**
	(0.299)	(0.319)	(0.580)	(1.046)
'High externality' treatment	0.485	0.199	1.178	0.706
	(0.385)	(0.415)	(1.165)	(1.048)
Female		-3.173***		-3.303***
		(0.850)		(0.723)
Age		-4.039**		1.031
		(1.570)		(1.687)
Age squared		0.0804**		-0.0233
		(0.0323)		(0.0333)
Economists		-1.345*		-2.246**
		(0.768)		(1.005)
Risk		0.281		-0.643**
		(0.215)		(0.285)
Religiosity		-0.227		0.617
		(0.261)		(0.382)
Catholic		-0.268		0.186
		(0.946)		(1.017)
Constant	3.865***	55.28***	4.663***	-4.453
	(0.263)	(18.57)	(0.540)	(22.83)
sigma				
Constant	3.831***	3.404***	5.015***	4.414***
	(0.295)	(0.354)	(0.664)	(0.621)
Observations	143	143	100	100

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix B The written instructions to participants

These are instructions for treatment 1, with low social loss and low private benefit.[15] Original text in Italian and available upon request to the authors.

Welcome!

Welcome! and thanks for participating. This is an experiment to evaluate how people make economic decisions .

If you pay attention, the instructions will help you make these decisions and earn a sum of money. This sum depends both on yours decisions during the experiment , and on the decisions of the other participants.

The earnings will be calculated in tokens, converted to Euro at the end of the experiment and paid in cash at the end of today's session. For every 5 tokens, you will receive 1 Euro. In addition, you will receive 5 Euros for your participation .

Your answers will remain confidential and final payments will be made in a sealed envelope. You are not obliged to communicate your earnings to anyone.

It is important not to communicate in any way with other people in the room until the end of the experiment . The use of cell phones is prohibited. Please take a moment to turn off your cell phones and put away all outside materials.

You can ask questions at any time. If you have a question, please raise your hand and we will respond in private .

Anonymity

When you entered the room , you took a piece of paper with a number. This and yours identification code and will be used to guarantee your anonymous. The decisions you make during the experiment will be matched to your identification code , never to your name.

No one (ie neither the people who will analyze the data, or other participants, nor any other person) will never be able to match your name to the decisions that you will take during the experiment.

Ten Rounds

This experiment consists of 10 rounds . Each round is completely independent of the others: the decisions taken in a round will not influence either your gain or the procedures in the following rounds.

At the end of the experiment, the computer will randomly choose a round for your payment.

Roles and Groups

At the beginning of today's experiment, you will be randomly assigned to one of four roles: a **Citizen**, a **Public Official**, **Other Member**, or a **Monitor**.

At the beginning of each round, the computer will randomly form groups of four participants, each with a different role. The composition of the groups will casually change from round to round.

[15] The instructions for other treatments and instructions in Italian language are available upon request.

You will never know the identity of the other members of the group, nor will the other members of the group know yours.

The Decisions

At the beginning of each round every player receives **50 tokens**.

Each round consists of four stages.

1. First stage: the Citizen's decision. If you are a Citizen, you have to decide whether to offer the Official a bribe or not.

If you do not offer a bribe to the public official , your payoff remains equal to 50 tokens and the round ends.

If you offer a bribe to the public official, you offer 3 tokens to the public official. These 3 tokens will only be returned to you if the public official rejects the bribe. Please see the second stage.

2. Second stage: the Official's decision. If you are an Official, you will be notified of the decision made by the citizen in your group.

If the Citizen in your group has offered you a bribe, you can decide whether to accept it or reject it. Accepting the bribe implies accepting the 3 tokens offered by the citizen and granting a favour to the citizen, which causes a loss to the other members of society. Specifically:

If you accept the bribe:

- You receive 9 additional tokens (in addition to the 3 tokens offered by the Citizen);
- The Citizen receives 9 additional tokens;
- Each Other Member in this room loses 3 tokens, and may lose more tokens if bribes are offered and accepted in the other groups in this room;
- You and the Citizen may receive a punishment by the Monitor (see Fourth Stage).

If you reject the bribe:

- You receive 0 additional tokens (and return the 3 tokens to the citizen);
- The Citizen receives 0 additional tokens;
- Each Other Member may lose some tokens if bribes have been offered and accepted in the other groups in this room;
- You and the Citizen may receive a punishment by the Monitor (see Fourth Stage).

3. Third stage: the role of the Other Members. If you are an other Member of the society, your role is passive: you cannot take any decision.

Per each bribe offered and accepted in this room, you lose 3 tokens.

While waiting for the other participants' decisions, we will ask you to answer some questions, with no effects on your earnings.

4. Fourth stage: the Monitor's decision. If you are a Monitor, the experimenter will show you the decisions made by the Citizen and the Official in your group.

You can choose an amount between 0 and 10 tokens to punish the Citizen, and an amount between 0 and 10 tokens to punish the Official. Your payoff will be reduced by the amount of punishment you have chosen. The amount of punishment that you choose will be multiplied by 2, and the payoffs of the Official and the Citizen will be reduced each by this doubled amount.

For example, if you spend 5 tokens to punish the citizen and 2 tokens to punish the public official: Your earnings are reduced by 7 tokens ($= 5 + 2$); the citizen's earnings are reduced by 10 tokens ($= 5 \times 2$); the earnings of the public official are reduced by 4 coins ($= 4 \times 2$).

All your decisions remain private information: the other participants will not receive any information about your choices.

After the fourth stage, the round ends and the next round starts.

Remember that each round is completely independent from the others: the choices made in one round do not affect neither payoffs nor procedures in the next rounds.

At the end of the experiment, one round will be randomly chosen for payment.

Questionnaire and Payment

After receiving information on your final earnings, please remain sit: we will ask you to answer a series of questions.