

Buffat, Justin; Senn, Julien

**Working Paper**

## Corruption and cooperation

Working Paper, No. 260

**Provided in Cooperation with:**

Department of Economics, University of Zurich

*Suggested Citation:* Buffat, Justin; Senn, Julien (2017) : Corruption and cooperation, Working Paper, No. 260, University of Zurich, Department of Economics, Zurich, <https://doi.org/10.5167/uzh-138912>

This Version is available at:

<https://hdl.handle.net/10419/173419>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



**University of  
Zurich** <sup>UZH</sup>

University of Zurich  
Department of Economics

Working Paper Series

ISSN 1664-7041 (print)  
ISSN 1664-705X (online)

---

Working Paper No. 260

# **Corruption and Cooperation**

Justin Buffat and Julien Senn

August 2017

---

# CORRUPTION AND COOPERATION\*

Justin Buffat      Julien Senn

August 03, 2017

## Abstract

Corruption is a widespread phenomenon. Nevertheless, causal evidence on the effects of corruption is still lacking. In this paper, we assess whether and how corruption affects cooperation using a public good game experiment. Overall, contributions to the public good are reduced by 30% when participants have the possibility to bribe the punishment authority. Two concurrent channels lead to lower levels of cooperation. First, the punishment of low contributors decreases both at the intensive and the extensive margin. Second, bribery discourages initially high contributors, who gradually decrease their contributions down to the level of initially low contributors.

**JEL CODES:** C91, D73, K42

**Key words:** corruption, bribery, cooperation, public good, institutions

---

\*Financial support from the Center for Economic and Social Behavior (C-SEB) from the University of Cologne is gratefully acknowledged. Buffat: University of Lausanne and C-SEB (justin.buffat@unil.ch). Senn: University of Zurich (julien.senn@econ.uzh.ch).

# 1 Introduction

Corruption is a widespread phenomenon. In 2016, it was estimated that more than 1 trillion of USD was spent on bribery worldwide (World Bank Group).<sup>1</sup> While some have argued that corruption might increase efficiency (see e.g. Huntington, 1968), the lion's share of the empirical evidence suggests that it imposes large costs on the society. Existing studies document, among other things, that corruption negatively affects investment and growth (Mauro, 1995; Méon and Sekkat, 2005) and that it creates important allocation distortions (see e.g. Bertrand et al., 2007).<sup>2</sup> The International Monetary Fund recently estimated the cost of bribery at about 1.5 to 2 trillions of USD, i.e. approximately 2 percent of the global GDP (IMF). Western countries are not spared by this plague: a 2014 anti-corruption report from the European Union evaluated the costs of corruption within the EU at 120 billions of euros per year.<sup>3</sup> While suggestive, none of the existing evidence can unambiguously establish the causal (negative) effects of corruption. Using a public good game experiment in which the punishment authority is centralized, our paper takes a first step in this direction by providing direct evidence that corruption has a negative effect on cooperation and by documenting two channels through which this effect operates.

In our experiment, participants are matched in fixed groups of four: three *citizens* and one *monitor* (she). We investigate the effects of corruption on cooperation using two treatments in a between-subjects design. In the NO BRIBE treatment, citizens decide on their contribution to the public good and the monitor subsequently allocates punishment points. In the BRIBE treatment, citizens have the possibility to send a small bribe to the monitor before she allocates punishment points. We purposefully restrict our attention to small bribes, i.e. transfers of money that do not make up for a large portion of the revenue of the monitor, because of the variety of corruption situations that they encompass: a traffic offender slips a bank note in his driver's license, a real estate agent sends a box of wine to the person in charge of deciding on which piece of land the construction of a new building will be allowed, etc. In all these examples, the briber sends a small gift to someone in the hope of getting a favor in return.

We conjecture that cooperation will be lower in BRIBE than in NO BRIBE

---

<sup>1</sup>Retrieved from <http://www.worldbank.org/en/topic/governance/brief/anti-corruption> on February 22, 2017.

<sup>2</sup>Mauro (1995) uses an instrumental-variable approach to show that corruption negatively affects investment, which in turn negatively impacts growth. Using a non-instrumental approach, Méon and Sekkat (2005) find that corruption has a negative impact on growth independently from its impact on investment. Finally, Bertrand et al. (2007) study the allocation of driver's license in India using a field experiment. They show that subjects that are (randomly) more likely to make extralegal payments are also more likely to obtain a license without knowing how to drive.

<sup>3</sup>See the 2014 EU Anticorruption Report (EC 2014).

for two reasons. First, we expect citizens – in particular the free-riders – to bribe the monitor in the hope of receiving fewer deduction points just as a traffic offender might bribe a policeman in the hope of not receiving a fine. The monitors are then expected to reciprocate to these bribes by inflicting *ceteris paribus* fewer punishment points. Because the punishment is softer, the contributions of low contributors will fail to increase in the subsequent periods. This chain of event is expected to gradually undermine the overall level of cooperation in BRIBE.

Second, we expect a discouragement effect to strike in in the BRIBE treatment. More specifically, we conjecture that citizens that initially contribute a lot to the public good will get discouraged by seeing that the initially low contributors of their group manage to avoid heavy punishment by bribing the monitor. We expect that this discouragement will lead them to decrease their contributions over time.

Our experimental approach has two main advantages. First, the exogenous manipulation of the institutional environment allows us to cleanly identify the effects of corruption on cooperation. Such an exogenous variation of corruption would be virtually impossible to either observe or achieve in a more natural setting. While instrumental variables can be used to circumvent endogeneity problems and reverse causation, it is not clear that an indisputably good instrument for corruption, i.e. a variable that is correlated with corruption but not with cooperation, exists. In such circumstances, laboratory experiments provide a unique way to gather causal evidence on the effects of corruption on cooperation. Second, and perhaps most importantly, our experimental design allows to nail down different channels through which bribery affects cooperation. This is important because only a better understanding of how bribery affects cooperation can help the society to come up with more effective ways to tackle it.

Several important findings emerge from our data. We first show that overall contributions to the public good are reduced by 33% when participants have the possibility to bribe the monitor. Interestingly, the average initial contribution to the public good in BRIBE is not significantly different than in NO BRIBE. This finding suggests that institutional differences take time before significantly affecting behavior. This is indeed the case: it is only after the fourth period that the average contribution in NO BRIBE significantly exceeds the average contribution in BRIBE. Interestingly, the prevalence of corruption does not lead to a complete collapse of cooperation. Indeed, citizens contribute on average 50% of their endowment to the public good in the BRIBE condition.<sup>4</sup>

While documenting such a large treatment effect is interesting in itself, our paper

---

<sup>4</sup>Using a  $n$ -person Prisoner's Dilemma framework in which players can bribe the punishment authority, Kosfeld (1997) provides theoretical support for this finding. He shows that there exists a maximal number of corrupting and defecting agents under which the public good is formed and the other members of the group cooperate.

also investigates how it originates. We provide evidence that at least two concurrent channels can explain this finding.

First, we show that bribery largely softens the punishment behavior of the monitor. On average, BRIBE reduces punishment both at the intensive and at the extensive margin (despite the fact that the average level of cooperation is lower in BRIBE). This is particularly true for low contributors. For example, a citizen that contributes between 12 and 20 points less than the other citizens of his group (the maximum contribution is 20 points) receives about 5 times more punishment points in NO BRIBE than in BRIBE (7.1 points versus 1.5 point). This difference in the severity of punishment has implications for cooperation: while harsh sanctions discipline low contributors in NO BRIBE, the soft punishments inflicted in BRIBE are not sufficient to enforce higher levels of cooperation in the subsequent periods.

Second, we show that the prevalence of bribery has a large discouraging effect on initially high contributors, i.e. citizens that contributed more than the average of the other members of their group in period 1. Quite astonishingly, the average contribution in period 1 of initially high contributors is not affected by the treatment. Neither is the average contribution in period 1 of initially low contributors. However, the monitors assign significantly more punishment points to initially low contributors in period 1 in NO BRIBE than in BRIBE. Over time, initially low contributors increase their contributions up to the level of initially high contributors in NO BRIBE. In contrast, initially high contributors decrease their contribution down to the level of initially low contributors in BRIBE. Ultimately, all the citizens end up contributing a relatively high proportion of their endowment (approximately 75%) in NO BRIBE, whereas all the citizen in BRIBE end up contributing a relatively low proportion of their endowment (less than 50%). We speculate that the main reason for which initially high contributors reduce their contributions so dramatically in BRIBE is that they get discouraged to see that initially low contributors can “*get away with murder*”, i.e. that they can free-ride, bribe and nevertheless receive few punishment points.

Our paper relates to the growing literature studying the causes and consequences of corruption using tailor-made laboratory experiments. Pioneering the experimental work on bribery, Abbink et al. (2002) show that reciprocity between the briber and a public official can establish bribery relationships. They also show that negative externalities imposed by the briber on the society do not affect bribery and that tougher sanctions imposed on bribers if bribery is discovered reduce corruption at the extensive margin. Their paper has paved the way for numerous experiments that have shed light on the effects of staff rotation (Abbink, 2004), wages of public officials (Abbink, 2005; Van Veldhuizen, 2013) and culture (Barr and Serra, 2010; Cameron et al., 2009) on corruption,

among others.<sup>5</sup>

Four recent papers are linked to our study. In a concurrent study that has been conducted independently, Muthukrishna et al. (2017) investigate the effects of structural factors (e.g. the monitor’s punitive power or the economic potential of the public good) as well as anti-corruption strategies (e.g. increased transparency) on cooperation. They provide evidence that corruption reduces cooperation and that anti-corruption strategies are effective under some conditions but can backfire in others (e.g. when the monitor is endowed with little power to discipline free-riders). While their experimental design shares some similarities with ours, important differences between the two designs exist.<sup>6</sup> Moreover, they focus the effects of structural factors and anti-corruption policies on cooperation but do not investigate the channels through which corruption affects cooperation, which is at the heart of the analysis we conduct.

Cagala et al. (2016) study the effects of corruption on the provision of public goods using dictator games. In their experiment, participants are matched in groups of four: three contributors and one administrator. In the first stage, the administrator decides whether she is willing to expropriate 10 percent of the donations that will be made by the contributors in the second stage. In the second stage, the contributors decide how much they are willing to donate to a charity of their choice (which they picked in stage 1). Despite large differences with our experimental design, their results are comparable to ours: Compared to a control condition in which the administrator cannot expropriate the contributors, subjects matched with an expropriating administrator reduce their donations by about 30% (recall that we document a 33% decrease in cooperation).

Malmendier and Schmidt (2017) study “*the dark side of prosocial behavior*” by focusing on the negative externalities of gift giving. In their main treatment, a decision maker buys a product from two potential sellers on behalf of a client. The sellers have the possibility to (unconditionally) make a gift to the decision maker. Malmendier and Schmidt show that the subjects strongly respond to the gift despite the fact that a) its value is small and b) they understand the (selfish) intentions of the gift giver. Subjects reciprocate to gifts even in the absence of monetary incentives for doing so or when gift

---

<sup>5</sup>See Abbink (2006) and Abbink and Serra (2012) for extensive overviews of existing laboratory studies on bribery.

<sup>6</sup>For example, in their experiment participants have to simultaneously split their initial endowment between a bribe to the monitor and a contribution to the public good whereas in our design i) the contribution and the bribe decisions are taken sequentially and ii) different endowments are used at the contribution stage and the bribery stage. Other important differences in design include the role given to the monitor in the contribution stage (in our setting the monitor does not contribute to the public good while in their experiment the monitor either can or is forced to contribute to the public good), the possibility to decline the bribes (impossible in our design, possible in theirs) and the procedure used to match the subjects between the rounds (we implement a partner matching design whereas they implement a stranger matching design), amongst others.

giving imposes large externalities on third-parties (i.e. the client).

Gneezy et al. (2016) investigates the origins of bribery. They report on an experiment aimed at distinguishing whether greed or reciprocity drives bribery. In their setup two participants compete for a prize. The participants have the possibility to send a bribe to a third participant (the referee) who then has to decide who wins the prize. In one treatment the referee only keeps the bribe of the participant to whom the prize is attributed. In the other treatment the referee keeps both bribes. They show that the referees' judgment is mainly distorted when only the winners' bribe can be kept, thereby excluding reciprocity as the main driver of bribery.

Quite astonishingly and despite the large body of laboratory experiments related to corruption and bribery, our paper is the first to offer clean causal evidence on the effects of corruption on cooperation and to documents different channels through which these effects operate.<sup>7</sup>

Our findings also speak to the burgeoning literature studying cooperation in environments in which the punishment authority is centralized. Nicklisch et al. (2016) show that, when information about others' contributions is imperfect, a large part of the individuals prefer to delegate their sanctioning rights to a centralized authority, provided that this centralized authority does not punish cooperative individuals. Baldassarri and Grossman (2011) show, using a lab-in-the-field experiment, that a centralized sanctioning authority can foster high levels of cooperation and that a democratically elected monitor is more successful at sustaining cooperation than a randomly elected monitor. Kosfeld and Rustagi (2015) show that different types of real-world leaders, revealed using a social dilemma experiment with third party punishment, manage their forest commons very differently. While these studies highlight the benefits of centralized sanctioning, our results suggest that bribery largely undermines the capacity of a centralized authority to foster cooperation.

The paper is organized as follows. In Section 2, we outline the experimental design and the key behavioral predictions in Section 3. In Section 4, we present our findings and discuss their implications. Finally, Section 5 concludes the paper.

## 2 Experimental design

Our experiment consists of a modified public good game with two treatments. The participants are matched in groups of 4 that remain constant for the entire duration of

---

<sup>7</sup>Note that Weisel and Shalvi (2015) study the collaborative roots of corruption. Using a die rolling experiment, they show that collaborative settings (settings in which the payoffs of the players are aligned) “provide fertile ground for the emergence of corruption”, i.e. reduce the honesty with which players report the outcome of their die roll.



the experiment. Within a group, one participant is randomly selected to be the monitor. The monitor acts as a centralized sanctioning authority; It is the only player who has the power to assign punishment points to other participants. The other three participants are assigned the role of citizens. Their main activity is to decide how much to contribute to a public good.

The two treatments we implement differ in only one dimension: the possibility to transfer tokens to the monitor, i.e., to bribe her.<sup>8</sup> A period consists of 3 stages: 1) the contribution stage, 2) the bribery stage and 3) the punishment stage. In the BRIBE treatment, the citizens decide how many tokens to transfer to the monitor during stage 2. In the NO BRIBE treatment, the citizens do not have the possibility to transfer tokens to the monitor.

The experiment is repeated over 20 periods. In order to avoid reputation effects, the citizens' identity is reshuffled at every period.<sup>9</sup> This is a realistic assumption: While the citizens of a country generally know whether the authority is corrupted (e.g. whether the police tends to accept bribes), the authority does not necessarily know what to expect from each particular individual. Therefore, it is reasonable to focus on situations in which the monitor does not know the identity, respectively the past behavior, of its citizens.

Let  $\pi_i^t$  be the profit of citizen  $i$  at the end of stage  $t = \{1, 2, 3\}$ , and  $\pi_m^t$  denote the profit of the monitor. At the beginning of a period, each participant (i.e. the monitor and the citizens) receives an endowment of 20 tokens. In the contribution stage, each citizen has to (individually) decide how much of his endowment to invest in the public good. The difference between the endowment and the contribution to the public good is directly transferred to the citizen's private account. The monitor keeps its endowment and cannot contribute to the public good. Each token contributed to the public good benefits all the group members equally, i.e. the citizens and the monitor all receive the same monetary payoff out of the public good. The revenue from the public good consists of the sum of contributions multiplied by 0.4, the marginal per capita return to the public good. At the end of the contribution stage, the revenue generated by the public good and by the private account is privately disclosed to each citizen. Citizens are also made aware of their stage 1 profit  $\pi_i^1 = 20 - c_i + 0.4 \sum_{j=1}^3 c_j$ , where  $c_i$  is the contribution of citizen  $i$  to the public good.

At the beginning of the bribery stage, each citizen receives an additional endowment of 5 tokens. In the NO BRIBE treatment, neither the citizens nor the monitor are asked to take a decision. In the BRIBE treatment, the citizens must decide how many

---

<sup>8</sup>In our experiment, we purposefully did not use of the word "bribe" in order to avoid a framing that is too negatively connoted. Nevertheless, survey data collected at the end of our experiment suggests that virtually all the participants did perceive the transfers of tokens as bribes.

<sup>9</sup>For example, citizen 1 in period 1 is not necessarily called citizen 1 in period 2.

of these additional 5 tokens to *transfer* to the monitor. Following recent experiments on bribery, we assume that the monitor cannot refuse bribes (see e.g. Gneezy et al., 2016; Malmendier and Schmidt, 2017).<sup>10</sup>

The profits at the end of the bribery stage are

$$\begin{aligned}\pi_i^2 &= \pi_i^1 + 5 - b_i \cdot \mathbf{1}(\text{BRIBE}) \quad , \text{ for } i = 1, 2, 3 \\ \pi_m^2 &= 20 + 0.4 \sum_{j=1}^3 c_j + \sum_{j=1}^3 b_j \cdot \mathbf{1}(\text{BRIBE})\end{aligned}$$

where  $b_i \in [0, 5]$  denotes the bribe sent to the monitor by citizen  $i$  and  $\mathbf{1}(\text{BRIBE})$  is the indicator function taking the value 1 if the participant is in the BRIBE condition and zero otherwise.

The third stage corresponds to the punishment stage. The monitor starts by receiving an additional endowment of 15 tokens (out of which the punishment points can be allocated) and to learn each citizen's contribution. In BRIBE, the monitor also learns whether the citizens bribed her and by how much. The monitor is then asked to decide how many punishment points  $R_i$  to allocate to each citizen  $i$ . One punishment point reduces the profit of a citizen by 3 tokens. The monitor cannot allocate more than 10 punishment points to one citizen, i.e. she cannot reduce its profit by more than 30 tokens, and cannot allocate more than 15 punishment points per period ( $\sum_i R_i \leq 15$ ).<sup>11</sup> The final profits are

$$\begin{aligned}\pi_i^3 &= \pi_i^2 - 3R_i \quad , \text{ for } i = 1, 2, 3 \\ \pi_m^3 &= \pi_m^2 + 15 - \sum_{i=1}^3 R_i.\end{aligned}$$

At the end of a period, full feedback regarding the actions taken by all the citizens and the monitor is publicly disclosed to all the members of the group. That is, each citizen is provided with a summary of his own actions, the number of punishment points he received and his final profit in that particular period. Each citizen is also informed about the contributions and (if applicable) the bribery decisions as well as the

---

<sup>10</sup>We believe that this is a reasonable choice because people seem to only very rarely decline small bribes in reality, although they generally would have the possibility to do so (Malmendier and Schmidt, 2017; Gneezy et al., 2016). Gneezy et al. (2016) provide laboratory evidence corroborating this hypothesis. While their main treatments do not incorporate the possibility to decline bribes, they report evidence from two additional treatments that indicate that only a very small fraction of the participants (approximately 10%) declines bribes when given the possibility to do so. Moreover, they provide evidence that the behavior of their participants is orthogonal to their possibility to decline bribes. Hence, because virtually no one refuses bribes and because giving participants the possibility to decline bribes has no incidence on their behavior, we chose to force the monitors to accept bribes.

<sup>11</sup>See Figure S1 in Appendix for a screenshot of the monitor's decision screen.

punishment points received by the two other citizens of his group.<sup>12</sup>

Three important features of our experimental design are worth being discussed. To begin with, the choice to offer complete feedback to all the citizens at the end of a period might seem – at first – unrealistic. Indeed, people do not usually observe the bribes sent by others. However, people do generally have a fairly good idea of the average degree of corruption that prevails in the society they are living in. Their perception of corruption has generally been constructed over time, through experience (see e.g., Olken, 2009).<sup>13</sup> By choosing to fully disclose the actions of all the members of the group at the end of a period we shorten the social learning process, i.e. we reduce the time that would have been needed by the participants to learn what strategy is most commonly being played by the other participants.

Second, note that a bribe sent to the monitor can only make up to 25% of the monitor’s endowment. Hence, the bribes that we consider are relatively small.<sup>14</sup> As discussed earlier, we focus on small bribes because of the variety of bribery situations they encompass, e.g., transferring a small amount of money to a police officer or offering a basket of wine to an official in the hope of obtaining a favor in return. The decision to give every citizen a “bribery endowment” and to force them to only bribe out of this additional endowment of 5 tokens is purely practical: it ensures that their bribery decision is unconstrained by their profit in stage 1, thereby improving comparability between the citizens.

Third, note that the monitor has a selfish interest in being part of a group in which citizens highly contribute to the public good. She is therefore expected to take action in order to foster high levels of cooperation in her group, i.e. punish non-contributors. This modeling choice, although highly stylized, is an accurate representation of many real life situations involving a centralized sanctioning authority. For example, the leader of a tribe – or the mayor of a small town – has an interest in promoting and sustaining high levels of cooperation because both herself and her community reap the benefits of living in a well-functioning society.

---

<sup>12</sup>See Figure S2 in Appendix for a screenshot of the end-of-period screen.

<sup>13</sup>Olken (2009) uses field data of a road-building project in an Indonesian village to measure the accuracy of corruption perceptions. His results suggest that villagers’ perception of corruption is a good predictor of a more objective measure of corruption: ‘missing expenditures’ in the project. While relying solely on measures of corruption perceptions (such as the Transparency International Corruption Index (CPI) or World Bank Governance Indicators) can be misleading, Olken’s results suggest that reported perceptions of corruption reliably predict actual levels of corruption.

<sup>14</sup>Note that Gneezy et al. (2016) and Malmendier and Schmidt (2017) also focus on bribes that are “small gifts”.

## Instructions and framing

At the beginning of the experiment, participants received a set of instructions describing the decision stages and the payoffs at stake for the citizens and the monitor.<sup>15</sup> The instructions were neutrally framed. A citizen was referred to as a participant A, the monitor was referred to as participant B and a bribe was referred to as a “*transfer to participant B*”. Note that, within a treatment, both citizens and monitors received the same set of instructions.

## 3 Behavioral predictions

The experimental design outlined above leads us to formulate the following behavioral predictions.

**Prediction 1.** *Cooperation in BRIBE is lower than in NO BRIBE.*

While we expect the monitor in the NO BRIBE condition to be able to sustain high levels of cooperation,<sup>16</sup> the possibility to bribe the monitor (BRIBE condition) is predicted to dampen cooperation. We expect two channels to simultaneously undermine cooperations rates in BRIBE: lower punishment of the free-riders (prediction 2) and a discouragement effect (prediction 3).

**Prediction 2.** *Bribery softens punishment, which reduces the monitor’s ability to discipline free-riders and to foster cooperation.*

Following the literature on reciprocity, we expect bribed monitors to assign weaker sanctions to bribers.<sup>17</sup> In turn, we predict that these weaker sanctions will not discipline low contributors as much as tougher sanctions would (see e.g. Fehr and Gächter, 2000), i.e. we expect the disciplining power of punishment to be lower in BRIBE. Over time, weaker sanctions might increase the tendency of citizens to free-ride in the contribution stage. *Ceteris paribus*, if a low contributor (i.e., a free-rider) can avoid a severe punishment by bribing the monitor, then free riding on the contribution of the other members of the group and bribing the monitor might become an attractive strategy.<sup>18</sup>

---

<sup>15</sup>See the Appendix for a translated version of the instructions.

<sup>16</sup>Indeed, previous literature has documented that a centralized sanctioning authority can foster cooperation (see for example Baldassarri and Grossman, 2011).

<sup>17</sup>For key contributions on reciprocity, see Berg et al. (1995) and Fehr et al. (1997), among others.

<sup>18</sup>To illustrate how bribing might be profitable and part of an equilibrium, consider two strategies available to a citizen. He can either (i) contribute  $c_i^0 \geq 0$  to the public good and transfer nothing ( $b_i^0 = 0$ ) to the monitor or (ii) fully free-ride (i.e.,  $c_i = 0$ ) and bribe the monitor an amount  $b_i > 0$ . Furthermore, suppose that he expects to receive  $R_i^0$  punishment points if he does not bribe and  $R_i^b < R_i^0$  if he bribes

**Prediction 3.** *Bribery discourages initially high contributors.*

Finally, we expect the prevalence of bribery to discourage initially high contributors. Because citizens have perfect information about the contributions, the bribes and the punishment points received by the other members of their group, observing that free-riders can get away with low punishment might ultimately discourage them from contributing high amounts in the subsequent rounds.

## 4 Results

We conducted 7 experimental sessions at the University of Cologne in July 2016. A total of 224 subjects participated in our experiment which was fully computerized using z-Tree (Fischbacher, 2007). Subjects were recruited using ORSEE (Greiner, 2015) and were allowed to take part in the experiment only once. Payoffs were converted in Euros (EUR) at an exchange rate of 100 tokens = EUR 1.20. On average, a session lasted 75 minutes and participants earned EUR 13.80, including a show-up fee of EUR 4. Treatments were randomized at the session level: half of the participants were assigned to the BRIBE treatment while the other half was assigned to the NO BRIBE treatment. Tables S1 and S2 in Appendix compare participants' observable characteristics across conditions. While some slight differences exist, these tables provide evidence that both monitors and citizens were well balanced across treatments. Overall, our data comprises 56 independent groups of 4 subjects (28 groups per treatment).

**Result 1.** *Contributions are on average 33% lower in BRIBE than in NO BRIBE.*

Figure 1 depicts the average contribution in the BRIBE and the NO BRIBE treatments. Overall, subjects in BRIBE contribute an average of 4.14 tokens less – i.e. approximately 33% less – than subjects in NO BRIBE (Wald test,  $p < 0.01$ , OLS regressions column 1 of Table S3 in the Appendix). This result is robust to an OLS estimation which controls for period fixed effects and individual covariates (Wald test,  $p < 0.01$ , see column 2 of Table S3 in the Appendix). Interestingly, the average initial contribution is approximately equal to 10, i.e. 50% of the endowment, in both treatments. This suggests that the treatments do not affect the behavior of the citizens from the outset of the experiment. This is indeed the case: While the average contribution increases by the monitor. Given these beliefs and the profit function defined above, bribing is profitable if and only if

$$20 + 5 - b_i + 0.4 \sum_{j \neq i} c_j - 3R_i^b > 20 + 5 - c_i^0 + 0.4c_i^0 + 0.4 \sum_{j \neq i} c_j - 3R_i^0 \Leftrightarrow b_i < 0.6c_i^0 + 3(R_i^0 - R_i^b),$$

i.e. if the bribe is lower than the cost of contributing,  $0.6c_i^0$ , and the additional punishment that can be expected from not bribing  $3(R_i^0 - R_i^b)$ .

50% over time in NO BRIBE (from 9.56 tokens in period 1 to 14.3 tokens in period 19, Wald test,  $p < 0.01$ ), it stays constant in BRIBE (from 10 tokens in period 1 to 9.05 tokens in period 19, Wald test,  $p = 0.43$ ).<sup>19,20</sup> These dynamics strike in very early in the experiment: In period 4 already, the average contribution in NO BRIBE exceeds the average contribution in BRIBE (Wald test  $p = 0.08$ ). This difference in contribution increases both in size and in significance over time ( $p < 0.05$  between periods 6 and 10,  $p < 0.01$  after period 10).

Two important findings are worth being highlighted. First, contributions continuously increase over time in NO BRIBE. This finding is in line with Baldassarri and Grossman (2011) who show that the presence of a monitor (either democratically elected or randomly chosen) significantly increases cooperation compared to a baseline condition with no monitor. Second, cooperation does not collapse in BRIBE. This finding suggests that cooperation and corruption can coexist. This result is in line with casual observation: corrupted countries generally still provide some positive level of public good to their citizens.<sup>21</sup>

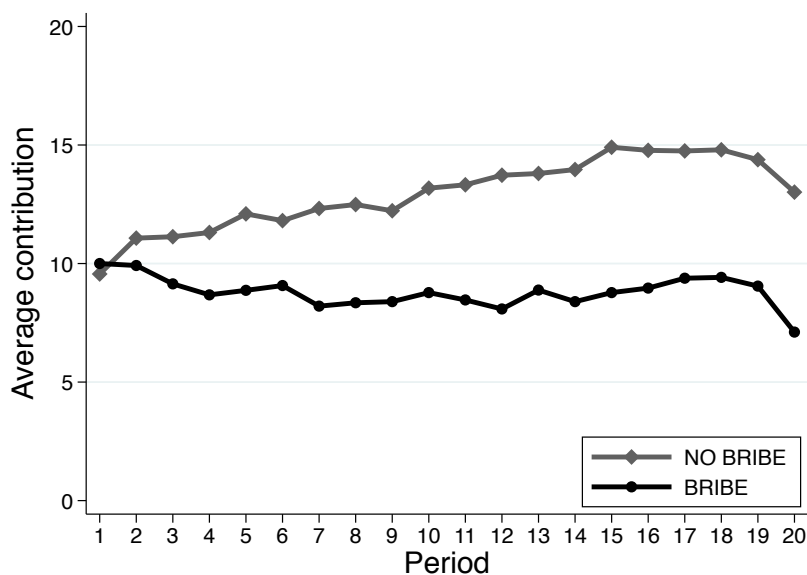


Figure 1: Average contributions over time (by treatment)

<sup>19</sup>Note that we do not consider the last period in order to account for the often-discussed end-game effects.

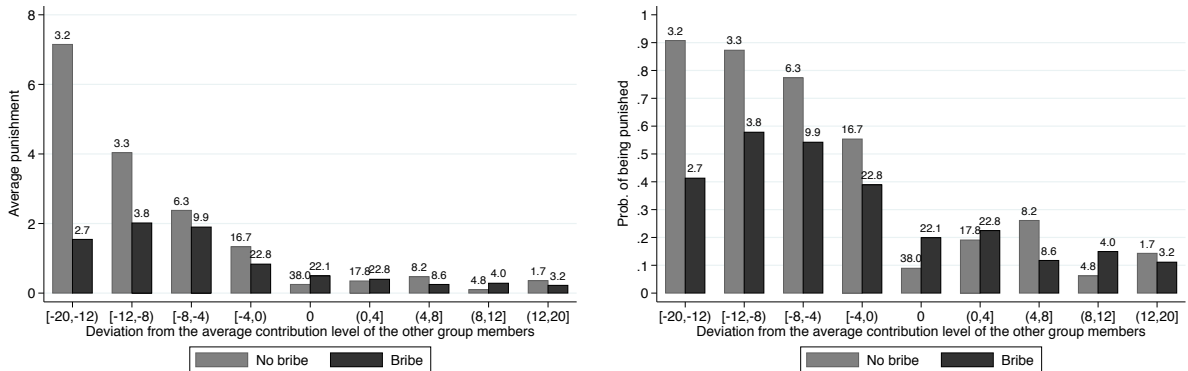
<sup>20</sup>OLS regressions of contributions on a linear time trend indicate a positive and significant trend in NO BRIBE (Wald test,  $p < 0.01$ ) and no trend in BRIBE (Wald test,  $p = 0.48$ ).

<sup>21</sup>As noted by Johnston (1998), “the most serious cases of corruption are entrenched political and bureaucratic corruption and such systems are tightly organized, internally stable and do not automatically result in collapse.”

**Result 2.** *Punishment is much softer in BRIBE than in NO BRIBE.*

In what follows, we investigate whether the different dynamics of cooperation observed across the two treatments are driven by the citizens' response to the behavior of the monitor. In particular, we investigate whether lower cooperation in BRIBE results from less severe punishment. We show that punishment in BRIBE is much softer than in NO BRIBE. More specifically, we show that BRIBE weakens punishment both at the *intensive* and the *extensive margin*. That is, not only does a citizen receive fewer punishment points for a given deviation from the average contribution of the other members of the group than in BRIBE than in NO BRIBE, he is also less likely to be punished.

We start with an analysis of the punishment points at the intensive margin. Figure 2a plots the average punishment points inflicted to a citizen as a function of his deviation (positive or negative) from the average contribution of the other members of his group. For example, a citizen that contributed between 12 and 20 tokens less than the average contribution of the other members of his group received an average of 7.15 punishment points in the NO BRIBE condition. Consistent with previous findings (see e.g. Fehr and Gächter, 2000), punishment is mostly targeted at below-average contributors in both treatments. Notice also that larger negative deviations from the group average result in larger punishments and that a zero deviation is virtually not sanctioned.<sup>22</sup>



(a) Punishment points received.

(b) Probability of being punished.

Figure 2: (a) Punishment points and (b) probability of being punished as a function of the deviation from the average contribution of the other group members. Numbers above bars indicate the relative frequency (in %) of observations in the different intervals within a treatment.

Figure 2a also reveals that there is a clear difference between the treatments: While individuals that contribute the group average or more than the group average are

<sup>22</sup>Note that the data reveal instances of anti-social punishment (0.34 punishment points of above-average deviations in treatment NO BRIBE, Wald test  $p < 0.01$ ).

not punished significantly more heavily in one treatment than in another (Mann-Whitney U tests,  $p = 0.37$  for those contributing the group average and  $p = 0.93$  for those contributing more than the group average), the punishment of below-average contributors is much softer in BRIBE than in NO BRIBE (Mann-Whitney U test,  $p < 0.01$ ).<sup>23</sup> For example, a negative deviation of 8 to 12 tokens from the group average is sanctioned by an average of 4.04 punishment points in NO BRIBE while it is sanctioned by an average of only 2.02 points in BRIBE (Wald test,  $p = 0.01$ ). Note that the difference between the average sanction inflicted to citizens in NO BRIBE and in BRIBE increases in the size of the (negative) deviation from the average contribution of the other members of the group, i.e., it is mostly large free-riders that benefit—on average—from an environment in which bribery is possible. These results are broadly supported by OLS regressions, as documented in Table 1. In particular, below-average contributors are sanctioned significantly less harshly (about 1.3 token less) in BRIBE compared to NO BRIBE (column 1,  $p < 0.01$ ). This result holds also after controlling for the average contribution of the other group members, period fixed effects and the individual characteristics of the monitor (column 2,  $p < 0.05$ ).

Very similar findings are reported at the extensive margin, as depicted in Figure 2b. The probability of being punished is much higher for below-average contributors than for citizens that contributed at least the group average (or more). However, BRIBE largely decreases the likelihood that a below-average contributor is punished. For example, a citizen that contributed between 12 and 20 less than the average is punished in 90.7 % of the cases in NO BRIBE, this probability drops to 41.3 % in BRIBE (Wald test,  $p < 0.01$ ). These results are robust to OLS regressions (see Table 1, columns 3-4): On average, a below-average contributor is 23.7 to 28.2 percentage-points less likely to be punished in BRIBE than in NO BRIBE ( $p < 0.01$ ).

---

<sup>23</sup>These results are supported by a two-sample Kolmogorov-Smirnov test of identical distributions ( $D = 0.33$ ,  $p = 0.01$ ).



Table 1: Effects of BRIBE on punishment (intensive and extensive margin)

	Severity of punishment		Probability of punishment	
	(1)	(2)	(3)	(4)
Treatment Bribe	0.103 (0.114)	0.082 (0.151)	0.055 (0.045)	0.027 (0.046)
Below-average contribution	2.199*** (0.455)	2.158*** (0.472)	0.541*** (0.067)	0.511*** (0.064)
Below-average contribution $\times$ Bribe	-1.328*** (0.493)	-1.270** (0.514)	-0.282*** (0.084)	-0.237*** (0.079)
Others' average contribution ( $\bar{C}_{-i,t}$ )		-0.029** (0.013)		-0.011*** (0.003)
Constant	0.292*** (0.072)	0.333 (0.535)	0.134*** (0.029)	0.128 (0.110)
Periods F.E.	No	Yes	No	Yes
Controls	No	Yes	No	Yes
$R^2$	0.191	0.222	0.185	0.250
# Clusters	56	56	56	56
Observations	3360	3360	3360	3360

*Notes:* OLS estimations. Standard errors (clustered at the group level) are displayed in parentheses. In columns 1-2, the dependent variable is number of punishment points assigned to citizens (0 to 10) in  $t$ , in columns 3-4 the dependent variable is 1 if the citizen received punishment points, and 0 otherwise. Below-average contribution is a dummy variable. Controls include monitor-specific dummies for gender, German mother tongue and economics major. \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

Overall, these findings provide compelling evidence that the monitors' behavior is largely affected by the treatment. While they tend to sanction the free-riders often and harshly in NO BRIBE, they are much more lenient in BRIBE. These findings are particularly remarkable when considered in light of Result 1. Because the average contribution is much lower in BRIBE than in NO BRIBE, one could have expected the punishment inflicted to below-average contributors in BRIBE to be tougher than in NO BRIBE. The fact that we observe a significant reduction in the punishment of below-average contributors between the two treatments despite lower contributions in BRIBE suggests that the punishment motives are indeed different in the two treatments.

**Result 3.** *In BRIBE, the punishment is too soft to discipline citizens.*

Figure 3 displays the average change in contribution from period  $t - 1$  to period  $t$ , conditional on having been punished in  $t - 1$  or not. The left-hand side of the figure indicates that subjects who did not receive punishment points in  $t - 1$  significantly reduce their contributions in  $t$  in both treatments ( $-0.68$  in NO BRIBE and  $-0.7$  in BRIBE, Wald tests, both  $p < 0.001$ ). These decreases in contribution are not significantly different from

each other (Wald test,  $p = 0.91$ ), indicating that the dynamics of contributions of citizens that do not receive punishment points are similar in the two treatments (OLS regressions confirm this finding, see Table S4 columns 1-2 in Appendix).

The right-hand side of the figure depicts the extent to which contributions increase after having received punishment points in the previous period. After having been assigned punishment points in  $t - 1$ , subjects significantly increase their contribution in  $t$  in the two treatments ( $+2.19$ ,  $p < 0.001$  in NO BRIBE and  $+1.15$ ,  $p < 0.001$  in BRIBE). However, the increase in contribution is twice as large in NO BRIBE than in BRIBE. This difference in increase in contributions is highly significant (Wald test,  $p < 0.01$ ) and provides compelling evidence that the punishment in BRIBE is too soft to discipline citizens.<sup>24</sup> These results are robust to OLS regressions including a variety of control variables such as group average contribution in the previous period, period fixed effects and demographics (see Table S4, columns 3-4 in Appendix).

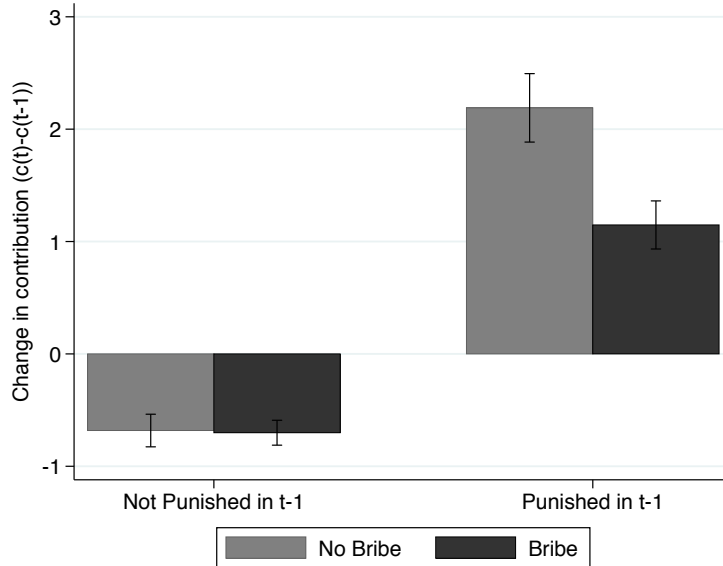


Figure 3: Average change in contribution ( $c_t - c_{t-1}$ ) by treatment, with clustered standard errors at the group level. “Not punished in  $t - 1$ ” indicates citizens that did not receive punishment points in  $t - 1$ . “Punished in  $t - 1$ ” indicates citizens that received any positive punishment point in  $t - 1$ .

**Result 4.** *Citizens who contribute more than the average in period 1 (initially high contributors) dramatically reduce their contribution in BRIBE but not in NO BRIBE, sug-*

<sup>24</sup>Note that Baldassarri and Grossman (2011) document an average response to punishment in their randomly elected monitor condition that is very similar to ours. While in our NO BRIBE condition being punished in  $t - 1$  translates into an approximate 15% increase in contribution in  $t$ , Baldassari and Grossman report that punishment leads to an average increase in contributions of 12%.

gesting that they get discouraged in BRIBE.

In order to study the dynamics of cooperation, we divide our sample into two categories: citizens that contributed more than the average contribution of the other members of their group in period 1 (*initially high contributors*) and citizens that contributed less (*initially low contributors*). We then plot the evolution of the average contribution of these two types of individuals as a function of the treatment. In addition, we also plot the average punishment received by these two types of citizens in the two different treatments. The results are depicted in Figure 4.<sup>25</sup>

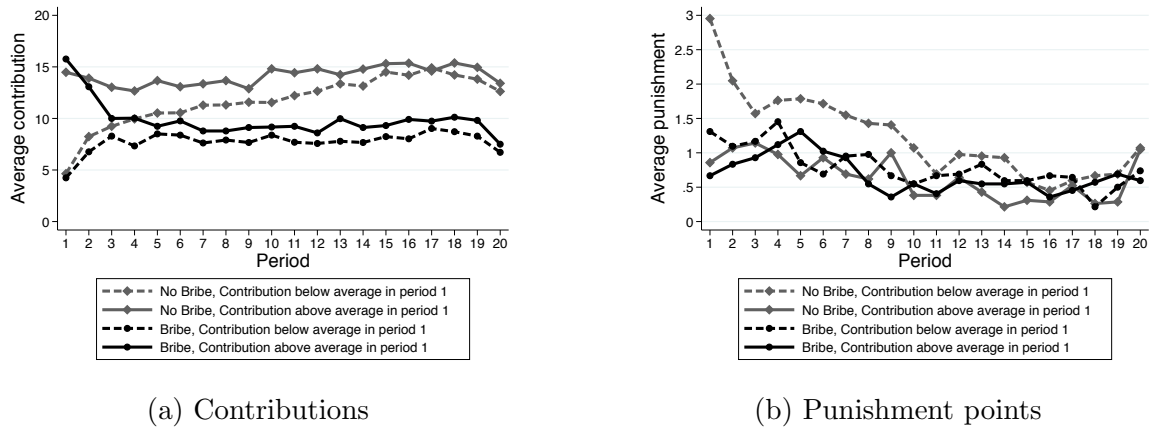


Figure 4: Evolution of (a) contributions and (b) punishment points over the 20 periods for *initially* low or high contributors in BRIBE and NO BRIBE. An initially low (high) contributor is defined as a citizen contributing strictly less (weakly more) than the average contribution of the other group members in period 1.

Two important findings emerge from these panels. First, note that the average contribution in period 1 of initially high contributors (solid lines) is equal to approximately 75% of their endowment, both in BRIBE and in NO BRIBE (Figure 4a). Similarly, the average contribution in period 1 of initially low contributors (dashed lines) corresponds to approximately 25% of their endowment in both treatments. Hence, the treatment does not affect the behavior in period 1 of these two types of citizens. However, the treatment immediately affects the behavior of the monitors, as indicated in Figure 4b. Indeed, the average punishment received by an initially low contributor in period 1 in the NO BRIBE condition is much higher than the average punishment received by an initially low contributor in BRIBE. In period 1, initially low contributors in NO BRIBE

<sup>25</sup>As a robustness check, note that we also divided the sample depending on how many times a citizen contributed more than the average of the other group members over either the first 2 or the first 3 periods (see Figures S4a and S4b in Appendix). We also performed the analysis by defining a high contributor as a citizen whose contribution in period 1 is higher than the average contribution at the *treatment* level (see Figure S4c) or higher than the group median contribution in period 1 (Figure S4d). All these robustness checks yield very similar results.

receive a punishment that is on average of 2.25 times higher than in BRIBE (Wald test,  $p < 0.01$ ), despite the fact that their average contribution to the public good are virtually identical.<sup>26</sup> As for initially high contributors, they are not sanctioned differently in the two treatments.<sup>27</sup>

Second, initially low contributors largely increase their contribution to the public good in NO BRIBE, whereas they keep on contributing only very little in BRIBE. Indeed, striking is that, in NO BRIBE, initially low contributors rise their contribution from 5 in period 1 to 12.62 in period 20. As of period 8, the average contribution of initially low and initially high contributors are statistically indistinguishable from each other in NO BRIBE (Wald test,  $p = 0.16$ ). The picture is very different in BRIBE as it is initially high contributors who strongly decrease their contributions down to the level of initially low contributors. While conditional cooperation can explain why initially good contributors reduce their contributions after observing free-riding in their group, it cannot explain why this phenomenon only occurs in the BRIBE condition. We speculate that this drop in contribution is the consequence of a discouragement effect: initially high contributors in BRIBE get discouraged to see that low contributors get away from harsh sanctions. This discouragement quickly leads them to reduce their contribution dramatically. Indeed, between period 1 and period 3 the average contribution of an initially high contributor in BRIBE drops from 15.76 to 10 (Wald test,  $p < 0.001$ ). In this treatment, the average contribution of initially high contributors is statistically indistinguishable from the average contribution of initially low contributors as of period 3 already.

Quite astonishingly, the contributions of all the citizens in BRIBE converge towards an equilibrium involving low levels of cooperation, whereas all the citizens in NO BRIBE converge towards an equilibrium in which cooperation is high. Interestingly, it is the initially low contributors that catch up with the initially high contributors in NO BRIBE, whereas it is the initially high contributors that decrease their contributions down to the levels of initially low contributors in BRIBE.

As initially low contributors increase their contributions in NO BRIBE, the punishment points they receive gradually drop (Figure 4b). The picture is less clear in BRIBE. While initially high contributors get more heavily punished as their contributions drop from period 1 to period 4, the average deduction they receive decreases again after period 4 despite the fact that they do not increase their contribution.

Up to this point, it is striking that the BRIBE treatment largely prevents contributions to increase over time, as opposed to what is observed in the NO BRIBE condition.

---

<sup>26</sup>Initially low contributors in the NO BRIBE condition receive on average 2.95 punishment points in NO BRIBE in period 1, while they receive an average of 1.31 punishment points in BRIBE.

<sup>27</sup>Initially high contributors receive on average 0.86 punishment points in the NO BRIBE condition while they receive an average of 0.66 punishment points in BRIBE ( $p = 0.65$ ).

The BRIBE condition dramatically softens the sanctions inflicted to below-average contributors, which undermines their disciplining power. It also discourages initially high contributors to continue to highly contribute to the public good. However, for bribery to be the main driver of these results, the extent to which citizens actually bribe their monitor still needs to be established. In what follows, we investigate the extent to which bribery is used (Result 5), whether it helps to prevent punishment and how it affects profits (Result 6).

**Result 5.** *Citizens condition their bribes on their deviation from the average contribution of the other group members.*

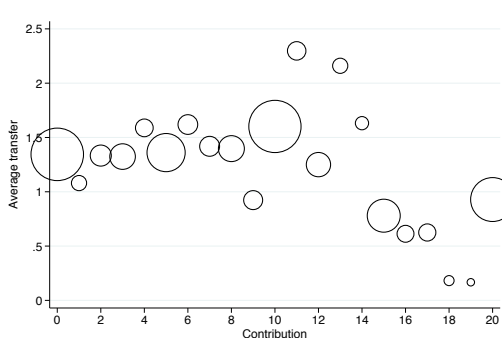
Throughout the experiment, citizens made a large use of the bribes. They transferred a positive amount of tokens to the monitor in almost 60 % of the cases and the maximum amount (5 tokens) in about 7 % of the cases (see Figure S3 in Appendix). The average transfer is 1.30 (i.e., 26% of the transfer endowment) and is constant over the 20 periods (OLS trend test,  $p = 0.85$ ). These findings are broadly consistent with the bribery rates documented in previous studies (albeit involving very different experimental designs). For example, Gneezy et al. (2016) report bribery rates ranging from 44% to 74%. In Abbink et al. (2002), subjects enter a bribery relationship in 22% to 25% of the cases and bribe the full amount in 12% of the cases.

Figure 5a indicates that there is no strong substitution between the average contribution to the public good and the average transfer sent to the monitor (Wald test,  $p = 0.11$ ).<sup>28</sup> Although surprising, this finding is in line with previous studies (see e.g. Gneezy et al., 2016). Nevertheless, Figure 5b indicates that the transfers depend on the deviation from the average contribution of the group: the average transfer sent to the monitor by below-average contributors is on average 30% higher than the transfer sent by above-average contributors ( $t - test = 2.91$ ,  $p < 0.01$ ).<sup>29</sup>

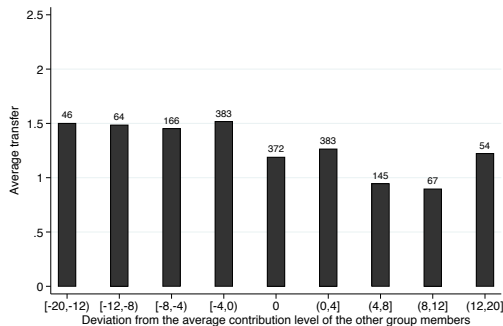
---

<sup>28</sup>While Figure 5a might suggest that the relationship between average transfers and contribution is zero up to contribution level 10 and slightly negative from 10 to 20, the negative slope is only very marginally significant (Wald test,  $p = 0.1$ ).

<sup>29</sup>The average transfer of an above-average contributor is 1.15 tokens whereas the average transfer of a below-average contributor is 1.5.



(a) Average transfers and contribution levels



(b) Average transfers and contribution deviations

Figure 5: Average transfers and contributions (panel a) and average transfer over contribution deviations (panel b). The size of the dots in panel a indicate the relative frequency of observations for each of the different contributions levels.

So far, we have established that below-average contributors receive fewer punishment points in BRIBE (Result 2) and that they bribe significantly more the monitor than above-average contributors (Result 5). An important question which remains to be answered is how this strategy affects their profit.

**Result 6.** *Below-average contributors can decrease punishment through bribery, but this strategy does not increase their profit.*

In Table 2, we regress the punishment points received (columns 1 to 6) and the profit of the citizens (columns 7 to 12) on their contributions and their bribes. We run separate regressions for citizens that contribute less, exactly or more than the average of their group. Columns 1 and 2 document that below-average contributors benefit the most from bribing the monitor: An additional token transferred to the monitor reduces the number of punishment points they receive by 0.17 ( $p = 0.06$ ). In contrast, citizens that contribute the average (columns 3-4) or more (columns 5-6) do not receive significantly fewer punishment points if they increase their transfers to the monitor. Interestingly, higher contributions to the public good would allow each type of citizen to reduce the number of punishment points received.

While the bribes allow below-average contributors to reduce the number of punishment points received, they do not significantly affect their profits (columns 7-8). Hence, it is not a profit-enhancing strategy to voluntarily free-ride on the contribution of others and subsequently bribe the monitor. A better strategy for below-average contributors would be to increase their contributions to the public good. Finally, note that bribes decrease the profit of citizen that contribute at least the average of their group, as indicated by columns 9 to 12 ( $p < 0.05$  and  $p < 0.01$ ).

Table 2: The effect of transfers and contributions on punishment points and payoffs

	Punishment						Profit					
	$C_{i,t} < \bar{C}_{-i,t}$		$C_{i,t} = \bar{C}_{-i,t}$		$C_{i,t} > \bar{C}_{-i,t}$		$C_{i,t} < \bar{C}_{-i,t}$		$C_{i,t} = \bar{C}_{-i,t}$		$C_{i,t} > \bar{C}_{-i,t}$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Transfer	-0.175* (0.088)	-0.173* (0.088)	-0.101 (0.105)	-0.075 (0.082)	0.029 (0.031)	0.038 (0.031)	-0.387 (0.233)	-0.411 (0.257)	-0.697** (0.315)	-0.786*** (0.252)	-1.008*** (0.210)	-0.952*** (0.187)
Contribution	-0.091** (0.033)	-0.116*** (0.036)	-0.069*** (0.024)	-0.092*** (0.025)	-0.019* (0.010)	-0.027** (0.012)	0.183* (0.095)	0.216** (0.093)	0.408*** (0.072)	0.469*** (0.074)	-0.117** (0.053)	-0.078 (0.060)
Constant	2.002*** (0.456)	1.536 (0.919)	1.335*** (0.394)	2.449* (1.432)	0.524*** (0.182)	0.444* (0.227)	24.049*** (1.326)	25.826*** (3.568)	20.995*** (1.183)	17.090*** (4.189)	23.928*** (0.508)	19.506*** (1.459)
Periods F.E.	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
$R^2$	0.080	0.133	0.141	0.351	0.019	0.070	0.027	0.129	0.484	0.613	0.139	0.276
# Clusters	28	28	28	28	28	28	28	28	28	28	28	28
Observations	659	659	372	372	649	649	659	659	372	372	649	649

*Notes:* OLS estimations. The dependent variable is either punishment points received (columns 1 to 6) or the period-profit (columns 7 to 12). Standard errors (clustered at the group level) are displayed in parentheses. Controls include citizen-specific (and monitor-specific, for columns 7-12) dummies for gender, German mother tongue and economics major. Levels of significance: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

Up to this point, our study has focused on the analysis and the interpretation of subjects' behavior. In order to gain additional insights into the motives driving participants' actions, we now turn to subjects' responses to a qualitative questionnaire distributed at the end of the experiment. Since we are mainly interested in understanding the causes and consequences of corruption, we restrict the following analysis to the subjects in the BRIBE treatment.

**Result 7.** *Citizens expect low contributors to use bribes to avoid punishment and expect monitors to be more lenient with bribers. Monitors find it normal to reciprocate bribes despite the fact that they understand that citizens use them to avoid punishment.*

From the citizens' perspective, it is clear that the bribes are used in the hope of reducing the number of punishment points received, as indicated by figure 6a. Moreover, the vast majority of the citizens in the BRIBE condition (more than 60%) find it normal that the monitor reciprocates the bribes by assigning fewer punishment points (Figure 6b).<sup>30</sup> Interestingly, the monitors' beliefs are quite similar: despite the fact that most of them understand that citizens bribe in the hope of receiving fewer punishment points (Figure 6c), they nevertheless find it normal to reciprocate (Figure 6d).<sup>31</sup>

Overall, our results suggest that citizens purposefully use bribes in order to reduce the amount of punishment points they receive and that monitors find it normal to

<sup>30</sup>The exact question was “*It is normal to assign fewer deduction points to a participant A who makes a positive transfer.*”

<sup>31</sup>About 70% of the monitors in BRIBE agree with the statement “*Citizens who don't contribute enough to the project bribe in the hope of receiving fewer deduction points.*”

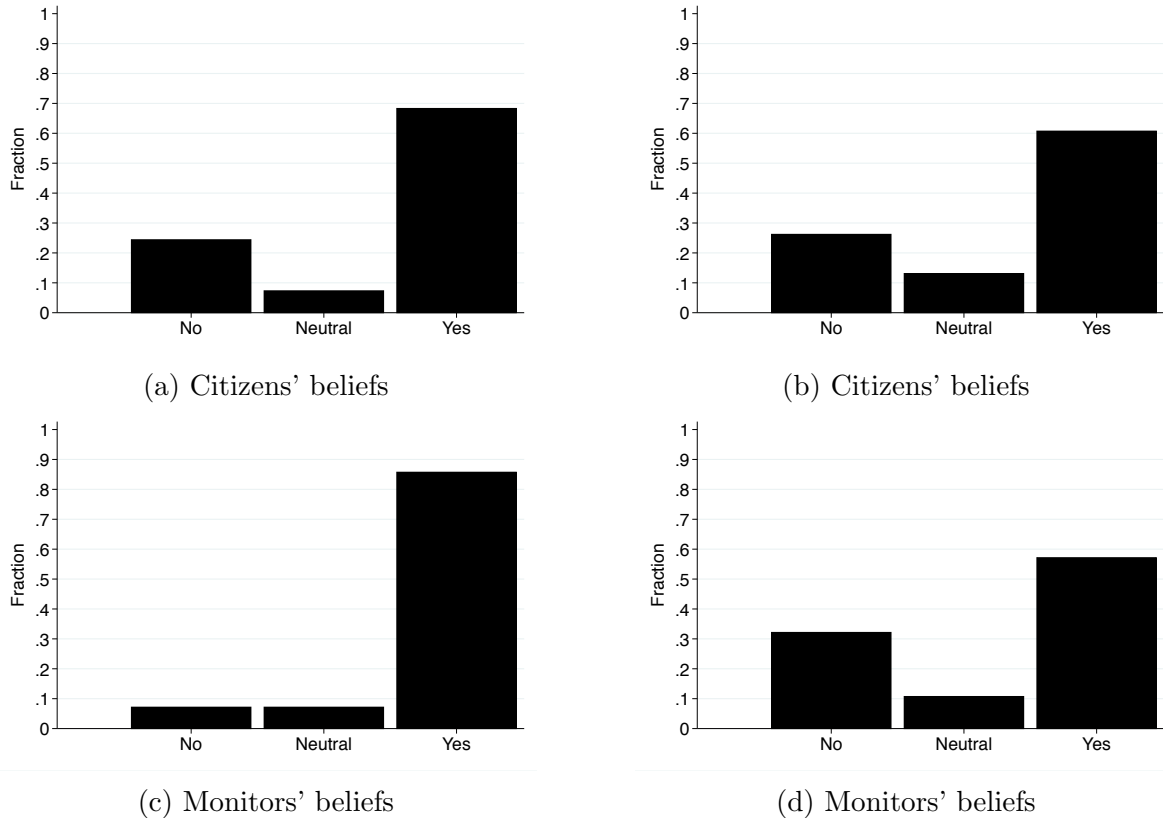


Figure 6: Citizens' (panel a and b) and monitors' beliefs (panel c and d). Panel (a) and (c): Subjects' agreement with the statements "It is normal that participant B [*the monitor*] assigns fewer deduction points to a participant A [*a citizen*] who transfers tokens". Panel (b) and (d): Subjects' agreement with the statement "Participants A [*Citizens*] transfer tokens to participant B [*the monitor*] in the hope of receiving fewer deduction points". Answers range from 1 ("I fully disagree") to 7 ("I fully agree") and are collapsed into three categories: "No" (1-3), "Neutral" (4) and "Yes" (5-7). Words in bracket were not displayed in the original questionnaire.

reciprocate bribes by allocating fewer punishment points to bribers. Citizens's behavior suggests that bribes are mostly used by below-average contributors, who indeed manage to decrease the number of punishment points they receive. In BRIBE, initially high contributors get discouraged to see that initially low contributors do not get punished. These mechanisms prevent cooperation rates to increase over time, as opposed to what is observed in the NO BRIBE treatment.

Before drawing final conclusions, we investigate the long run profit consequences of both treatments.

**Result 8.** *In the long run, citizens are better off in NO BRIBE. Monitors' profit are unaffected by the treatment. Overall efficiency is higher in NO BRIBE.*

Figure 7 depicts the evolution of the average payoff of citizens (panel a) and



monitors (panel b) by treatment. Over the first 10 periods, the average profit of citizens in BRIBE and in NO BRIBE are not significantly different from each other ( $p = 0.40$ ). However, in the long run, citizens are better off in NO BRIBE. Indeed, over periods 11-20 a citizen earns about 8% more in NO BRIBE than in BRIBE (32.06 vs 29.6,  $p < 0.01$ ). Monitors' payoff remain unaffected by the treatment, as indicated by panel b. The overall efficiency (See Figure S5 in Appendix), measured as the average payoff of all the members of a group, is significantly higher in NO BRIBE than in BRIBE (taking periods 11 to 20,  $p = 0.01$ ).

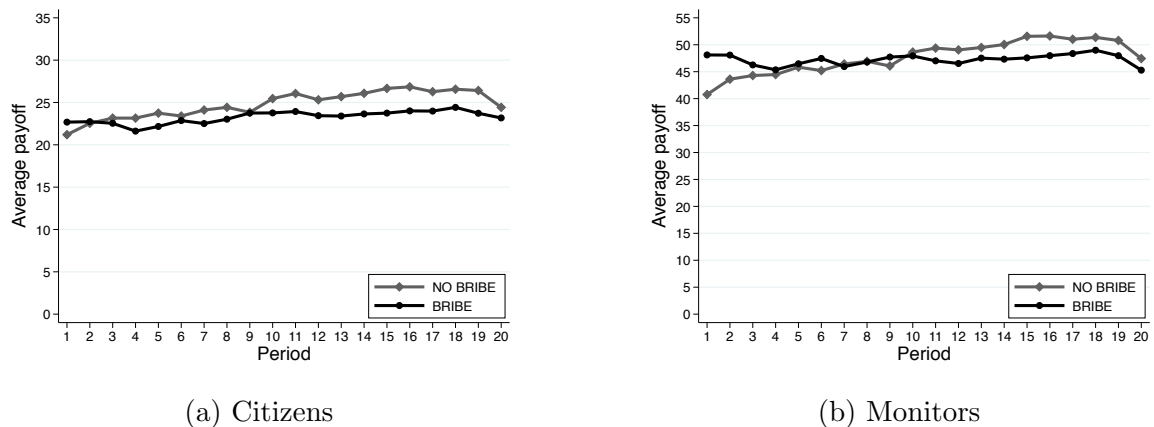


Figure 7: Evolution of payoffs by treatment for (a) citizens and (b) monitors.

## 5 Conclusion

Corruption is a widespread phenomenon. In some countries, it has for long been thought to be one of the causes of low growth, poor investment rates, etc. However, clean causal evidence on the effects of corruption is still lacking. We take a first step in this direction by studying how corruption affects cooperation in a tightly controlled laboratory experiment. Using a public good game experiment in which the players either have the possibility to bribe the sanctioning authority (BRIBE treatment) or not (NO BRIBE treatment), we provide causal evidence that corruption largely undermines cooperation. In our experiment, an institutional setting that prevents bribery leads to cooperation rates that are approximately 50% higher than an environment in which bribery is possible.

In addition to this finding, our paper also makes precise two channels through which corruption undermines cooperation. First, corruption reduces both the frequency and the size of the punishment inflicted to low contributors, i.e. to free riders. This decrease in punishment has important consequences: Because low contributors do not feel threatened by subsequent punishment in the BRIBE treatment, they keep on contributing only very little to the public good. The second channel that we document is a

discouragement effect. Very rapidly, initially high contributors in the BRIBE condition reduce their contribution down to the level of initially low contributors, suggesting that they get discouraged by seeing that the free-riders can get away from punishment despite contributing very little to the public good.

To our knowledge, this paper is the first to provide clean causal evidence that corruption undermines cooperation and to document different channels through which this effect operates. While previous empirical studies have documented that corruption is negatively correlated with aggregate variables such as GDP and investment (see e.g. Mauro, 1995; Méon and Sekkat, 2005), our results unambiguously establish that corruption decreases cooperation. Our results also speak to the recent literature investigating the benefits of centralized sanctioning authorities in social dilemmas. While previous studies have shown that a centralized authority can successfully foster cooperation in groups at a relatively cheap cost compared to decentralized sanctioning systems, our results clearly highlight a weakness of centralized authorities: their sensitivity to corruption. An interesting avenue for future research would be to investigate whether a decentralized authority indeed does better at fostering high levels of cooperation in the presence of bribery than a centralized authority. We conjecture that this is the case, but leave the study of this particular question for future research.

## References

- Abbink, Klaus**, “Staff rotation as an anti-corruption policy: an experimental study,” *European Journal of Political Economy*, 2004, 20 (4), 887–906.
- , “Fair Salaries and the Moral Costs of Corruption. Advances in Cognitive Economics. BN Kokinov. Sofia,” 2005.
- , “14 Laboratory experiments on corruption,” *International Handbook on the Economics of Corruption*, 2006, p. 418.
- **and Danila Serra**, “Anticorruption policies: Lessons from the lab,” *New Advances in Experimental Research on Corruption*, 2012, 15.
- , **Bernd Irlenbusch**, and **Elke Renner**, “An experimental bribery game,” *Journal of Law, Economics, and Organization*, 2002, 18 (2), 428–454.
- Baldassarri, Delia** and **Guy Grossman**, “Centralized sanctioning and legitimate authority promote cooperation in humans,” *Proceedings of the National Academy of Sciences*, 2011, 108 (27), 11023–11027.
- Barr, Abigail** and **Danila Serra**, “Corruption and culture: An experimental analysis,” *Journal of Public Economics*, 2010, 94 (11), 862–869.
- Berg, Joyce**, **John Dickhaut**, and **Kevin McCabe**, “Trust, reciprocity, and social history,” *Games and Economic Behavior*, 1995, 10 (1), 122–142.
- Bertrand, Marianne**, **Simeon Djankov**, **Rema Hanna**, and **Sendhil Mullainathan**, “Obtaining a driver’s license in India: an experimental approach to studying corruption,” *Quarterly Journal of Economics*, 2007, pp. 1639–1676.
- Cagala, Tobias**, **Ulrich Glogowsky**, **Veronika Grimm**, **Johannes Rincke**, and **Amanda Tuset Cueva**, “Does Corruption Affect the Private Provision of Public Goods?,” *Working Paper*, 2016.
- Cameron, Lisa**, **Ananish Chaudhuri**, **Nisvan Erkal**, and **Lata Gangadharan**, “Propensities to engage in and punish corrupt behavior: Experimental evidence from Australia, India, Indonesia and Singapore,” *Journal of Public Economics*, 2009, 93 (7), 843–851.
- EU Commission**, “Report from the commission to the council and the European parliament–EU anti-corruption report,” 2014.

- Fehr, Ernst and Simon Gächter**, “Cooperation and Punishment in Public Goods Experiments,” *American Economic Review*, 2000, *90* (4), 980–994.
- , – , and **Georg Kirchsteiger**, “Reciprocity as a contract enforcement device: Experimental evidence,” *Econometrica*, 1997, pp. 833–860.
- Fischbacher, Urs**, “z-Tree: Zurich toolbox for ready-made economic experiments,” *Experimental Economics*, 2007, *10* (2), 171–178.
- Gneezy, Uri, Silvia Saccardo, and Roel Van Veldhuizen**, “Bribery: Greed versus reciprocity,” *WZB Discussion Paper No. SP II*, 2016.
- Greiner, Ben**, “Subject pool recruitment procedures: organizing experiments with ORSEE,” *Journal of the Economic Science Association*, 2015, *1* (1), 114–125.
- Huntington, Samuel P.**, *Political order in changing societies*, Yale University Press New Haven, CT, 1968.
- Johnston, Michael**, “What can be done about entrenched corruption?,” in “Annual world bank conference on development economics 1997” World Bank Washington DC 1998, pp. 69–90.
- Kosfeld, Michael**, “Corruption within a Cooperative Society,” Economics Series 48, Institute for Advanced Studies July 1997.
- and **Devesh Rustagi**, “Leader punishment and cooperation in groups: Experimental field evidence from commons management in Ethiopia,” *American Economic Review*, 2015, *105* (2), 747–783.
- Malmendier, Ulrike and Klaus Schmidt**, “You Owe Me,” *American Economic Review (forthcoming)*, 2017.
- Mauro, Paolo**, “Corruption and growth,” *Quarterly Journal of Economics*, 1995, pp. 681–712.
- Méon, Pierre-Guillaume and Khalid Sekkat**, “Does corruption grease or sand the wheels of growth?,” *Public Choice*, 2005, *122* (1-2), 69–97.
- Muthukrishna, Michael, Patrick Francois, Shayan Pourahmadi, and Joseph Henrich**, “Corrupting cooperation and how anti-corruption strategies may backfire,” *Nature Human Behaviour*, 2017.
- Nicklisch, Andreas, Kristoffel Grechenig, and Christian Thöni**, “Information-sensitive Leviathans,” *Journal of Public Economics*, 2016, *144*, 1–13.

**Olken, Benjamin A**, “Corruption perceptions vs. corruption reality,” *Journal of Public Economics*, 2009, *93* (7), 950–964.

**Veldhuizen, Roel Van**, “The influence of wages on public officials’ corruptibility: A laboratory investigation,” *Journal of Economic Psychology*, 2013, *39*, 341–356.

**Weisel, Ori and Shaul Shalvi**, “The collaborative roots of corruption,” *Proceedings of the National Academy of Sciences*, 2015, *112* (34), 10651–10656.

# Appendix

## Additional Tables and Figures

Table S1: Check of randomization of treatments to citizens

	BRIBE		NO BRIBE		Test	
	Mean	S.D.	Mean	S.D.	F stat	P-value
Male (= 1)	0.333	(0.474)	0.452	(0.501)	2.503	0.116
German mother tongue (= 1)	0.833	(0.375)	0.893	(0.311)	1.254	0.264
Economics (= 1)	0.452	(0.501)	0.488	(0.503)	0.213	0.645
Observations	84		84		168	

*Notes:* Variables include a dummy for gender, a dummy for German mother tongue and a dummy for economics major.

Table S2: Check of randomization of treatments to monitors

	BRIBE		NO BRIBE		Test	
	Mean	S.D.	Mean	S.D.	F stat	P-value
Male (= 1)	0.321	(0.476)	0.500	(0.509)	1.839	0.181
German mother tongue (= 1)	0.893	(0.315)	1.000	(0.000)	3.240	0.077
Economics (= 1)	0.321	(0.476)	0.429	(0.504)	0.669	0.417
Observations	28		28		56	

*Notes:* Variables include a dummy for gender, a dummy for German mother tongue and a dummy for economics major.

Table S3: Contribution in  $t$  ( $C_{i,t}$ ), Periods 1-20

	Periods 1-20	
	(1)	(2)
Treatment Bribe	-4.136*** (1.291)	-4.006*** (1.313)
Constant	11.848*** (0.952)	10.849*** (1.486)
Periods F.E.	Yes	Yes
Controls	No	Yes
$R^2$	0.092	0.096
# Clusters	56	56
Observations	3360	3360

*Notes:* OLS estimation. Standard errors (clustered at the group level) are displayed in parentheses. The dependent variable is the contribution level in period  $t$ . Controls include citizen-specific dummies for gender, German mother tongue and economics major. Levels of significance: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

Table S4: Change in contribution from  $t - 1$  to  $t$  ( $C_{i,t} - C_{i,t-1}$ ).

	Not punished at $t - 1$		Punished at $t - 1$	
	(1)	(2)	(3)	(4)
Treatment Bribe	-0.020 (0.183)	0.275 (0.186)	-1.042*** (0.372)	-0.848** (0.384)
Others' average contribution ( $\bar{C}_{-i,t-1}$ )		0.068*** (0.016)		0.040 (0.043)
Constant	-0.682*** (0.145)	-2.911*** (0.641)	2.190*** (0.305)	1.204 (1.046)
Periods F.E.	No	Yes	No	Yes
Controls	No	Yes	No	Yes
$R^2$	0.000	0.035	0.010	0.039
# Clusters	56	56	53	53
Observations	2238	2238	954	954

*Notes:* OLS estimation. Standard errors (clustered at the group level) are displayed in parentheses. The dependent variable is the change in contribution from period  $t - 1$  to period  $t$ . Controls include citizen-specific dummies for gender, German mother tongue and economics major. Levels of significance: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

Figure S1: Screenshot of the monitors' decision screen in the BRIBE treatment. Note that in the NO BRIBE treatment, the line "transfer to the monitor" did not appear.

Period 1 of 10 Remaining Time [sec]: 109

The decisions of the participants A of your group

Contributions to the project: ....

Transfer: ....

Your deduction points to participants A:

To assign deduction points, please enter a number from 0 to 10.

Your intermediary income: ....

Additional endowment: 15.0

Cost of your deduction points: ...

Your income this period: ...

Calculate

OK

HELP

Please enter for each participant A the amount of deduction points you want to assign. Click on « calculate » before clicking on « ok ».

Figure S2: Screenshot of the citizens' final screen in the BRIBE treatment. Note that in the NO BRIBE treatment, the line "transfer to the monitor" did not appear.

Period 1 of 10 Remaining Time [sec]: 51

You The other participants A of your group

Contribution to the project: ...

Transfer to participant B: ...

Deduction points received: ...

Your income this period: ...

OK

HELP

You can see the results of this period.

After you click on « ok » or if the time is up, the experiment will proceed



Figure S3: Distributions of transfers

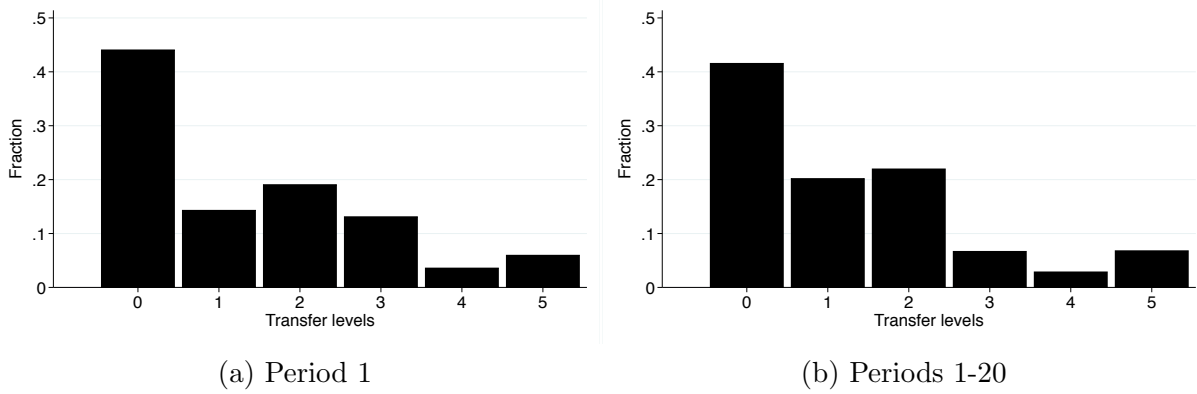


Figure S4: Robustness checks for the definition of initially high contributors.

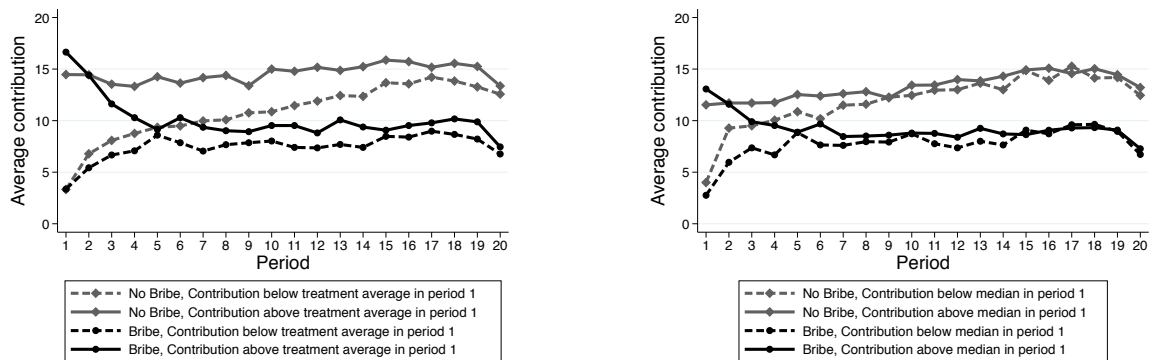
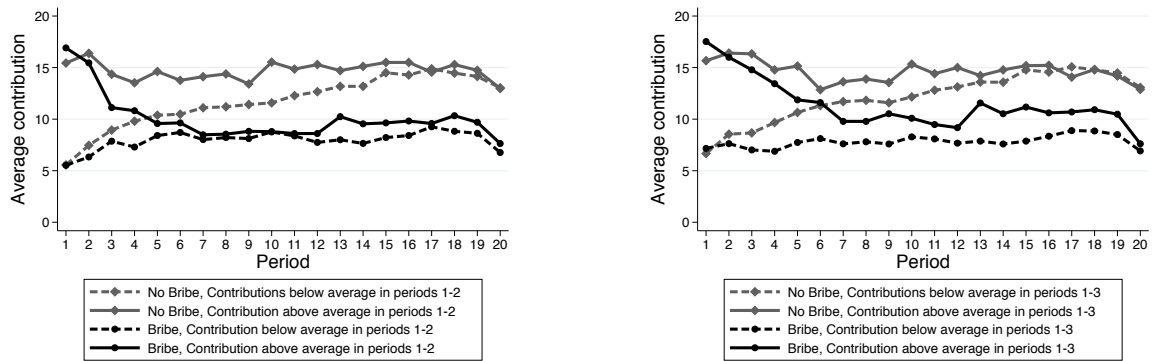


Figure S5: Overall efficiency (average payoff of the monitor and the 3 citizens)

