Modelling Domestic Corruption Deterrence Through Self-Reporting and Collaborations

Lucas Alves Chacha

October 16, 2020

Abstract

Past studies have shown that giving sanction reductions (leniencies) to self-reporters who collaborate with investigations is an effective way to prevent corruption. However, poorly designed policies may induce criminal activities. In this sense, this theoretical work models the crime of collusive corruption or bribery. It focuses on the differences between games with complete and incomplete information. Additionally, the model allows agents to self-report their crimes in exchange for some judicial benefit before or after being detected by the authorities. Like similar past researches, it shows that collusive corruption suffer from a coordination constraint, since agents cannot enforce their deals through legal contracts. This is known as the 'hold-up' problem. In this case, giving leniencies to self-reporters provide agents with a credible threat, which may overcome this problem of corruption. The paper shows that the 'hold-up' constraint for collusion is smaller under incomplete information. However, it also shows that uncertainties from agents are strong constraints to engage in collusive corruption. Lastly, this work shows that giving leniencies before or after being detected by the authorities induces crime and have no cumulative effects under complete information. Conversely, under uncertainty both policies can jointly deter corruption. Consequently, the level of information about players is crucial to determine the deterrent effect from anti-corruption policies.

1 Introduction

Using one party against the other to disrupt criminal agreements is an old prosecutorial practice. However, the design of the strategies varied throughout time. Notably, corruption or bribery requires at least one agent to pay a bribe and other receive it. Therefore, stimulating self-reporting can be particularly effective against this type of crime.

Since the seminal work from Becker (1968), economists formalized a way to analyse the decision to engage in crimes[1]. Notably Kaplow and Shavell (1994) first described how giving fine reductions to self-reporters could deter crimes[2] for sufficiently risk-averse individuals. However, their analysis was only one sided or static.

[1]More recently Polinsky and Shavell (2007) devised a general static model based on contributions of a vast literature. The model accommodates a series of effects such as different liability rules, different risk preferences, non-monetary sanctions, expected fines, judicial errors, deterrence effect, incapacitation effect, Principal-Agent relation, settlements, repeated offenders, corruption on judgements, social norms and self-reporting. Also Garoupa (1997) resumes the relevant literature.

[2]Specific literature in corruption starts earlier. Rose-Ackerman (1975) shows that incentives to corruption rise on different market structures. Another strand of literature derived from Tirole (1986), which explored corruption as an output from agency problems in a three-tier hierarchy. Later on, Polinsky and Shavell

The dynamic features from the interaction between agents were initially added in the work of Motta and Polo (2003)[3], which was focused on anti-trust offences[4], and later Spagnolo (2005) acknowledged that the methodology could be extended to other types of criminal organizations. Later on, Buccirossi and Spagnolo (2006) framed a similar methodology as a corruption game[5]. All of them found situations in which leniency for self-reporters may deter or increase criminal activities, depending on the policy design.

Even though the models of Motta and Polo (2003) and Chen and Rey (2013) allow the agents to self-report not only before an eventual detection by the authorities, but also after. The literature that followed their strand mostly do not allow for it. Here, the model devised allows agents to report their crimes before or after being detected[6]. Hence, this model resembles some features from plea bargaining studies. The literature on plea bargaining, starting with Landes (1971), analyses specifically the case in which defendants would collaborate with justice and report their criminal actions after an accusation. Later, Grossman and Katz (1983) devised a dynamic model of plea bargaining pointing out to screening effects that would help prosecutors in their choices. Kobayashi (1992) developed a model accounting for multi-defendant plea bargain which is closer to this study's case of corruption analysis[7]. Conversely, the studies on plea bargaining are interested in knowing the optimal solution for prosecutors or judges. Here, it is assumed that prosecution and judiciary are making exogenous (and optimal) decisions.

Note that parties can have distinct decision power over a corrupt deal. If the relation is asymmetric, one party may be subjected to pay bribes as result from extortion. This is the typical case of civil servants requiring some bribe to emit a due license. Works from Dufwenberg and Spagnolo (2014) and Lambsdorff and Nell (2007), show that leniency to self-reporters or asymmetric punishments may be effective against these cases. This happens because one party would always prefer not to engage in corruption. However, if power is balanced, collusive corruption arises by the benefit of both parties. Consequently, this type of corruption may imply in bigger and more complex crimes, such as procurement frauds and overpricing. In this case both agents have incentive to hide their wrongdoings. Buccirossi and Spagnolo (2006) approach this type of bribe and found that second movers can act

(2001) deal with the problem of corruption for framing and extorting individuals. These works and many others were written on the same line (Burguety et al., 2016). However, here the focus is on the self-reporting dynamics and how it can deter corruption.

[3]One important difference between the static and dynamic approaches so far is that the first were set on individual level, or else, the decisions were taken by natural persons. The later dynamic models were set on the corporate level, with the exception of Spagnolo (2005) extents on criminal organizations.

[4]Notable theoretical contributions in the leniency for cartels are Harrington (2008), Motchenkova (2004), Aubert et al. (2006), Motchenkova and van der Laan (2011), Buccirossi and Spagnolo (2000), Chen and Rey (2013) and Chang and Harrington (2015). Interesting experimental findings are Bigoni et al. (2012) and Bigoni et al. (2015). Also, empirical approaches are explored by Brenner (2009) and Miller (2009).

[5]Other remarkable literature on theoretical games of corruption are Dufwenberg and Spagnolo (2014) and Basu and Cordella (2016), and experimentally Engel et al. (2016) and Abbink and Wu (2017).

[6]At this point the reader may wonder why is it relevant to analyse both policies together. First, it is more realistic, being better suited to most jurisdictions, as agents face both choices to report their actions and different benefits in each case. It is indeed similar to Brazilian recent judicial reform that is going to be used as an empirical example in further developments of this work. Additionally, if we consider that corruption schemes are strategic interactions, the intelligent use of both policies together can increase jointly the deterrence or disruption of the criminal schemes

[7]Previous literature goes further on the analysis of optimum decision both for defendants and justice taking into account the trade-off between certain of conviction and wrongful pleas from risk-averse defendants (Grossman and Katz, 1983; Kobayashi, 1992; Givati, 2014; Mungan and Klick, 2016; Berg and Kim, 2018). However, since the objective here is to analyse the defendant's decision making, the prosecutor's choice is assumed as exogenous.

opportunistically, not delivering a promised favour after receiving a bribe. It is known as the 'hold-up' problem. However, leniencies can provide a credible threat to first movers, overcoming the 'hold-up' problem of corruption. This work focuses on the last type of corruption, and try to assess the role of uncertainties in this setting[8][9].

Importantly, there are two channels in which well-designed anti-corruption policies can be effective. It can work ex-ante, deterring the crime or impeding agents from performing it. Or it can be effective after the crime is committed, as a good framework to prosecute or to disrupt an eventual ongoing crime. This paper focuses on the deterrent aspect of the policies, the disrupting effect is explored in further opportunity[10].

Lastly, in most of the past works on leniency or corruption, information was complete. Or else, agents had perfect knowledge about other agents. The uncertainties were mostly about the enforcement variables. The exception comes from Spagnolo (2005), which make a riskiness analysis of the agent view of the game. Here, the innovation relies on addressing some uncertainty to the type of player that the agents are playing against.

In this sense, this paper tries to answer how deterrent can leniency policies against collusive corruption be. What are the characteristics of an effective leniency policy against corruption? Does the possibility of self-reporting before and after have a joint effect in deterring collusive corruption? How does the fear of being reported affects the decision to engage in collusive corruption?

At the end, beyond confirming some past findings from the specific literature, it is clear that the level of information that agents have plays an important role on the effectiveness anti-corruption policies. If agents do not have complete information about other players (do not know what type of players they are playing against), the constraint of the leniency policies is bigger. Moreover, leniencies for self-reporters before and after detection may independently deter collusive corruption[11].

In the next section the basic model is devised. In sections 3 and 4, the games are played under complete and incomplete information respectively. Moreover, in each session all the different timing protocols are explored. At the end, section 5 concludes the study resuming the relevant results from the model, also further studies and extensions are proposed.

2 Domestic Corruption Model

This section contains a theoretical analysis of domestic collusive corruption. Furthermore, it focuses on the deterrent effects from the leniency policies. Notably, a lot has been written on leniency for cartels (Marvão and Spagnolo, 2014) and a lot less for corruption criminals (Burguety et al., 2016). In this case it is necessary to frame the model as a corruption game in order to verify if conclusions from past works on leniency for cartels apply to the case of corruption. Importantly, this study considers only one dimension of the corruption activities, the bribery related crimes between players with symmetric bargain power.

[8] Rose-Ackerman (2010) uses four different categories instead of the two used here. For her benefits from bribes can be illegal (such as the case here) or legal and scarce by nature. Also can be only scarce because corruption allocates it inefficiently and, lastly, not scarce but corruption allocates it inefficiently.

[9] Needless to say, private-to-private corruption is out of the reach of the present work.

[10] The deterrent effect of self-reporting is somehow simpler to show, since agents make their forecasts of the future and then make a decision. The parameters in which the decision was made are fixed, so there is a lot less degrees of freedom in the model.

[11] If the value of corruption is decided by the parties, self-reporting policies can increase the values of bribes and advantages from corruption. However, it is doubtful that players would choose values of bribe and advantages less than maximum. If values are tied to the probability of detection and conviction, this result may not appear at all.

The model begins with the agent's choice. In this sense, considering that the relevant objective of the policy (policymaker) is to maximize society's welfare. Also, supposing that the enforcement expenditures are exogenous[12]. Then, the welfare maximization is given by restricting the set of possible enforcement combinations in which corruption would occur. Note that it is implicitly assumed that there is no 'grease money hypothesis'(Kaufmann and Wei, 1999), i.e. there are no gains for society from corruption. Therefore, society's welfare is monotonically decreasing in the number of corruption crimes[13].

Off course, costs of deterring corruption may increase as the number of crimes decreases. Therefore, there may be an optimum level of enforcement that does not eradicate corruption (non-trivial solution), as many similar works have already found (Motta and Polo, 2003). Although, this analysis points out to the trade-offs of the anti-corruption policies, they do not try to measure the welfare gained from the absence of corruption. The reasons for that are several, it goes from the difficulty to point out the negative externalities from corruption (Rose-Ackerman, 2006) to the relative subjectiveness of the measure of welfare. Nonetheless, the analysis is insightful in pointing out the direction of the deterrent and the inducing effects of the policies described here.

The study's overall methodological approach to achieve this goal is to devise an economic model of public enforcement against corruption. The model accounts for the possibility of defendants or potential criminals to report their acts to the authorities in exchange for some judicial benefit. If the model is consistent with reality, the theoretical predictions extracted here should be confirmed by an eventual empirical test.

The model consists of risk-neutral agents which can either be a firm [14] or an individual. Here, corruption is defined as bribes paid in a *quid pro quo* exchange for some advantage. This deal is illegal and if detected an investigation/accusation is brought against involved parties. If the parties are convicted, they pay a monetary sanction[15]. The bribery can be performed in one-shot (one time only) or in repeated interactions.

The game arises when two agents[16], the government bureaucrat and private entrepreneur[17], decide whether to play or not a strategy of corruption (cooperation), constrained by probabilities of being detected and later convicted.

Given the above conditions, let:

[12]This model has an endogenous revenue from fines. To consider it as welfare neutral it is needed to assume that fines are mere transfers from corrupt agents back to society. In other words, the welfare lost from paying a fine from one agent is diluted in a gain for the society, i.e. no dead-weight loss.

[13]Polinsky and Shavell (2001) found this result applying traditional methodology from Becker (1968). More specifically, by weighting the costs and benefits from crimes to the welfare from society. For simplicity this is not done here.

[14]Note that, for corporate crime deterrence model, there is an intrinsic assumption that there is no principal-agent conflict which sometimes is not always explicit in models (Søreide and Rose-Ackerman, 2018), e.g. no conflict between the bureaucrats performing the corruption activity and the politicians or political parties that back them up. Also, there is no conflict between the firm's shareholders or owners and their employees that pay the bribe. Aubert et al. (2006) discuss the principal-agent and the extents of individual and corporate decisions when addressing the whistle-blowing issues. Also Garoupa (2007) deals with agency problems when agents can rely on plea bargains. In his model, it is a superior equilibrium to pay low wages for agents and give little information that can lead to the principal arrest. Lastly, Tirole (1986) proposes a three-tier agency problem, in which agents and supervisors collude against a principal. These are not the cases modelled here.

[15]Non-monetary sanctions are discussed in the extension of the model.

[16]Since corruption requires at least two players, a payer and a receiver, this assumption will always hold true. Notably there can be more players, the extensions to account more players is discussed further.

[17]The model can account for more agents, such as politicians, government agents and private companies. Also, briberies could happen between private parties, although the case of bribery between public and private parties is more emblematic and easier to picture.

B = Bureaucrat;
E = Entrepreneur;
π = Advantage from corruption;
b = Bribe; and
c = Cost of the bureaucrat for generating π,

where $\pi > 0$ is the gain from the entrepreneur from corruption and b is the gain from the bureaucrat, such that $0 < b < \pi$. Note that, the existence of $\pi > 0$ rules out crimes of extortion from the bureaucrat, since in that case the entrepreneur would not face any positive benefit from corruption[18].

Additionally, agents are assumed to be under a fault-based liability regime[19]. Furthermore, judiciary and prosecution have jurisdiction over all agents (domestic bribery).

Let the enforcement variables be:

α = probability of detection[20]; and
β = probability of conviction.

In the traditional Beckerian approaches of crime deterrence, the probability of detection and conviction are merged. Here, in line with the study of Motta and Polo (2003) and Chen and Rey (2013) the probabilities are split. This feature allows for the specific propose of analysing self-reporting after detection.

Finally, for $i = B, E$, let:

S_i = Sanction; and
F = Fine.
For the moment, it is considered that S_i is given by the gains of each player with addition of F which is the same for the payer and the receiver. Consequently, $S_E = \pi + F$ and $S_B = b + F$[21]. It is assumed that there are only monetary sanctions, or they can be represented in a monetary fashion at least[22]. Lastly, it is assumed that fines are costless to implement[23].

In order to derive the agent's decision rule, and for the sake of simplicity let,

[18] In this sense, the best way to tackle extortions would be by following the hypothesis of Basu (2011) better explored in Dufwenberg and Spagnolo (2014) and Basu and Cordella (2016).

[19] Since the agents can be either firms or individuals, there can be differences between each type of agent in different countries or for each distinct purpose. Previous literature discussed the optimal liability regimes. For instance, for individuals, Polinsky and Shavell (2007) conclude that, for risk-averse defendants, fault-based liability is preferable than strict-liability, however strict-liability is easier to implement. Arlen and Kraakman (1997) discuss corporate liability. The authors identify that, based on optimum activities of monitoring and compliance from firms along with preventive measures to avoid criminal activity, duty-based liability together with residual strict liability can be optimum. In fact, distinct liability can induce different incentives on self-enforcements, in this case, there can be a compound corporate liability that maximizes social welfare, as Oded (2012) would suggest.

[20] Innocent agents will never be subjected to detection so they can never be wrongfully convicted, i.e. there is no Type I judicial errors.

[21] Shleifer and Vishny (1993) found that efficient fines should be tied to agents benefits from corruption. In this sense, here, since fines F are constant and agents are required do give back all gains from corruption, they can be considered linearly correlated to the size of the corruption gains. However, $\frac{F}{b-c}$ or $\frac{F}{\pi-b}$ are strictly decreasing in the gains from corruption. This may not be optimal, but it is certainly realistic for most jurisdictions.

[22] This simplification may be more real for corporate agents than it is for natural persons. More specifically, imprisonment would be expected for individuals but not for firms. Nonetheless, there are also other types of non-monetary sanctions that can be applied to firms. In expanded model, this specific issue is revisited.

[23] Agents are not liable for damages by third parties or for crimes in other jurisdictions regarding the reported offence.

p_k = probability that k happens for $k = \pi, F, b$,
$p_F \equiv \alpha\beta$ and,
$p_b = p_\pi \equiv (1 - \alpha\beta)$.

Under risk neutrality, the agents' utilities U_i are given by the the differences between the costs and the expected benefits for each agent:

$$U_E \equiv -b + (1 - \alpha\beta)\pi - \alpha\beta F = -b + p_\pi - \pi p_F F,$$

and

$$U_B \equiv -c + (1 - \alpha\beta)b - \alpha\beta F = -c + p_b b - p_F F.$$

Let us define some profitable bribery as being the cost paid by agents which make their utilities at least greater than zero:

$$b < p_\pi \pi - p_F F, \tag{1}$$

and

$$c < p_b b - p_F F, \tag{2}$$

Consequently, the enforcement level $\alpha\beta$ needs to be lower than the relation between premium from corruption and the sanctions,

$$p_{F_E} = \alpha\beta < \frac{-b + \pi}{F + \pi}, \tag{3}$$

and,

$$p_{F_B} = \alpha\beta < \frac{-c + b}{F + b}. \tag{4}$$

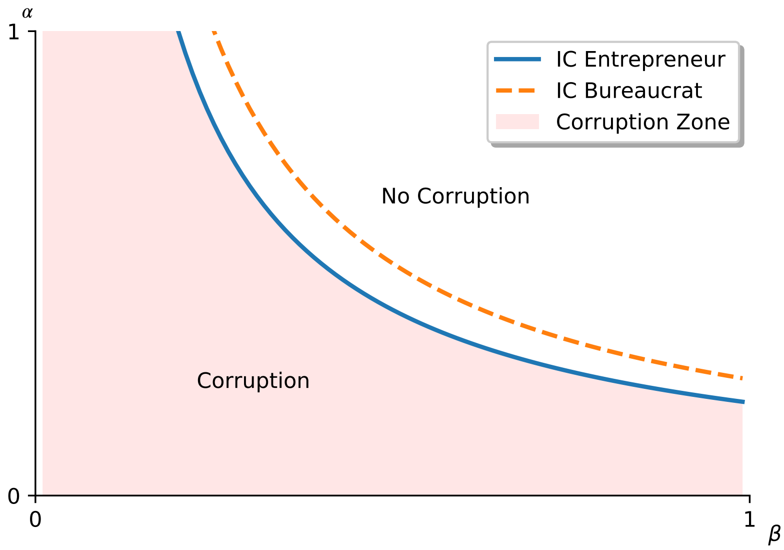
Importantly, given fixed levels of $\alpha\beta$, higher fines will always lead to higher constraints for bribe acceptance in both cases. Thus, if F is costless to implement, policymakers should set it as high as feasible (generally up to the agents' total wealth) lowering the enforcement levels $\alpha\beta$, which are usually costly. This conclusion is in line with traditional Beckerian approaches [24].

Analogously, for fixed feasible conditions of c, b, π and F[25] there will be some combination of α and β which makes bribes profitable for bureaucrats and entrepreneurs. The Figure 1 shows the Indifference Curves between prosecutorial or judicial efficiency against the agents .

[24] Assuming that enforcement variables α and β are functions of some exogenous government expenditure $f(e)$, such that, for both cases, $f'(e) > 0$ and $f''(e) < 0$. It is possible to note from (3) and (4) that it is always possible to exchange (F) for lower levels of $\alpha\beta$ and maintain the decision rule unchanged (deterrence level). This conclusion is in line with what Becker (1968) suggests.

[25] Since the minimum bribe is given by the bureaucrat and the maximum is given by the entrepreneur, it is straightforward to see that feasible bribes would lie in $b \in (\frac{c+\alpha\beta F}{(1-\alpha\beta)}, (1 - \alpha\beta)\pi - \alpha\beta F)$. Additionally, the minimum advantage from corruption necessary for the bribery is $\pi > \frac{c+\alpha\beta F-\alpha\beta(1-\alpha\beta)F}{(1-\alpha\beta)^2}$.

Figure 1: Indifference Curves for Profitable Bribes



Given the above conditions, bribes are profitable if $\alpha < \frac{\pi - b}{\beta(\pi + F)}$ or $\beta < \frac{\pi - b}{\alpha(\pi + F)}$for the entrepreneur and $\alpha < \frac{b}{\beta(b+F)}$ and $\beta < \frac{b}{\alpha(b+F)}$ for the bureaucrat. Therefore, an increase in F moves both curves to the left, increasing the deterrent effect. Meanwhile, a decrease in $(b - c)$ moves the bureaucrat indifference curve towards a more deterrent set, whereas for the entrepreneur's indifference curve, it moves left if $(\pi - b)$ decreases.

If agents can mutually enforce some the bribery agreement, then conditions (1) and (2) suffice for parties to collude. However, since the deal is illegal, agents cannot rely on traditional contracts to enforce their commitments. This is known in contract theory as a 'hold-up' problem. Consequentially, parties would need to rely on other means for it, such as threats of violence, trust and other strategies. In this sense, information about parties and timing protocols are important to determine the conditions for agreements to hold in an environment without legally enforceable contracts.

Lastly, note that p_k depends on the level of uncertainty (private information) that agents have over the k variables. Therefore conditions (3) and (4) may change under uncertainty. Additionally, since players interact sequentially with each other, the conditions will also depend on the dynamics of the game (timing). These relations become more explicit if one allows agents to self-report. Therefore, profitable bribes ((1) and (2)) are necessary but not sufficient for agents to start colluding. The different cases regarding the level of information of agents and different timings are discussed below.

Timing and Information

In the next sections different stylized cases of the model are set. The main difference between them is the level of information of the agents and the timing protocol. This is precisely the intended contribution from this paper, to isolate and highlight effect of different timing rules under distinct levels of information about players.

Regarding the timing protocol, the judicial frameworks that allow criminals to self-report in exchange for fine reductions are numerous, such as leniency agreements, whistle-blowing and plea bargaining. They differ between each other in their procedures, intentions and, consequently, in their outcomes. For simplicity, in this model there are only two types of self-reporting: before and after detection.

There are four different self-reporting regimes. In other words, there are four distinct timings for agents to self-report in exchange for some benefit or leniency in their sanctions: 1- No fine reductions for self-reporters (benchmark); 2- Fine reductions for agents who self-report only before detections from the authorities; 3- Fine reductions for agents who self-report after detected; and 4- Distinct rules for agents who self-report before or after detected.

For convenience, the representative game in which the entrepreneur plays first is the standard approach. However, the same results can be mirrored if the bureaucrat plays first. The timing for the benchmark, without fine reductions for self-reporters, is the following:

- $t_0 =$ Policymakers set the rules for fines F;
- $t_1 =$Nature sets the distribution of the advantages π that the entrepreneur can explore from a bureaucrat. Then, the entrepreneur decides whether to offer a bribe b;
- $t_2 =$The bureaucrat decides to accept or not. If the bribe is accepted, the bureaucrat decides to perform or not the corruption act;
- $t_3 =$Prosecution authority investigates (detects) with probability α; and
- $t_4 =$If detected by the prosecution, judicial authority convicts with probability β.

Agents can decide to keep 'hard evidence' from corruption (cooperation). This evidence can never be manipulated and if it is shown to the authorities, it would convict the other party with certainty[26][27].

If the policymakers decide to include a leniency policies in t_0 to self-reporters. Then, there are three more timing possibilities: agents can only self-report if not detected (before t_3); agents can only self-report after detected (before t_4); and agents can either self-report before (before t_3) or after detection (before t_4) and there are specific rules for fine reductions[28].

Since this model explores the deterrent effect from the anti-corruption policy. All the variables are chosen or set before the decision whether or not to bribe (accept a bribe) (t_1 and t_2), and remain constant after. This happens because the agents must decide their actions ex-ante and have no reason to believe that variables would change over time[29].

Past literature points out to counter-productive dynamics from giving fine reduction to self-reporters (breaking a sequential prisoner's dilemma) (Buccirossi and Spagnolo, 2000; Spagnolo, 2005; Buccirossi and Spagnolo, 2006; Dufwenberg and Spagnolo, 2014). However, there are positive deterrent effects from the fear of being reported (Spagnolo, 2005;

[26]This assumption somewhat similar to the ones in Aubert et al. (2006) and Buccirossi and Spagnolo (2006). However, in the last, agents always create the evidence and if detected they are convicted for sure because for that. Note that this is not the case here, since agents will still have the opportunity to go to trial, they can still try to hide the evidence. Therefore, since agents will only use the evidence if they want to self-report, the assumption here is less strong than it is in Buccirossi and Spagnolo (2006).

[27] The role of information on self-reporting is more elaborated in Garoupa (2007). However, his concept of criminal organization may not apply to this situation of corruption between public and private sectors. His modelling cares more about agency problems inside a hierarchical criminal organization.

[28]Note that, in practice, agents who can self-report only after detection, can do so before $t = 3$ as well, since they can turn their selves in to the authorities. The difference between only self-reporting after, and self-reporting both before and after is that in the last case, it is allowed for different leniency rules for agents who self-report before or after detection.

[29]This is consistent with rationality. Also, this feature makes the deterrent effect somewhat easier to model than the disruptive effects from self-reporting. In order to model a run for court effect or a desistance effect, one would need to set the variables free to changes over time. This was done by Motchenkova (2004) and Harrington (2008) and is a natural improvement of this analysis, or at least, an intended further analysis.

Harrington, 2008; Chang and Harrington, 2015). Moreover, these findings are confirmed by experimental evidence (Bigoni et al., 2012, 2015; Engel et al., 2016; Abbink and Wu, 2017). Therefore, in order to isolate these two effects, different assumptions about completeness of information are set. In particular, the complete information model extracts the dynamics from the game. Meanwhile, the incomplete information settings analyse the difference in expected outcomes resulting from the addition of a noisy uncertainty to the system.

In the following subsections all different specifications of the games are discussed.

3 Complete Information

In this section, agents have complete information, meaning that, they know their and other's benefits from corruption ($\pi - b$ for the entrepreneur and $b - c$ for the bureaucrat). Moreover, agents know their fines F and probability of being caught α and if so convicted β. Since agents are risk neutral and act rationally, the agents' utilities U_i are given explicitly by the game pay-offs which they are trying to maximize.

3.1 One-Shot Complete Information

In this setting, agents play a sequential game in which the first mover offers a bribe then the second mover delivers the favour. Importantly, the deal is never repeated between the players, neither there are any reputational gains or any other ways for parties to enforce their claim. This particular framing resembles typical routine interactions between entrepreneurs wishing to explore some procurement contract and bureaucrats or civil servants entitled to execute them. In this case, entrepreneurs pay a bribe then bureaucrats place some adverse condition (overpricing, defrauded quantities) in the contract that generates an uncompetitive value. Other examples are queuing hold-ups, in which someone can bribe a bureaucrat to put a process further in a waiting list.

Note that there must be symmetry of bargaining power[30], a sequential order of paying, some value generated to the bribe payer and some non-negative cost to a bureaucrat. This leaves out some specific types of corruption, such as extortion and avoiding framing individuals. Since in extortions there is no value added, bribe payer is just avoiding a cost. Such as paying a civil servant emit a licence that one would be entitled to. In the case of framing, if the bribe payer refuses to pay, they are going to be immediately detected by authority. For an example, if people try to bribe a police officer to avoid a speed limit ticket or bribe a judge to acquit a guilty defendant.

Needless to say, in this case, agents are very sophisticated and they know all the variables involved for them and for the other parties. The dynamics of different cases are discussed below.

3.1.1 Benchmark Without Fine Reductions for Self-Reporters

In this first case, the one-shot (one-time) game implies that agents have perfect information, meaning that they know where they are in the sequential game. In other words, players can observe what the others have done. This is a natural real assumption of paying a bribe, agents know whether they paid or received a bribe.

[30] Put differently, in this study, the relevant corruption cases are the ones involving equally interested parties. Or else, the players have some symmetry of bargaining power. Therefore, there is collusion. If agents were in an asymmetric relation, the bribes would be paid by extortion.

The game consists of an entrepreneur making a 'take-it-or-leave-it offer' of a bribe to a bureaucrat who decides whether to reject, accept and perform or accept and not perform the corruption act. Figure 2 shows the extensive form of the game, in which the nodes E and B are the entrepreneur and the bureaucrat. The, strategies nb, b, r, p and np represent respectively, 'no bribe', 'bribe', 'reject the bribe', 'perform the act' and 'not perform the act'. Finally, the upper pay-offs are from the entrepreneur and the lower from the bureaucrat.

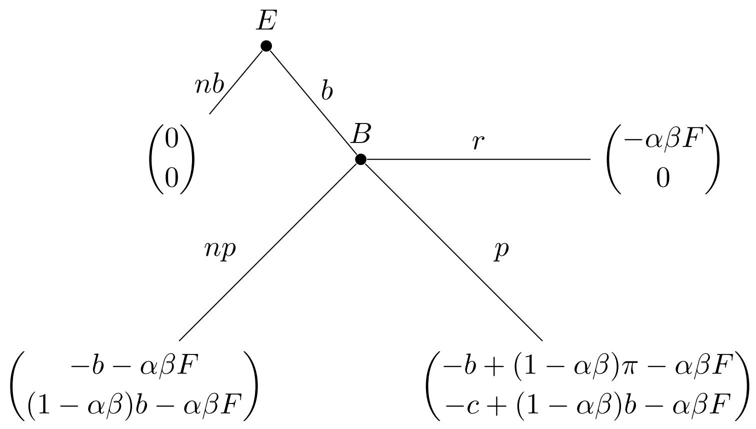


Figure 2: One-Shot Sequential Game Tree

Following the timing structure, after the bureaucrat chooses either to perform or not the act, the prosecutors play and after them the judges play. Note that these moves are exogenous to this model. Thereafter, the expected values from each node are condensed in the pay-off following B's decision. Thus, in this reduced form only the nodes in which the relevant players E and B play are displayed.

It is also important to point out that the expected fine $\alpha\beta F$ is the same in cases of the bureaucrat performing or not the corruption, both for the public official and for the entrepreneur. This assumption is in line with many national laws[31], since the mere attempt or promise to corrupt is considered a crime. For simplicity, it is considered that the rejection does not constitute a crime but the attempt to pay does.

Solving the game backwards. For any non-negative cost of corruption $(c \geq 0)$[32], if expected return from bribes is not positive $(1-\alpha\beta)b - \alpha\beta F \leq 0$, or else the bribe is too low $b \leq \frac{\alpha\beta F}{(1-\alpha\beta)}$, than the bureaucrat chooses rejection. Whereas, if the expected pay-off from the bribe is positive, the bureaucrat would chose to 'not perform' the corruption since $(1-\alpha\beta)b - \alpha\beta F \geq -c + (1-\alpha\beta)b - \alpha\beta F$. Therefore, no benefit π for the entrepreneur would be generated. Consequently, the entrepreneur is not going to offer the bribe. Hence, no corruption (not colluding) is a unique Subgame Perfect Nash Equilibrium (SPNE)[33]. Furthermore, by symmetry, since the bribe will always be costly to the entrepreneur, there is no corruption if the bureaucrat plays first. Once again, in both cases, second mover would prefer to take the benefit and not perform the activity, then the first mover is not offering the bribe. This is the 'hold-up' problem, from contracting theory.

[31]For example, Brazilian Law No. 2.848/40 Art. 317 and 333.

[32]Dufwenberg and Spagnolo (2014) allow for a $c < 0$. This is possible in their framing since they deal with extortive or harassment bribes. In this case, if the bureaucrat denies a licence or other service, she can be caught-in-act and fired. Of course, this condition does not apply for this current model.

[33]Considering that in the special case in which $c = 0$, then it is assumed that agents weakly prefer not to act, since acting would represent at least a change in current status quo.

One can argue that this design is too restrict, and there can be a way in which agents can intelligently and tacitly collude[34]. However, this is a particular case of a 'trust game'. Consequently, if players approach it rationally, then it will always lead to non-cooperation. This may appear unrealistic, since occasional corruption is often observed in the society. Perhaps, the occurrence of such behaviour may be due to the use of other enforceable agreements (threats of violence, shadow contracts, trust, cheap talk, etc.). Or even due to the presence of bounded rationality, and effects such as reciprocity (positive and negative) or inequity aversion[35].

The important message is that the dynamics of corruption does not favour collusion, since it generates a 'hold-up' problem for parties. This is true to any other illegal sequential transactions that cannot be enforced through contracts or legal systems. They resemble trust games and rational players would not collude if they do not have any enforceable mechanism to assure their agreements.

Proposition 1. *There is a crime deterrent effect from the dynamics of corruption if bribes are paid in a sequential one-shot game without enforceable agreements.*

This outcome is in line with previous literature. Buccirossi and Spagnolo (2006), in their work with a similar game structure found that corruption would not be implemented because parties' dominant strategy would lead to mutual defection. Nonetheless, Lambsdorff and Nell (2007) found that, if punishments are asymmetric (either by being different between agents or between the action of reciprocating), there can be a punishment structure that allow players to collude. However, asymmetric punishments are more similar to the cases ahead (fine reductions to self-reporters) than they are to the benchmark.

3.1.2 Self-Reporting Before Detection

In this setting, the agents receive fine reductions if they self-report only before an eventual detection from the prosecutorial authority. This kind of setting approaches the ones used on 'leniency policy' studies, most of them aiming to analyse anti-trust policies (Marvão and Spagnolo, 2014). However, extensions from these approaches acknowledge the possibility of analysing other crimes such as corruption (Buccirossi and Spagnolo, 2006; Dufwenberg and Spagnolo, 2014) and other criminal organizations (Spagnolo, 2005).

This feature implies that the information in the sub-game after the bureaucrat's decision to perform or not the favour is imperfect. In other words, the agents cannot know what the other players have done, or else, they do not know with certainty in which node they are in the game. For an example, if the entrepreneur decides to self-report, she can only know if the bureaucrat did the same after taking the decision to the authority. On the other hand, if an agent does not self-report, she can only know if she was reported when the authority pronounces about it, or when the game ends (or round, if repeated)[36].

In this section, the model allows agents to self-report after committing a corruption crime but before any detection from competent authority. The self-reporter pays a reduced fine

[34]The non-cooperation equilibrium persists under different payment strategies, such as paying a part of the bribe before the act and other after (Buccirossi and Spagnolo, 2006). Nonetheless, fractioning the payment would approach the design to a repeated game, which is going to be explored ahead.

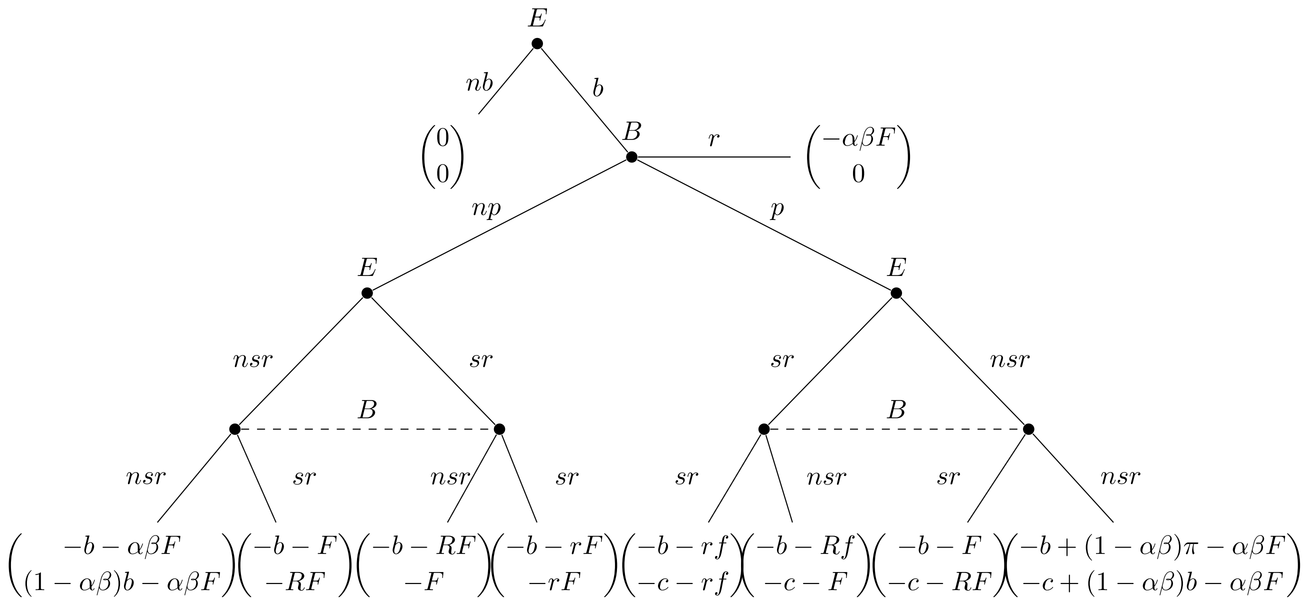
[35]Engel et al. (2016) explore the effect of asymmetric punishments on collusive bribery. Their benchmark is very similar to this one, and they observed that corruption (offers and acceptance) occurs frequently. The authors discuss several behavioural explanations for their results. However, they also show that they're benchmark is more deterrent than their control with asymmetric punishments.

[36]Although the absence of a clear game ending may be confusing. It is possible to assume that, for a one-shot corruption, the game finishes at the prescription date.

of RF, with $R \in (-\infty, 1]$. Lastly, if both players self-report at the same time they pay a reduced fine rF. Note that, under the usual rule that leniencies are only given to the first to self-report, then $E[r] = \frac{1+R}{n}$, i.e. the real fine is still $r = R$ for the first to come up and $r = 1$ for all the $(n-1)$ others[37].

The Figure 3 shows the extensive form of the game for the players, where '*nsr*' and '*sr*' stand for 'not self-report' and 'self-report' respectively:

Figure 3: Corruption Game With Self-Reporting Before Detection



Note that the crime can only be performed if bribery is profitable, i.e. if conditions in (1) and (2) are met. First, it is interesting to check in which ex-ante conditions would agents commit the crime and self-report. They would only do so if the premium for self-reporting is greater than the expected pay-off of corruption, or else $-RF > -b+(1-\alpha\beta)\pi-\alpha\beta F$ for the entrepreneur and $-RF > -c+(1-\alpha\beta)b-\alpha\beta F$ for the bureaucrat. This condition is only true if $-RF > 0$, this implies that $R < 0$, which means that agents would need to earn some bonus from self-reporting, otherwise they would certainly prefer to stick with the expected returns from corruption. More specifically, if bonuses are sufficiently big ($R < -\frac{-b+(1-\alpha\beta)\pi-\alpha\beta F}{F}$ for the entrepreneur or $R < -\frac{-c+(1-\alpha\beta)b-\alpha\beta F}{F}$ for the bureaucrat), agents would enter the game only to exploit it. In summary, it is possible to conclude that self-reporting is non-exploitable if:

$$R < max\left[-\frac{-b+(1-\alpha\beta)\pi-\alpha\beta F}{F}; -\frac{-c+(1-\alpha\beta)b-\alpha\beta F}{F}\right],$$

in other words,

$$-RF > min\left[-b + p_\pi \pi - p_{F_E}F; -c + p_b b - p_{F_B}F\right] \tag{5}$$

[37]Note that, since decisions are taken ex-ante, there is no impact from the 'first come first serve' rule. This happens because agents are only interested in the expected value from the fine reduction, not from the actual value. Other works in leniency for cartels that allow for ex-post decisions for self-reporting, such as Motchenkova (2004), Harrington (2008) and Chen and Rey (2013), found that leniency only to the first-comers is effective to disrupt collusions (run-for-court effect).

One player would pay the fine and the other would earn the bonus. Consequentially, the other party is never going to accept any bribe offer. Thereby, deterring eventual corruption[38]. However, if $-\frac{1+R}{n}F = -rF > (-b + (1-\alpha\beta)\pi - \alpha\beta F)$ for the entrepreneur and $-\frac{1+R}{n}F = -rF > (-c + (1-\alpha\beta)b - \alpha\beta F)$ for the bureaucrat, then both players would enter in corruption and both would self-report. Such high enough bonus, as any other bonus are not reasonable to assume. From now on, this situation where agents accept a bribe only to self-report (5) are referred as exploitable fine reductions.

Solving the game backwards[39], it is possible to see that if the bureaucrat does not perform the corruption, then the entrepreneur would self-report if $-\alpha\beta F < max[-RF, -rF]$, or else if $min[r, R] > \alpha\beta$. Furthermore, under the rule that leniency is divided between self-reporters or only given to the first to self-report, then $r > R$. Consequently, if RF is smaller than the expected fine $\alpha\beta F$, self-reporting here acts like a credible threat to the bureaucrat inducing her to perform the illegal act. From now on, this condition that generates credible threats is going to be referred as feasible fine reductions.

Given the above conditions, if leniency is either moderated[40] up to full amnesty ($R \in [0, \alpha\beta]$) like in most jurisdictions, self-reporting can only happen in pure strategies if α or β vary. Since, in the deterrence analysis agents assess the probabilities before playing the game, then they have no incentive to change their strategy, leading to corruption being performed.

In summary, allowing agents to self-report before being detected by authority's in exchange for benefits can induce corruption. This output is in line with Buccirossi and Spagnolo (2006), in which the addition of a self-reporting mechanism breaks the hold-up problem by providing a credible threat that can make collusion possible. On the other hand, if prosecutors allow bounties for collaborators $R < 0$ then, there can be situations in which collusion would be profitable, but enforcements deter the criminal act. Because agents would want to accept (offer) a bribe just to self-report. Therefore, in this case no one would offer (propose) a bribe. This result is in line with Spagnolo (2005) and Aubert et al. (2006).

3.1.3 Self-Reporting After Detection

In this setting, agents can earn judicial benefits if they self-report after being detected by the authorities. This type of self-reporting has been investigated by a particular strand of literature that mostly deals with plea bargaining such as the works of Grossman and Katz (1983), Kobayashi (1992), Franzoni (1999) and Mungan and Klick (2016). Literature about the subject traditionally focuses on the screening effects of plea bargaining. Prosecutors use them to decrease type I judicial errors, i.e. to prevent the conviction of innocent defendants. However, here the prosecutor choice is exogenous and agents cannot be detected if they are innocent. In this sense, the objective here is to assess the conditions in which self-reporting would be preferable and how they can contribute with the ex-ante deterrence of crimes.

Like in the past example, the game has imperfect information. However, now uncertainty lies on the node after the detection of the authorities. Due to the fact that players cannot know the other's action until the end of the round.

[38]Spagnolo (2005) using endogenous enforcement expenditures, finds that giving all the profits from corruption to the other party as bonus constitutes a first best solution to the problem. Because they would achieve full deterrence with costless enforcements.

[39]Note that the entrepreneur could self-report if rejected. However, since the node is never going to be played in any non-exploitable design, the extension from this node is severed.

[40]Note that ex-ante, if policymakers set the reduced fine RF bigger than the expected value of being caught and convicted $\alpha\beta F$, they would not expect self-reporting from wrongdoers, only from enough risk-averse agents (Kaplow and Shavell, 1994).

Once again, corruption happens if the bureaucrat has any incentive to perform the bribery act. As seen before, this happens when returns from bribery are positive and the entrepreneur can credibly threat the bureaucrat with some punishment. In this case the threat of self-reporting.

Note that the agents control the probability of being detected α, because they can, at any time, turn themselves in to the authorities. In practice, being able to self-report after being detected necessarily allows agents to do so before as well. In this case, the analysis focuses in the fine reduction rule as fixed for self-reporting after detection. In the next case, the analysis allows for different fine reduction rules.

Given a fine reduction rule where PF is the reduced fine, for $P \in (-\infty, 1]$ in case of a unilateral self-reporting, and pF if both parties self-report simultaneously, where $PF < pF$ or $p > P$. In this case, if the entrepreneur wants to threat the bureaucrat with self-reporting, the reduction P must meet the same conditions as in the past example, i.e. $P < \alpha\beta$. Note that if $\beta > P > \alpha\beta$, the bureaucrat would expect that the entrepreneur would only self-report if detected. Thus, the threat is only credible in the case of certain detection ($\beta = \alpha\beta$). Therefore, if $P > \alpha\beta$ then self-reporting is not a credible threat.

It is important to observe that conditions for exploiting the game (5) are similar, i.e. agents would enter in a collusion just to self-report if

$$P < max\left[-\frac{-b + (1-\alpha\beta)\pi - \alpha\beta F}{F}; -\frac{-c + (1-\alpha\beta)b - \alpha\beta F}{F}\right].$$

Therefore, in order to the agent not exploit the game, it is necessary that the following condition is also present,

$$-PF > min\left[-b + p_\pi \pi - p_{F_E}F; -c + p_b b - p_{F_B}F\right]. \tag{6}$$

Consequently, there is no collusion. Similarly, given the rule that agents divide the reduced fine PF if they self-report at the same time, then both players would enter in corruption and both would self-report if $-\frac{1+P}{n}F = -pF > (-b + (1-\alpha\beta)\pi - \alpha\beta F)$ for the entrepreneur and $-\frac{1+P}{n}F = -pF > (-c + (1-\alpha\beta)b - \alpha\beta F)$ for the bureaucrat.

3.1.4 Self-Reporting Before and After Detection

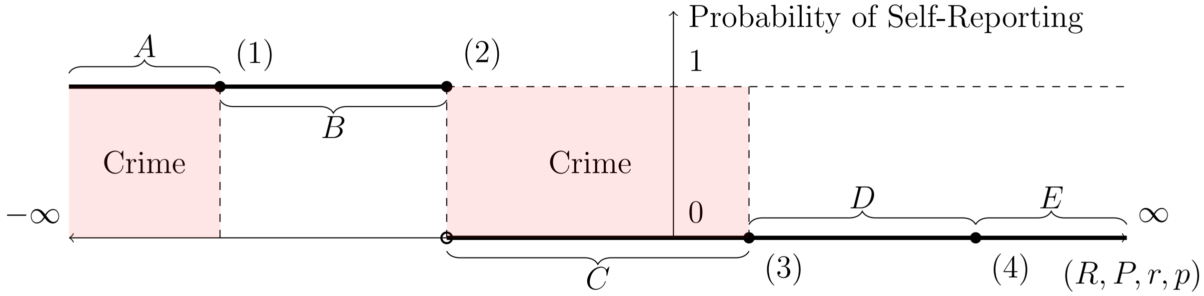
In the last case, agents would receive the same fine reduction P if self-report both before or after detected. Here, agents face different fine reduction rules if they self-report before authority's detection RF or after PF[41]. As previous examples show, the existence of profitable corruption schemes (1) and (2), non-exploitable leniencies (5) and (6), together with feasible self-reporting programs (R or $P < \alpha\beta$) are together sufficient conditions for inducing corruption. Once again, it happens because bribe givers gain a credible threat under these conditions.

In the case of different rules for fine reductions, the only remarkable issue lies in the special case in which the reduced fine for early self-reporters RF is non-negative and lower than the expected fine $\alpha\beta F$. Simultaneously, the reduced fine for post-detected self-reporters is $\alpha\beta F < PF < \beta F$. Consequently, in this particular case, there would be corruption led by the credible threat from RF. However, if the crime is detected, one would expect agents

[41] The leniencies are mutually exclusive, i.e. agents can only choose one regime of fine reduction. In other words, agents cannot self-report before detections and then plea bargain. No jurisdiction which I am aware would allow for it, nor would judges or prosecutors want, since one self-report is enough to convict the other involved party.

to self-report. The Figure 4 resumes the expected output from agents given the domains of R, r, P and p.

Figure 4: Fine Reduction Domain



Notes: (1) r or $p < min\left[-\frac{-b+(1-\alpha\beta)\pi-\alpha\beta F}{F}; -\frac{-c+(1-\alpha\beta)b-\alpha\beta F}{F}\right]$; (2) R or $P < max\left[-\frac{-b+(1-\alpha\beta)\pi-\alpha\beta F}{F}; -\frac{-c+(1-\alpha\beta)b-\alpha\beta F}{F}\right]$;(3) $\alpha\beta$; and (4)β. (A) Agents Collude and Self-Report; (B) Agents do not collude fearing other party's self-reporting; (C) Agents collude because first mover uses self-reporting as a credible threat; (D) No collusion or Collusion and Self-Report if detected only if $2 < R < 3$ and $3 < P < 4$; and (E) No Collusion and no Self-Reporting (Trust Game).

In summary, under complete information the fine reduction dynamics may induce corruption. It doesn't matter when the agents are allowed to self-report, only the size of the reduction[42]. The only crime deterrent case for leniencies is when a self-reporter receives a bonus greater than expected return from the corruption act, but not big enough that both agents could explore. In this case, the fear that the other party will exploit the game deters bribery.

Proposition 2. *There is a crime inducing effect from non-negative and high enough fine reductions for agents, independently if they can self-report before or after detection.*

3.2 Repeated and Complete Information

The one-shot bribe game can be seen as a small bribery act. This is the type of interaction that occurs occasionally, such as paying a civil servant to have priority in some process in a queue. However, if the game is repeated, this can represent a more serious offence. Generally, it is the indication of other associated crimes and is a common characteristic of big corruption cases. Of course, it does not mean strictly that one-time corruption is a small bribe and the repeated ones are big. Therefore, no judgement about the size (value involved) of the corruption scheme is done based on the periodicity.

In fact, the examples suggested in Section 3.1 are the same as here. However, agents can expect to repeat the bribery indefinitely. Needless to say, it is assumed that there is no job-rotation of bureaucrats.

3.2.1 Benchmark Without Fine Reductions for Self-Reporters

An indefinitely repeated game can be represented in a matrix form. The benefits of iterated collusion are represented as the present value from infinite collusions given a certain time

[42] In Figure 4, (1) is a minimum because both players need to show the condition for the equilibrium to change. Whereas, (2) if only one agent satisfy the condition then there is no collusion.

discount δ_i for $i = [E, B]$ and $\delta_i \in [0, 1]$. Lastly, defection is a one-time gain[43]. Given these characteristics, the matrix form of repeated collusion can be given as:

E/B	Perform	Not Perform	Reject
Bribe	$\frac{-b+(1-\alpha\beta)\pi-\alpha\beta F}{1-\delta_E}$ $\frac{-c+(1-\alpha\beta)b-\alpha\beta F}{1-\delta_B}$	$-b-\alpha\beta F$ $(1-\alpha\beta)b-\alpha\beta F$	$-\alpha\beta F$ 0

The possibility of repetition induces corruption when the entrepreneur plays first if:

$$\frac{-c + (1 - \alpha\beta)b - \alpha\beta F}{1 - \delta_B} > (1 - \alpha\beta)b - \alpha\beta F,$$

or,

$$c < \delta_B[(1 - \alpha\beta)b - \alpha\beta F],$$

or even,

$$\delta_B > \frac{c}{(1 - \alpha\beta)b - \alpha\beta F} = \frac{c}{p_b b - p_F F}. \tag{7}$$

This result shows that the entrepreneur is going to propose the bribe if bureaucrat's time discount is sufficiently bigger than the bureaucrat's cost benefit from corruption (right side of (7)). Symmetrically, it is straightforward to see that, corruption is going to happen when the bureaucrat plays first, if:

$$\delta_E > \frac{b}{(1 - \alpha\beta)\pi - \alpha\beta F} = \frac{b}{p_\pi \pi - p_F F}. \tag{8}$$

Note that, for a given set of pay-offs, corruption always happens if the time discount is big enough (Folk Theorem). In other words the time discount from agents needs to be bigger than the cost benefit from bribery in order to overcome the one-time defection constraint. Furthermore, if bribes are profitable, i.e. if conditions (1) and (2) are true than $\delta_i \in (0, 1]$.

Proposition 3. *Without fine reduction for self-reporters, if returns from briberies are positive and agents have high enough time discounts, there is a crime inducing effect from repetitions.*

3.2.2 All cases

After analysing individually the one-shot corruption game with complete information in each distinct self-reporting regimes, it becomes easier to proclaim the results for the repeated game. This happens because, if the game is not exploitable (5 and 6) and fine reductions are lenient enough (RP or $RF < \alpha\beta F$), then one would simply expect that corruption would be induced like a sequence of one-shot games. Therefore, choosing collusion is a Subgame Perfect Nash Equilibrium (SPNE) every round. Since, in this deterrence analysis, the parameters do not change over time, there is no need to expect a change in strategies.

Note that for less lenient fine reductions $R \in (\alpha\beta, 1]$, the game is similar to the benchmark in which there must be a sufficiently high time discount δ_i that make the collusion profitable. In other words, more lenient fine reductions drop the necessity for the time discount conditions (7) and (8). This result is similar to the one in Buccirossi and Spagnolo (2006).

[43]The underlining assumption is that, if agents see a defection, they punish the other player by never colluding again (grim trigger).

Proposition 4. *In a repeated non-exploitable corruption game with risk-neutral agents, repetitions only induce bribes if fine reductions are less lenient ($R \in (\alpha\beta, 1]$,) and agents cannot use them as credible threats.*

The result that repetition may induce corruption was also observed by Buccirossi and Spagnolo (2006). Nonetheless, this conclusion that collusion is less likely in short-run relationships in similar hierarchical structures dates back, at least, to Tirole (1986).

4 Incomplete Information

With incomplete information agents cannot know for sure what type of player one is playing against. However, they know what are the pay-offs for each type of player. In this sense, agents estimate the probability of each type and solve their decision problem accordingly.

Note that there are two distinct sources of uncertainties. Either the first mover does not know if the second mover will reject, accept and perform or accept and not perform the act. In other words, agents have incomplete information about the execution of the corruption act. Or, agents may not know if other players will be self-reporters or not, i.e there is incomplete information about self-reporting. Both cases are treated separately bellow.

4.1 One-Shot with Uncertainty About the Favour Execution

In this case, one can use the same examples as in Section 3.1. However, entrepreneurs now have some idea about the environment they are in. In other words, they estimate some probability of bureaucrats being of a certain type. In this specific example, entrepreneurs know that there are bureaucrats who are never going to accept the bribe, conversely there are some that will, and between them, some that one can trust to perform the favour and others that are going to defect. The bribe payer estimates the probability of each type and makes her decision accordingly.

4.1.1 Benchmark Without Self-Reporting

In this setting, suppose that there are three type of bureaucrats: honest who never accepts the bribes; dishonest and trustworthy who accepts the bribe and perform the act; dishonest and not trustworthy who accepts the bribe and do not perform the act[44]. Assuming that the probability of being of each type is λ_h, λ_t and λ_{nt} respectively, such that $\lambda_h + \lambda_{nt} + \lambda_t = 1$.

If the game is played under the conditions from Figure 2, the pay-off matrix with the type specific probability is the following:

E/B	Perform (λ_t)	Not Perform (λ_{nt})	Reject (λ_h)
Bribe	$-b + (1 - \alpha\beta)\pi - \alpha\beta F$	$-b - \alpha\beta F$	$-\alpha\beta F$
	$-c + (1 - \alpha\beta)b - \alpha\beta F$	$(1 - \alpha\beta)b - \alpha\beta F$	0
NB	0	0	0
	0	0	0

The entrepreneur decides what to do based on the bureaucrat's best option. In this game, for positive expected pay-offs from corruption, collusion will be a dominant strategy

[44]One can argue that these types of bureaucrat are not rational. However, this could be fixed by imposing a moral cost for each type of bureaucrat, such that pay-offs would be negative to honest bureaucrats that receive bribes, or trustworthy bureaucrats that do not perform the act.

for the entrepreneur if:

$$\lambda_t(-c + (1-\alpha\rho)b - \alpha\rho F) > \lambda_{nt}((1-\alpha\rho)b - \alpha\rho F),$$

rearranging,

$$\frac{\lambda_t - \lambda_{nt}}{\lambda_t} > \frac{c}{(1-\alpha\rho)b - \alpha\rho F}. \tag{9}$$

In conclusion, following the inequality (9), corruption occurs if the share of not trustworthy bureaucrats (left side) is bigger than the cost-benefit of corruption (right side) for bureaucrats. Likewise, the same results can be mirrored if the bureaucrat plays first and expect different types of entrepreneurs in the market, or else[45]:

$$\frac{\lambda_t - \lambda_{nt}}{\lambda_t} > \frac{b}{(1-\alpha\rho)\pi - \alpha\rho F}. \tag{10}$$

In this way, if the first mover believes (trust) that there are enough trustworthy people in the bribery market ((9) and (10)). Consequently, this disrupts the trust game structure of the case with complete information. Therefore, leading to bribes being offered.

Proposition 5. *With incomplete information over the performance of corruption and without enforceable agreements, a one-shot corruption is induced if there are enough trustworthy agents in the bribery market.*

4.1.2 All cases

As seen above. The uncertainty about the type of agent can induce some agents to propose bribes. In this setting, there are no cases in which first movers would rely on credible threats in order to propose the bribe. This due to the fact that the execution of the favour is now taken for granted. In other words, the proposal of the bribe does not depend on the possibility of the other party self-reporting or not[46].

The second mover, conversely, can play rationally or according to the first mover expectations. If bureaucrats are still playing rationally, it is expected that the number of bribes paid is not going to change with the addition of fine reductions. Hence, there are corruption acts that are going to be performed if R or $P < \alpha\beta$. Because bureaucrats would fear retaliation. However, there are bribes that are not going to be performed if R or $P > \alpha\beta$. Consequently, the not performed corruption is not going to be self-reported before any detection, with the exception of the case where $\beta > P > \alpha\beta$, in which agents would report if detected.

Importantly, it is expected that, if R or P are non-exploitable (5 and 6) and lower than $\alpha\beta$, then there would be fewer corruption cases under incomplete information than it would under complete information. This happens because the trust conditions ((9) and (10)) that disrupts the trust game, act like a constraint in the case where conditions would favour corruption. In other words, for these levels of fine reduction, profitable bribes (1) and (2)

[45]One can suggest that the proportion of honest and dishonest is some function of the public enforcement variables α, β and F. However, for now, the proportions are considered constant. This result may support the hypothesis that there are some environments more prone to corruption than others.

[46]One can propose that credible threats could be used against untrustworthy bureaucrats to induce collusion. This is possible, but for simplicity it is assumed that the untrustworthy agent has impeditive (moral) costs to retribute the favour. However, the reader may also suppose there that could be some hold-up problem in this case.

are enough to induce corruption under complete information. Whereas under incomplete information agents would need to fulfil conditions (1), (2), (9) and (10).

If agents play according to the expectations[47]. Then, due to miscalculated trust, some bribes are going to be paid to untrustworthy agents. In this case, if the bureaucrat is not performing the act and if fine reductions R or P are lower than $\alpha\beta$, then the entrepreneur is going to self-report as retaliation. It would be the only case in which one would observe self-reporting as retaliation for non-executed corruption.

Proposition 6. *With incomplete information over the performance of corruption, fine reductions for self-reporters do not deter bribe offering. They may only determine if corruption is going to be performed or not.*

4.2 Repeated with Uncertainty About the Favour Execution

If agents expect to have repeated interactions, then through Folk Theorem there is a time discount δ_i that makes it profitable to enter collusion. The question here is to compare the time discount between the case with complete and incomplete information.

In this way, constraints from incomplete information over bribery performance (9) and (10) must be lower than the incentives given by repetitions (7) and (8). More specifically, agents would offer bribes if outcomes from repeated interactions are greater than the probability of the one-time defection. Combining both restrictions then:

$$\delta_i > \frac{\lambda_t - \lambda_{nt}}{\lambda_t}$$

In other words, if the time discount for both players is bigger than the expected proportions of trustworthy bribe receivers in the system, then corruption is going to be performed in repeated games and not offered in one-time games.

Proposition 7. *With incomplete information, repetition induces crime if time discounts are bigger than the share of trustworthy bribe receivers.*

4.3 One-Shot with Uncertainty About Self-Reporting

In this case agents do not have complete information about their or the other parties' decision to self-report, i.e. agents do not know which type of agent they are facing (a potential self-reporter or not). This is the same as if agents consider an autonomous exogenous probability of self-reporting sometime after starting a corruption scheme[48]. Note that it does not mean that game has lost its dynamic. It means that agents consider some ex-ante probability of self-reporting due to uncertainty, but they can still rationally use self-reporting in their advantage whenever pay-offs are favourable. Be it by using self-reporting as a credible threat or to exploiting the game[49].

[47] Second movers would rationally play according to first movers expectations if they have some different cost structure for each type. For simplicity, note that it is not necessary to build the structure, just necessary to know that there are players that would always reject, always perform the bribe or not.

[48] Be it because they do not know the future or they do not know perfectly what are the variables or any other characteristic that could make the other party self-report.

[49] Harrington (2013) approach the possibility of the other party self-reporting as signs, that determine the probability of being detected. Here, the non-lenient enforcements (α and β) and the probability of being reported are independent. Nonetheless, there are similar results, even though his model approaches the problem in a ex-post state with a simultaneous game. Therefore, the results can be trusted in slightly different hypothesis.

Note that, differently from the past cases, there is no benchmark without fine reductions to compare to. Obviously, there are only cases in which agents can self-report. Therefore, the impact of uncertainty about parties' decision to self-report is discussed below.

4.3.1 Self-Reporting Before Detection

With complete information, the decision to self-report was completely deterministic. In this case agents consider that there is some source of uncertainty, in other words, there is a probability of agents self-reporting after committing a crime given by γ_i for $i = [E, B]$[50]. For simplicity, let γ_i be exogenous[51] and constant in all the feasible domain of R.

Given the above conditions, profitable bribes (1) and (2) are now a function of ($\alpha, \beta, \gamma_B, \gamma_E, R$ and R)[52]. In this case,

$$p_{F_E} \equiv \underbrace{\alpha\beta(1-\gamma_E)(1-\gamma_B)}_{\text{Prob. being convicted}} + \underbrace{(1-\gamma_B)\gamma_E}_{\text{Prob. reporting}} R + \underbrace{(1-\gamma_E)\gamma_B}_{\text{Prob. being reported}} + \underbrace{\gamma_E\gamma_B}_{\text{Prob. reporting together}} r,$$
$$p_\pi \equiv \underbrace{(1-\alpha\beta)(1-\gamma_E)(1-\gamma_B)}_{\text{Prob. of success}},$$

aggregating for $(1-\gamma_E)(1-\gamma_B) = \Gamma$, then,

$p_{F_E} \equiv \alpha\beta\Gamma + (1-\gamma_B)\gamma_E R + (1-\gamma_E)\gamma_B + \gamma_E\gamma_B r$, and
$p_\pi \equiv (1-\alpha\beta)\Gamma$.

The chance of being fined now depend on another player's decision. Therefore, condition (1),

$$-b + p_\pi\pi - p_F F > 0,$$

can be rewritten as,

$$-b + \pi\Gamma(1-\alpha\beta) - F(\alpha\beta\Gamma + (1-\gamma_B)\gamma_E R + (1-\gamma_E)\gamma_B + \gamma_E\gamma_B r) > 0.$$

By the same way, for the bureaucrat, if:
$p_{F_B} \equiv \alpha\beta(1-\gamma_E)(1-\gamma_B) + (1-\gamma_B)\gamma_E + (1-\gamma_E)\gamma_B R + \gamma_E\gamma_B r$,
$p_b = p_\pi \equiv (1-\alpha\beta)(1-\gamma_E)(1-\gamma_B)$,
or,
$p_{F_B} \equiv \alpha\beta\Gamma + (1-\gamma_B)\gamma_E + (1-\gamma_E)\gamma_B R + \gamma_E\gamma_B r$,
$p_b \equiv (1-\alpha\beta)\Gamma$ than condition (2),

$$-c + p_b b - p_F F > 0,$$

is now written as,

$$-c + b(1-\alpha\beta) - F(\alpha\beta\Gamma + (1-\gamma_B)\gamma_E + (1-\gamma_E)\gamma_B R + \gamma_E\gamma_B r)\Gamma > 0.$$

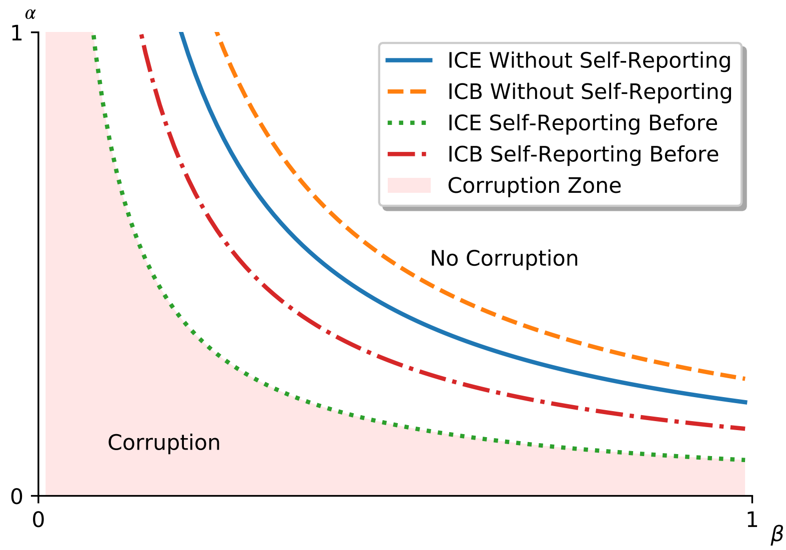
[50]Note that it is algebraically equal to use an exogenous probability of γ_i to the situations in which self-report is happening than it is to use a probability of γ_i for an agent being a 'self-reporter' and a $(1-\gamma_i)$ for a 'not self-reporter'.

[51]Of course, one might think it is a function of α, β, R or γ_j. This would add a lot of complexity to the model, and conditions for an internal solution to the agent's maximization problem would turn dynamic. Nonetheless, this is an excellent topic for further studies.

[52]Technically p_{F_i} is not strictly the probability of F_i being paid, but the expected value of F_i divided by F_i, since $E[F_i] = p_{F_i}F_i$. And it is given by the probabilities of F_i, RF_i or rF_i happening. The same interpretation can be used in the following example with self-reporting after detections.

Given the above conditions, p_{F_i} is different for each player. And both expected values of being fined or earn the corruption premium are now dependent on the other player's action. More specifically, for any positive probabilities of self-reporting $\gamma_i > 0$ the expected fine $p_F F$ is bigger and expected advantages from corruption ($p_b b$ and $p_\pi \pi$) are smaller in relation to the case without self-reporting. Therefore, the self-reporting uncertainty decreases the combination set in which corruption would be feasible for the same level of enforcements α and β. Hence, the deterrence effect can be seen by the retraction of the indifference curves in relation to the ones with no self-reporting. Figure 5 shows the indifference curves for the same level of π, b and F from Figure 1 along with $\gamma_i > 0$ and $R > 0$.

Figure 5: Indifference Curves With Self-Reporting Before Authorities' Detection



In summary, there is a deterrent effect from the uncertainty of the other party self-reporting $\gamma_i > 0$. Agents would require a softer combination of α and β to start colluding [53]. In other words, if one does not know if the other party will self-report or not, then it would be necessary that, at least, the crime is more difficult to detect or to be convicted for.

Note that, γ_i is a probability of self-reporting after the crime is committed, agents still need to figure out if crime is going to be performed. Like in the past examples, if entrepreneur plays first, than the bureaucrat performs the crime if agents can credibly threat the other party[54]. Since, the condition in R for credible threats (feasible fine reductions) is $RF < p_{F_i} R$, i.e. reduced fine is smaller than the expected fine. The constraint moved up from $\alpha\beta$ to p_{F_i}[55]. Fine reductions need to be less lenient for second movers to perform their act. By the same way, conditions for non-exploitable fines 5 and 6) are now less lenient than $R < max\left[-\frac{-b+(1-\alpha\beta)\pi-\alpha\beta F}{F}; -\frac{-c+(1-\alpha\beta)b-\alpha\beta F}{F}\right]$[56].

[53]Notice that, b and π are fixed in this example. If one allows them to vary, it could be the case that, in order for corruption to be profitable under self-reporting constraints, the price of bribe and the size of the corruption may go up as Bigoni et al. (2012) found in their experiment with higher cartel prices.

[54]Assuming incomplete information from the second mover also would result in always performing the act and the analysis loses its dynamic characteristic.

[55]There is a $\hat{R} \in (0, 1]$ such that $p_{F_i}(R) = \hat{R}$ and $p_{F_i} > \alpha\rho$. The quick interpretation is that, if $p_b = p_\pi = f(1 - p_{F_i})$, if γ_i and γ_j are positive, then the decrease in p_π or p_b follows an increase in $p_{F_i}(R)$. Proof is in Appendix X.

[56]Once again, this result only holds in a setting where γ_i is not a function of R.

The conditions above imply that, circumstances in which agents would use fine reductions as credible threats have increased. However, conditions in which agents would exploit the game are less lenient, thereby less costly to implement.

Lastly, one can note that p_{F_i} is monotonically increasing in R, $\left(\frac{\partial p_{F_i}}{\partial R} > 0\right)$. In this case, optimum leniency R should be the maximum feasible p_{F_i}. However, if γ_i is some function of R it is likely that optimal fine reduction is an interior solution within the non-exploitable and feasible domain of R. To solve this problem, one must have a good theory on how agents imperfectly calculate their probability of self-reporting γ_i. This issue is revisited in the next section with endogenous probability of self-reporting.

4.3.2 Self-Report After Detection

Like in the past examples with complete information. The bribe is going to be performed if fine reductions for self-reporters after detection are feasible and non-exploitable.

In this case, let ω_i be the probability of an agent to self-report for $i = [E, B]$ and $\Omega = (1-\omega_E)(1-\omega_B)$ be the probability of not self-reporting, then p_{F_i} and $p_\pi = p_b$ are $f(\alpha, \beta, \omega_B, \omega_E, P, p)$ which is different for each i, such as,

$p_{F_E} \equiv \alpha[\beta\Omega + (1-\omega_B)\omega_E P + (1-\omega_E)\omega_B + \omega_B\omega_E p]$ and
$p_\pi = p_b \equiv (1-\alpha) + \alpha\Omega(1-\beta)$.

By symmetry, for the bureaucrat, if

$p_{F_B} \equiv \alpha[\beta\Omega + (1-\omega_B)\omega_E + (1-\omega_E)\omega_B P + \omega_B\omega_E p]$,

So, the profitability condition for the entrepreneur (1), can be written as,

$$-b + p_\pi \pi - p_F F > 0,$$

or else,

$$-b + \pi((1-\alpha) + \alpha\Omega(1-\beta)) - F\alpha[\beta\Omega + (1-\omega_B)\omega_E P + (1-\omega_E)\omega_B + \omega_B\omega_E p] > 0.$$

Likewise for the bureaucrat, the decision rule is also,

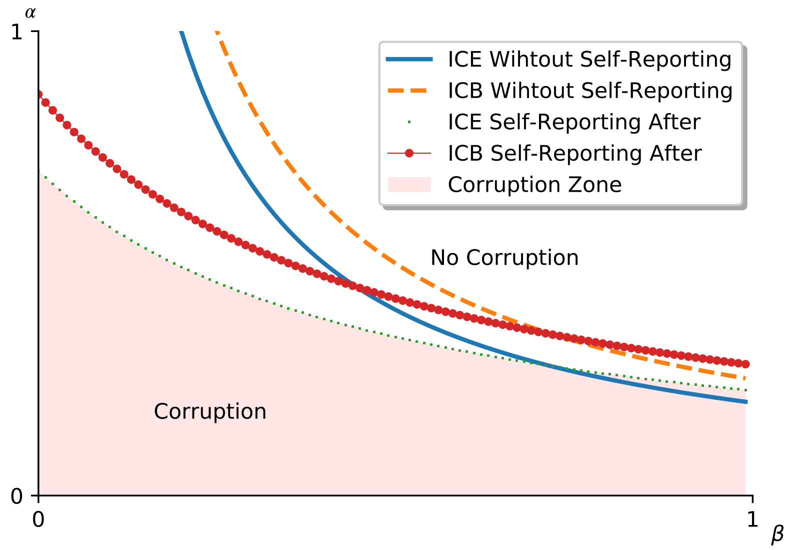
$$-c + p_b b - p_F F > 0,$$

or else, when substituting the terms ,

$$-c + b((1-\alpha) + \alpha\Omega(1-\beta)) - F\alpha[\beta\Omega + (1-\omega_B)\omega_E + (1-\omega_E)\omega_B P + \omega_B\omega_E p > 0.$$

The most notable difference between self-reporting before or after detection is given by the impact of α in the decision choice. In the current case, by definition, agents would fear self-reporting only if detected. Or else, they are more afraid of prosecutors than they are afraid of the judges. This peculiar effect can be shown in the Figure 6, in which indifference curves are more restrictive to increases in the probability of detection α than they are with probabilities of conviction β.

Figure 6: Indifference Curves for Self-Reporting After Detection



Note that, for a given fine reduction P, crimes with a high enough probability of conviction $\tilde{\beta}$ (crimes that produces undeniable evidence), self-reporting after detection may be crime inducing. This is due to the fact that, if agents know they can self-report after being detected, they will do it, because there is almost certainty of conviction in any case.

4.3.3 Self-Reporting Before and After Detection

If fine reductions are feasible and non-exploitable, then corruption is performed if profitability conditions are met. In order to calculate them in a staged game with incomplete information ($\gamma_i > 0$ and $\omega_i > 0$) and distinct rules for fine reductions R and P, let:

$p_{F_i} = f(\alpha, \beta, \gamma_E, \gamma_B, \omega_B, \omega_E, P, R, p, r)$,
$p_{F_E} \equiv \alpha\Gamma(\Omega\beta + (1-\omega_B)\omega_E P + (1-\omega_E)\omega_B + \omega_E\omega_B p) + (1-\gamma_B)\gamma_E R + (1-\gamma_E)\gamma_B + \gamma_E\gamma_B r$,
$p_{F_B} \equiv \alpha\Gamma(\Omega\beta + (1-\omega_B)\omega_E + (1-\omega_E)\omega_B P + \omega_E\omega_B p) + (1-\gamma_B)\gamma_E + (1-\gamma_E)\gamma_B R + \gamma_E\gamma_B r$,
in the same way,
$p_\pi = p_b \equiv \Gamma[(1-\alpha) + \alpha\Omega(1-\beta)]$.

Given that agents still face the same problem of having a positive expected pay-off from collusion (1 and 2). Then, the constraints given by the positive probability of being reported may decrease expected pay-offs from corruption through bigger expected fines $p_{F_i}F$ and lower probability of going unpunished and receiving the advantage of corruption p_b and p_π. Therefore, the decision rule for the entrepreneur can be expressed as:

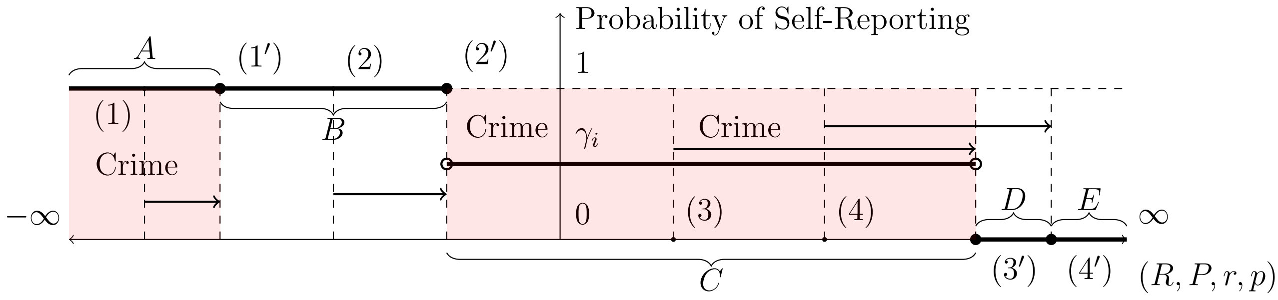
$$-b + \pi\Gamma[(1-\alpha) + \alpha\Omega(1-\beta)] - F(\alpha\Gamma(\Omega\beta + (1-\omega_B)\omega_E P + (1-\omega_E)\omega_B + \omega_E\omega_B p) + (1-\gamma_B)\gamma_E R + (1-\gamma_E)\gamma_B + \gamma_E\gamma_B r) > 0,$$

and for the bureaucrat,

$$-c + b\Gamma[(1-\alpha) + \alpha\Omega(1-\beta) - F(\alpha\Gamma(\Omega\beta + (1-\omega_B)\omega_E + (1-\omega_E)\omega_B P + \omega_E\omega_B p) + (1-\gamma_B)\gamma_E + (1-\gamma_E)\gamma_B R + \gamma_E\gamma_B r) > 0.$$

Regarding the dynamic constraint from using self-reporting as a credible threat to induce collusion, Figure 7 resumes how the domain in R and P changed with the addition of self-reporting uncertainty γ_i and ω_i.

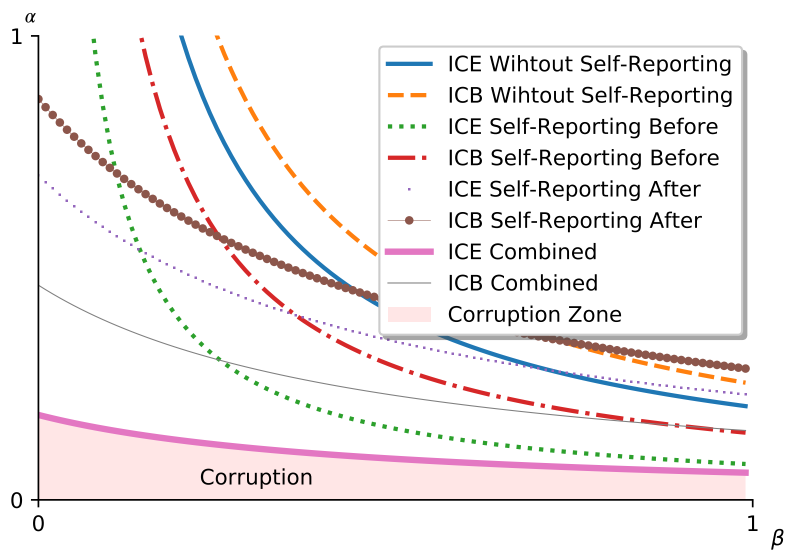
Figure 7: Fine Reduction Domain



Notes: Arrows from i to i', show the direction of the effect when the probability of self-reporting γ_i and ω_i grow. (1) r or p $<$ min $\left[\frac{-b+p_\pi\pi-p_{F_E}F}{F}; -\frac{-c+p_b b-p_{F_B}F}{F}\right]$;(2) R or P $<$ max $\left[-\frac{-b+p_\pi\pi-p_{F_E}F}{F}; -\frac{-c+p_b b-p_{F_B}F}{F}\right]$;(3) p_{F_i} ; and (4) $p_{F_i}|\alpha = 1$. (A) Agents collude and self-report; (B) Agents do not collude fearing other party's self-reporting; (C) Agents collude because first mover uses self-reporting as a credible threat; (D) No collusion or Collusion and Self-Report if detected only if $2' < R < 3'$ and $3' < P < 4'$; and (E) No Collusion and no Self-Reporting (Trust Game).

Computing the indifference curves from the above conditions and comparing with the previous calculated ones, from Figure 8 it is clear to see the deterrent effect from the combined policies when $\gamma_i > 0$ and $\omega_i > 0$. The deterrence effect is given by the reduced set of combinations of the public enforcements in which corruption would be feasible.

Figure 8: Indifference Curves With Combined Policies

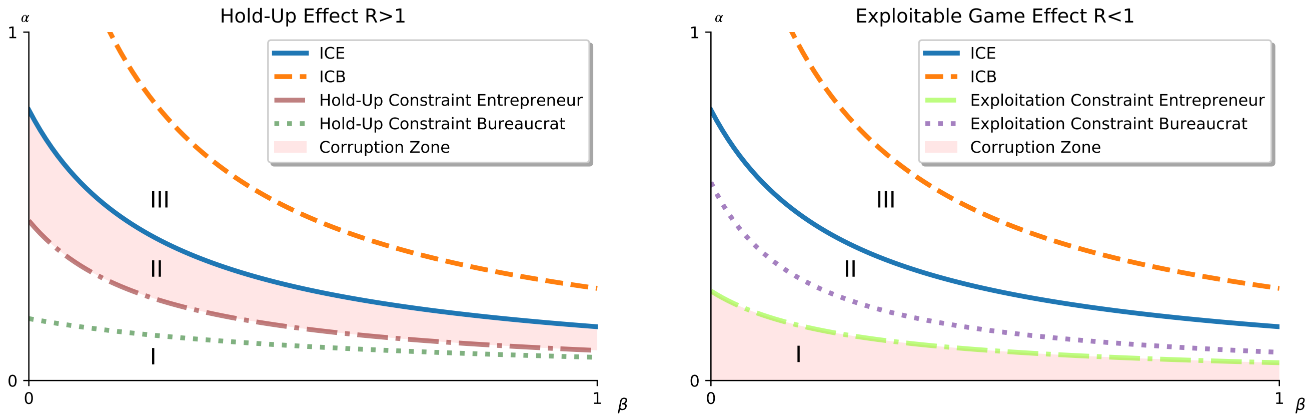


Proposition 8. *With incomplete information over self-reporting, the possibility of reporting ($\gamma_i \geq \gamma_j > 0$) and/or being reported ($\omega_i \geq \omega_j > 0$) cumulatively deter corruption by decreasing expected pay-offs from corruption. Whereas the size of de reductions P and R induces the crime trough decreasing expected fines.*

The effect above is similar to the 'pre-emption effect' pointed by Harrington (2013), although he assesses the effect from private signals in an environment after the crime is committed. Nonetheless results point to the same mechanism, the fear of being reported feeds fear on the other party, mutually affecting the expected value from self-reporting. Spagnolo (2005) also considers the role of uncertainties by another methodology, using riskiness analysis, he finds that the risk from being reported affects the decision to collude more than the current level of detection.

In summary, there are 3 different constraints to engage in corruption, all of them are affected by uncertainties, as Figures 7 and 8. First, there are the exploitation constraints (5 and 6), in which agents would not collude because they know that the other agent is going to self-report. Up to the point that it is interesting for both agents to self-report, then they collude. Second, is the feasibility condition, in which less lenient fine-reductions $(R, P > p_{F_i})$ would create a hold-up problem, i.e. the second mover is not going to perform the favour and the first mover is not going to propose a bribe. Last, the profitability of the bribes (1 and 2). The Figure 9 shows all the constraints in an environment with a noisy uncertainty about the other player self-reporting.

Figure 9: Indifference Curves With Combined Policies



In the Figure 9, both graphics show the same profitability conditions, where ICE is the indifference curve for the entrepreneur (1) and ICB for the bureaucrat (2).

The left curves are less lenient fine reductions R or P, there is some decrease in fines but no bonuses. In this case, the Hold-up constraints are represented as the feasibility condition $(R, P > p_{F_i})$. The area I, is the hold-up area, therefore there is no corruption. In II there is no hold-up and bribes are profitable, consequently agents would engage in corruption. Lastly, in III there is no hold-up problems but also no profitable bribes, then agents would not collude.

In the right case, there is a bonus for self-reporting. Additionally, the exploitation constraints are represented by (5) and (6). Therefore, the area I represent a profitable and non-exploitable combination of enforcements, so there is corruption. However, in II, bribes are profitable but exploitable, so there is no corruption, since agents know other parties are going to defect. Lastly, area III is both exploitable and not profitable.

4.3.4 Endogenous Probability of Self-Reporting

From past examples it is possible to observe that uncertainty about self-reporting represents a major constraint to corruption. However, uncertainties were not related to the fine reductions R and P. By this way, it was possible to isolate the deterrent effect from introducing leniency to self-reporters in a game without complete information from agents.

Note that the only effect from fine reductions changes R and P in the profitability from corruption (1 and 2), for given probabilities of self-reporting γ_i or ω_i, is that less lenient fine reductions deter corruption through higher expected fines $p_{F_i}F$. However, it is likely that the probability of self-reporting γ_i or ω_i is some negative function of R and P. Although, there is no good theory about the functional form of this relation[57], the linear case is the one addressed here.

In order to address the effect from leniencies over corruption deterrence, let θ_k be the linear coefficient that relates the k rules of fine reduction R and P to the probability of self-reporting γ_i and ω_i, in which $f'(\theta_k) < 0$, and $f''(\theta_k) = 0$ such that,

$$\gamma_i = \gamma_{i0} + \theta_R(1 - R),$$

and,

$$\omega_i = \omega_{i0} + \theta_P(1 - P).$$

If the functions above are bounded between 0 and 1, then, there must be a sufficiently high θ_k^* that makes the deterrent effect from a more lenient R bigger than the crime inducing effect. In other words, there is a point where the fear of other parties self-reporting (pre-emption effect (Harrington, 2013)) is higher than the crime inducing effects of the reduced expected fines from leniency. The proof of this statement is on Appendix X.

Lastly, the reader is invited to try different combinations of parameters in an interactive graph and check how indifference curves behave. An interactive application from this model is available at `https://caxaxa.github.io/Interactive_corruption_deterrence_model/`.

4.4 Repeated With Uncertainty About Self-Reporting

As in the previous example of repetitions with complete information. If fine reductions R or P are greater than p_{F_i}, in other words, within the interval of D and E from Figure 7, than repetition can induce corruption. Since they give the necessary incentive to overcome one-time defection constraint (Folk Theorem). In this case, agents need to have high enough time discount δ_i that disrupts the prisoners dilemma and make them collude (7) and (8). If such condition is met, agents only need expected returns of corruption to be positive (1) and (2).

4.4.1 All cases

From previous examples, $p_{Fi} = f(R, P)$, thereby $\delta_i = f(R, P)$, and it must be different in every different type of leniency policy.

Note that, if $\gamma_i > 0$ and $\omega_i > 0$, then expected fine $p_{Fi}F$ is now bigger and probabilities of receiving the advantages of corruption p_b and p_π are smaller. Consequently, through (7) and (8), $\delta_i^{nsr} \leq min[\delta_i^{srb}, \delta_i^{sra}] \leq max[\delta_i^{srb}, \delta_i^{sra}] \leq \delta_i^{srba}$ for 'nsr' 'srb', 'sra' and '$srba$' being respectively no probability of self-report ($\gamma_i = 0, \omega_i = 0$), positive probability of self-reporting only before detection ($\gamma_i > 0, \omega_i = 0$), positive probability self-reporting after detection ($\gamma_i = 0, \omega_i > 0$) and positive probability of self-reporting at any time ($\gamma_i > 0, \omega_i > 0$). In other words, if the probabilities of self-reporting grow (before or after detection), then the time discount necessary to engage in corruption must be higher than they would need to be in previous cases.

[57] Perhaps, hints about how fine reductions affect the probabilities of self-reporting could be better explored studying the disruption effect from self-reporting, i.e. ex-post behaviour of agents in a stochastic system.

5 Conclusion

This paper explores the ex-ante deterrent effect from leniency programs in corruption crimes. Literature suggests that giving leniencies to self-reporters can be counterproductive against corruption. Sophisticated players could use the law in their favour, thereby incentivizing criminal behaviour (Buccirossi and Spagnolo, 2000, 2001, 2006). Nonetheless, recent studies in leniency policies (for cartels) point out to the crime deterrent role of risk and uncertainty generated by the possibility of being reported (Spagnolo, 2005; Harrington, 2008, 2013). Along with other positive effects that the policy addresses to the problem (Marvão and Spagnolo, 2014). In this sense, most of the conclusions about fine reduction to self-reporters come from antitrust literature, a lot less is confirmed using the specific framing of corruption (Burguety et al., 2016). This paper goes deeper into this reasoning and shows that most of the potential deterrent effect from the policies come from generating uncertainties to the wrongdoers. Moreover, it shows that the potential differences between types of wrongdoers and how they devise their perceptions about other players are crucial for the success of the policies.

First, the study shows that, in order to perform corruption, bribes need to be profitable (conditions (1) and (2)). Then, it is crucial to define the level of information from the players about their environment.

If agents are very sophisticated and they know well the players they are playing against, the information is assumed to be complete. In this setting, the absence of leniency to self-reporters deters one-shot non-enforceable agreements through the sequential nature of the game. Therefore, this example is similar to a trust game, in which rational players would always defect. It is also known as the 'hold-up' problem from contract theory. However, if leniencies are given to self-reporters and if sanction reductions are non-exploitable (conditions (5) and (6)) and feasible (Figure 4) self-reporting is used as a credible threat by first movers, inducing corruption. Importantly, it does not matter if agents can self-report before or after being detected by an authority. Additionally, if policymakers allow bonuses to self-reporters, then players can enter in corruption just to exploit the game and self-report. Consequently, first movers can predict this behaviour and not engage in corruption. However, this constraint for corruption holds until the point that bonuses are so high, that it is profitable for both agents to self-report. Lastly, repetition can only induce corruption when fine reductions are less lenient and cannot be used as credible threats.

If information is incomplete, if there is some autonomous exogenous probability of agents self-reporting, then the outcomes change. Uncertainty reduces the expected benefits from corruption and increases expected fines. Therefore, it constrains the profitability ((1) and (2)) from bribes and the necessary time discounts in repeated interactions (7) and (8), deterring corruption. Furthermore, it also changes the domain in which fine reductions are feasible and non-exploitable ((5) and (6)), changing the set of conditions in which corruption would be performed. Notably, it decreases the bonuses necessary to exploiting the game, making less costly to prevent corruption by this channel. On the other hand, it also decreases the set of conditions in which 'hold-up' problems would deter corruption. Consequently, even less lenient fine reductions can be used as credible threats. Lastly, under uncertainty, allowing agents to self-report before or after being detected by the authorities may have a cumulative effect over corruption deterrence.

Lastly, this study speculates on the role of fine reductions in the agents probability of self-reporting. If this relation is negative, there must be a point in which the crime inducing effect from lower expected fines is smaller than the crime deterrent effect from fear of being reported. In this case, more lenient fine reductions would deter corruption. This relation must be better explored by a more robust theory. Perhaps the ex-post analysis of leniency policies can help in this task. Certainly a more challenging set would involve the decision from agents after they committed a crime. Or else, the disruptive effects of self-reporting. This approach would require letting variables change over time. Perhaps, a state space model with dependent laws of motion may be a good way to start.

Other possible extensions to this simple model may account for: the welfare maximization problem; corporate Vs individual self-reporting; non monetary sanctions; different liability regimes (agency costs); asymmetric punishments; recidivism; endogenizing variables or letting them change over time; allowing Type I errors; risk-preferences; and social norms or other behavioural biases.

References

Abbink, K. and Wu, K. (2017). Reward self-reporting to deter corruption: An experiment onmitigating collusive bribery. *Journal of Economic Behavior and Organization.*

Arlen, J. and Kraakman, R. (1997). Controling coorporate missconduct:an analysis of coorporate liability regimes. *New York University Law Review.*

Aubert, C., Rey, P., and Kovacic, W. E. (2006). The impact of leniency and whistle-blowing programs on cartels. *International Journal of Industrial Organization.*

Basu, K. (2011). Why, for a class of bribes, the act of giving a bribe should be treated as legal. Technical report, Ministry of Finance Government of India.

Basu, K. and Cordella, T. (2016). Asymmetric punishment as an instrument of corruption control. *Journal of Public Economic Theory.*

Becker, G. (1968). Crime and punishment: An economic approach. *Journal of Political Economy.*

Berg, N. and Kim, J.-Y. (2018). Plea bargaining with multiple defendants and its deterrence effect. *International Review of Law and Economics.*

Bigoni, M., Fridolfsson, S., Le Coq, C., and Spagnolo, G. (2012). Fines, leniency, and rewards in antitrust. *RAND Journal of Economics.*

Bigoni, M., Fridolfsson, S., Le Coq, C., and Spagnolo, G. (2015). Trust, leniency and deterrence. *Journal of Law Economics and Organization.*

Brenner, S. (2009). An empirical study of the european corporate leniency program. *International Journal of Industrial Organization.*

Buccirossi, P. and Spagnolo, G. (2000). Counterproductive leniency programs against corruption. *Working Paper.*

Buccirossi, P. and Spagnolo, G. (2001). The effects of leniency on illegal transactions: How (not) to fight corruption. *Working Paper.*

Buccirossi, P. and Spagnolo, G. (2006). Leniency policies and illegal transactions. *Journal of Public Economics.*

Burguety, R., Ganuzaz, J.-J., and Montalvo, J. G. (2016). The microeconomics of corruption. a review of thirty years of research. *Working paper.*

Chang, M.-H. and Harrington, J. E. (2015). When can we expect a corporate leniency program to result in fewer cartels? *Journal of Law and Economics.*

Chen, Z. and Rey, P. (2013). On the design of leniency programs. *Journal of Law and Economics.*

Dufwenberg, M. and Spagnolo, G. (2014). Legalize bribe giving. *Economic Inquiry.*

Engel, C., Goerg, S., and Yu, G. (2016). Symmetric vs. asymmetric punishment regimes for collusive bribery. *American Law and Economics Review.*

Franzoni, L. A. (1999). Negotiated enforcement and credible deterrence. *The Economic Journal.*

Garoupa, N. (1997). The theory of optimal law enforcement. *Journal of Economic Surveys.*

Garoupa, N. (2007). Optimal law enforcement and criminal organizations. *Journal of Economic Behavior Organization.*

Givati, Y. (2014). Legal institutions and social values: Theory and evidence from plea bargaining regimes. *Journal of Empirical Legal Studies.*

Grossman, G. and Katz, M. (1983). Plea bargain and social welfare. *American Economic Review.*

Harrington, J. E. (2013). Corporate leniency programs when firms have private information: The push of prosecution and the pul of pre-emption. *The Journal of Industrial Economics.*

Harrington, J. J. E. (2008). Optimal corporate leniency programs. *The Journal of Industrial Economics.*

Kaplow, L. and Shavell, S. (1994). Optimal law enforcement with self reporting of behavior. *Journal of Political Economy.*

Kaufmann, D. and Wei, S. (1999). Does "grease money" speed up the wheels of commerce? *National Bureau of Economic, Research Working Paper 7093 Cambridge.*

Kobayashi, B. (1992). Deterrence with multiple defendants: An explanation for "unfair" plea bargains. *The RAND Journal of Economics.*

Lambsdorff, J. and Nell, M. (2007). Fighting corruption with asymmetric penalties and leniency. *Working Paper.*

Landes, W. (1971). An economic analysis of the courts. *Journal of Law and Economics.*

Marvão, C. and Spagnolo, G. (2014). What do we know about the effectiveness of leniency policies? a survey of the empirical and experimental evidence. *Working Paper.*

Miller, N. H. (2009). Leniency and cartel enforcement. *American Economic Review.*

Motchenkova, E. (2004). Effects of leniency programs on cartel stability. *CentER discussion papers series, Tilburg University.*

Motchenkova, E. and van der Laan, R. (2011). Strictness of leniency programs and asymmetric punishment effect. *International Review of Economics.*

Motta, M. and Polo, M. (2003). Leniency programs and cartel prosecution. *International Journal of Industrial Organization.*

Mungan, M. C. and Klick, J. (2016). Reducing false guilty pleas and wrongful convictions through exoneree compensation. *Journal of Law and Economics.*

Oded, S. (2012). Inducing corporate proactive compliance: Liability controls and corporate monitors.

Polinsky, M. and Shavell, S. (2001). Corruption and optimal law enforcement. *Journal of Public Economics.*

Polinsky, M. and Shavell, S. (2007). Theory of public enforcement of law. *Handbook of Law and Economics.*

Rose-Ackerman, S. (1975). The economics of corruption. *Journal of Public Economics.*

Rose-Ackerman, S. (2006). *International Handbook on the Economics of Corruption.* Edward Elgar Publishing Limited.

Rose-Ackerman, S. (2010). The law and economics of bribery and extortion. *The Annual Review of Law and Social Science.*

Shleifer, A. and Vishny, R. W. (1993). Corruption. *The Quarterly Journal of Economics.*

Spagnolo, G. (2005). Divide et impera optimal deterrence mechanisms against cartels and organized crime. *Working Paper.*

Søreide, T. and Rose-Ackerman, S. (2018). Corruption in state administration. *The Research Handbook on Corporate Crime and Financial Misdealing - Edited by Jennifer Arlen.*

Tirole, J. (1986). Hierarchies and bureaucracies: On the role of collusion in organizations. *Journal of Law Economics and Organization.*