Modelling the Corruption Game

Lucas Alves Chacha

January 12, 2022

Contents

1 Introduction

Chapter 2 presents the multiple types of corruption. Additionally, it narrows the scope of this study to the analysis of an specific set of features from anti-corruption policies. Notably, as stated before, the game design from this work must have a criminal transaction setting (simple game) and allow for non-trial resolutions (complete game). In this sense, this Chapter revisits the corruption game from Chapter 2 and formalizes it in a more structured way. Furthermore, the game is solved for the players' best decision.

It is important to highlight that, despite the fact that in the first examples (sections 2 and 3) the game can be solved using only the expected returns from corruption[1]. Here, as the problems get more complex, they require a

[1]As it was shown in Chapter 2, Section $XXXX$ in a straightforward way. The idea was to quickly introduce the study's problem.

more sophisticated solution technique. Notably, they are solved using dynamic programming. Therefore, the formal definition of the game presented here allows it to be solved for both the simpler and more complex cases.

The objectives from this Chapter are twofold. The first objective is to understand the relation between the players' decisions and the relevant variables in the game. Consequently, it is necessary to model the game to predict the decisions. The second objective is to simulate what happens with the distribution of the crime of corruption in the society under distinct scenarios. These simulations are also crucial for the next chapter's empirical objective.

2 Simple Corruption Game

The goal from this chapter is to solve the corruption game with non-trial resolutions. However, in order to smoothly introduce the the analytical framework. First, it is convenient to introduce a simpler game where agents are convicted when detected and they cannot self-report. From this simple example is possible to set most of the game requirements. Additionally, it is also possible to draw the first conclusions regarding the role of the bribe and the decision to enter in corruption.

2.1 Setting the Game

The necessary ingredients to set the game are presented here. Namely, a set of players and a set of states or stages where the system can be. Related to each estate, there are a set of possible actions or decisions. Furthermore, each pair of state and action has a pay-off that depends on the constant parameters from the game. Lastly, the game needs a timing protocol which describes how and when the players are allowed to play. These necessary features are better presented in the following subsections.

2.1.1 Players

The game here is played by two players[2]. A bribe payer and a bribe receiver $i \in [payer, receiver]$.

2.1.2 Timing and Information

The game has discrete time t. Here, it is assumed that in each period t there is one corruption opportunity. Or, conversely, t can be understood as a constant time between each corruption opportunity.

As defined in the second chapter, the players here play simultaneously. This design seems more realistic than a sequential game [3].

[2] As stated in Chapter 2, there might be differences between individuals and corporations. Whenever relevant, the differences are going to be explored.

[3] This would logically imply in a take-it-or-leave-it proposition of the bribe. Which seems also not realistic.

The design of the game imposes how information is available for players. The first simpler example has naturally more information. As the game gains complexity, agents are not able to fully observe the relevant information in the game.

For now, agents can perfectly observe each other's actions, states and, consequentially, the pay-offs. Therefore, the information here is said to be complete. Moreover, the present states contain all past information (Markovian Property)[4].

2.1.3 Parameters

The relevant parameters for this first examples are the price of the bribe b^5,the advantage from corruption a, the cost for the *payer* to perform the bribe c_b, the sanctions s that players pay if detected by the authority and the probability of being detected by the authorities α. Note that, for this first example, the sanction s is given by a fine f^6 and everything that the agents have gained from corruption (a for the *payer* and b for the *receiver*).

It is possible to state that this a common criminal structure. Therefore, the conclusions made here can be applied to other types of crime. Notably, this work resembles a lot the structure used in leniency studies for anti-trust offences.

2.1.4 States

Any position in a game (or system) can be summarized as a state variable. Or else, in an extensive game the states are random variables that indicate each possible node in the network.

In this example, the set of states S are the 'states of the world' where agents can be[7]. For instance, if agents decide not to bribe, they are in a state of '*no corruption*'. As well as if they decide to enter in bribery, then they can be in two other states. If corruption is well succeeded, agents gain their pay-offs from corruption in the state '*colluding*'. Or, they can be detected by the authorities and fined in the state '*detected*'. The Figure 1, shows the states and how agents can go from one stage to the other. Lastly, the laws that describe

[4]For any dependent X_t in x_t then $p\left(X_n = x_n | X_{n-1} = x_{n-1}, ..., X_0 = x_0\right) = p\left(X_n = x_n | X_{n-1} = x_{n-1}\right)$. In words, future events depend only on the present set of states. Or else, all the past information is embedded in the current states.

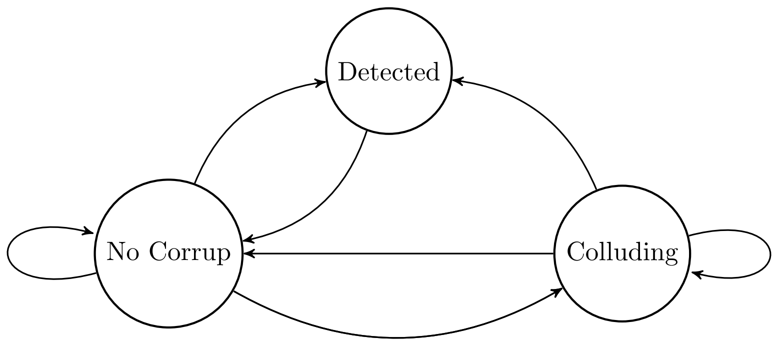
[5]It is possible to make the bribe b as a decision for the *receiver* and not a constant. The next section solves the corruption game for a random continuous bribe and for a binary decision over a constant value for b.

[6]In many jurisdictions there are non-monetary sanctions both for individuals (imprisonment) and corporations (licence and activity impediments). As Becker (1968) points out, they are complementary and necessary to optimal deterrence. Here, for the moment, one can assume that agents are able to translate the non-monetary fines in the value of f. Nonetheless, the argument on non-monetary fines is further developed in the following sections.

[7]States can be very subjective, and they depend on the analyst design of the game. The same game can be designed using different set of states. Moreover, some solutions (such as Linear Quadratic) require the use of states that are very unintuitive. As Ljungqvist and Sargent (2012) state, finding the right state is an art.

how agents go from one state to the other depend on the actions that they take and are better described bellow.

Figure 1: Game States and Transition Rules



The arrows represent the transition from the states[8]. It is given by a rule or 'law of motion' which depends on the state of the players and their actions. These concepts are better explained below.

2.1.5 Actions

Each possible state has a set of allowed actions d for each player i. In this first simple example the players are not allowed to report their misconducts. Nonetheless, this simpler version shows some expected relations between variables and the decisions towards bribing.

The *payer* can decide whether or not to pay $d_{payer} \in [pay, not\ pay]$ a bribe of value b to earn an advantage a if they are not detected[9]. The *receiver* decides to accept it or not $d_{receiver} \in [accept, not\ accept]$ at a cost of c_b. So that $d = [d_{payer}, d_{receiver}]$.

2.1.6 Laws of Motion

As shown in the Figure 1, agents go together from one state to the other. Since the game is stochastic, the next state S_{t+1} is function $S(.)$ of the actions taken d_t and the current state S_t given some probability of detection α.

$$S_{t+1} = S((S_t, \epsilon)|d_t) \tag{1}$$

Where $\epsilon \sim$ *Bernoulli* (α).

In other words, each arrow in Figure 1 happens depending on the actions taken in that state and the probability associated to that combination[10].

[8]For now, agents can be in only one state at a time. In this example, they would always go together to each state at each time t.

[9]For now, if agents are detected, they cannot pay or receive a bribe in that state.

[10]For instance, the arrows from Figure 1 that represents staying in the '*no corruption*' state,

2.1.7 Pay-offs, Costs and Rewards

In order to draw the player's decision rules in a similar fashion. It is possible to write the pay-offs in terms of costs ϕ and benefits π from corruption to each player i. Let then $\pi_{payer} = a$ and, $\pi_{receiver} = b$. Also, $\phi_{payer} = b$ and, $\phi_{receiver} = c_b$. The notion of costs and benefits extends also to the decision of not entering in bribes, where ϕ_i and π_i assume the value of zero.

Each state S_t has an outcome or pay-off associated to it. Lets call the rewards y so that $y_{i,t} = y(S_t)$, or more specifically:

$$y_{i,t}(S_t) = \begin{cases} 0 \text{ if not colluding} \\ \pi_i \text{ if colluding} \\ -f \text{ if detected} \end{cases} \tag{2}$$

As discussed in Chapter 2, rational individuals are trying to maximize some utility function. While firms may pursue the maximization of profits. For now, let the pay-off be the relevant objective from players. In the next section, the budget set and the notion of consumption is added to the agent's problem.

2.1.8 Expected Returns

Expected return from corruption can be written the costs of corruption ϕ_t plus the expected value of all possible rewards in the next period y_{t+1} weighted by a time discount γ, or else:

$$E[y] = -\phi + \gamma \sum_S E\left[(\pi - s)\right],$$

Since s is equal to zero if not detected but equal to f if detected. Therefore, in this example, the cost of corruption $-\phi$ and the pay-offs from corruption y to each player i weighted by its expected probabilities can be written as

$$E[y_i] = -\phi_i + \gamma\left[(1-\alpha)\pi_i - \alpha f\right]. \tag{3}$$

2.2 Solution and Equilibria

Firstly, the algebraic analytical solutions are enough to understand the equilibria. However, as the game acquires more complexity, the equilibrium concept is narrowed to the Markovian perfect equilibria (MPE), which also requires numerical solutions.

There are multiple ways to solve this game. It is possible to solve the game by taking b as constant and solving for the best d_t in a one shot or repeated

---

or going to the '*Colluding*' state can be written respectively as

$p(S_{t+1} = No\,Corruption \mid S = No\,Corruption,\ d_{t,payer} = pay,\ d_{t,receiver} = not\ accept) = 1,$ and

$p(S_{t+1} = colluding \mid S = No\ Corruption\ ,\ d_{t,payer} = pay,\ d_{t,receiver} = accept) = (1-\alpha).$

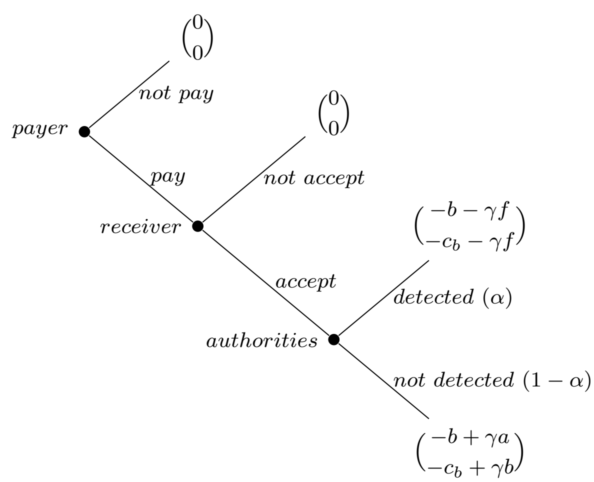
iterations. Note that in this setting, the timing of the decision is unimportant, since corruption only happens when both parties agree[11]. On the other hand, it is possible to solve the game for the best bribe b. Where, b is the relevant action to be chosen b the bribe payer.

2.2.1 Solving for a Given Bribe

If the value of the bribe b is given. Then agents must decide if they pay $d_{payer} = pay$ or accept $d_{receiver} = accept$ the bribe. In this sense[12], the solution is straightforward. If b lies in any profitable interval[13], then players choose to enter in bribe, otherwise they do not.

It is possible to solve this game with backward induction. The Figure 2 shows the decision tree and the pay-offs (above for the *payer* and below for the *receiver*) associated to each decision:

Figure 2: Simple Corruption Game Tree



By backward induction, the agents are going to enter in bribery if [14]

$$E[y_i] = -\phi_i + \gamma\left[(1-\alpha)\pi_i - \alpha f\right] > 0.$$

More specifically, for the *payer*,

[11] Buccirossi and Spagnolo (2006) explore the game in which players have pay a bribe and only after the receiver perform the favour. In this setting, there is a 'Hold-up' problem which can be fixed with enough repetitions or using leniencies as hostages.

[12] First order conditions for the equation to maximize (3) in b would lead to $\frac{\delta y_i}{\delta d_i} = y_i = 0$, since y_i is a function of d which is true whenever d_i is true.

[13] It is assumed that agents are neutral towards risk.

[14] It is assumed that agents weakly prefer not to enter in corruption. Therefore, it is possible to use the strictly less than or greater than 'preference'.

$$E[y_i] = -b + \gamma\left[(1-\alpha)a - \alpha f\right] > 0,$$

rearranging,

$$b < \gamma\left[(1-\alpha)a - \alpha f\right] \tag{4}$$

And for the *receiver*,

$$E[y_{receiver}] = -c_b + \gamma\left[(1-\alpha)b - \alpha f\right] > 0,$$

rearranging,

$$b > \frac{(\gamma\alpha f + c_b)}{\gamma(1-\alpha)} \tag{5}$$

The equations (4) and (5) show that there will always be corruption as long as the bribe lies in the interval $\left(\gamma\left[(1-\alpha)a - \alpha f\right], \frac{(\gamma\alpha f + c_b)}{\gamma(1-\alpha)}\right)$. Importantly, it does not matter who proposes to enter in bribery first.

2.2.2 Choosing the Bribe

If agents can choose the bribe, then the game design is naturally sequential. Because, one party must come up first and propose to enter in corruption. In this case, agents will choose the bribe that maximises their pay-offs. Or else, the bribe *payer* pays a bribe b which is slightly bigger than $\gamma\left[(1-\alpha)a - \alpha f\right]$ the *receiver* asks for a bribe b bigger enough then $\frac{(\gamma\alpha f + c_b)}{\gamma(1-\alpha)}$ [15]. Importantly, depending in the standard deviation σ[16], the utility from the uncertain returns from corruption vary according to the other player's relative risk aversion.

2.3 Final Remarks

Note that repetition is unimportant to this game equilibria. Here, even if the allowed actions depend on the player's state, they do not change as the stages of the game repeat. For instance, if an attempt to bribe fails and players are detected, in the next stage the decision rule is the same as if they had succeeded. Therefore, this specific design leads to the one-shot game equilibrium being equal to the same subgame perfect equilibrium (SPE) in each stage.

3 Corruption Game with Non-Trial Resolutions

In this section the game is expanded to encompass the non-trial resolutions (NTR). For this work, NTRs are summarized as the possibility of agreement

[15] Notably, since b is constant, its value is bigger than $E[y_i] = b$ for a risk averse agent.

[16] Which can be extracted from the binomial distribution $\sigma = nalpha(1 - alpha)$, where n is the number times that one repeats the event.

between offenders and judicial/prosecutorial authorities to avoid a trial[17]. The agreements offer judicial benefits to the agents and can only happen if the they agree to self-report and disclose their misconduct.

It is important to remember that this work focuses in the ex-ante decisions. The ex-post benefits such as facilitation of prosecutions, costless judicial decisions or screening effects are not discussed here[18].

3.1 New Setting

As a consequence of the new setting the current model must account for three new features. First, there must be a decision to self-report in the action-space d. Second, the detection and trial phases must be separated in the state-space S. Lastly, the new states must have a new set of pay-offs in y that account for the sanction reductions. Furthermore, the model must include a probability of being convicted β in a trial after being detected.

Most importantly, the model must now incorporate a set of distinct sanctions s for each type of conviction/agreement. Each agreement has a rule for reducing the fine f, such that. If agents unilaterally self-report before detections they receive a reduced fine of Rf, where $R < 1$. Moreover, if both *payer* and *receiver* self-report before being detected the fine reduction is rf, where $1 > r > R$. However, if the agents are detected, they can plea-guilty, in this case, the fine reduction is lower, or Pf if for unilateral reporting and pf for simultaneous reporting, where $1 > p > P$. In summary, the fine reductions are bigger (more lenient) for unilateral self-reporting than they are for simultaneous reporting. In the same way, fine reductions are more lenient to people that self-report before than after being detected.

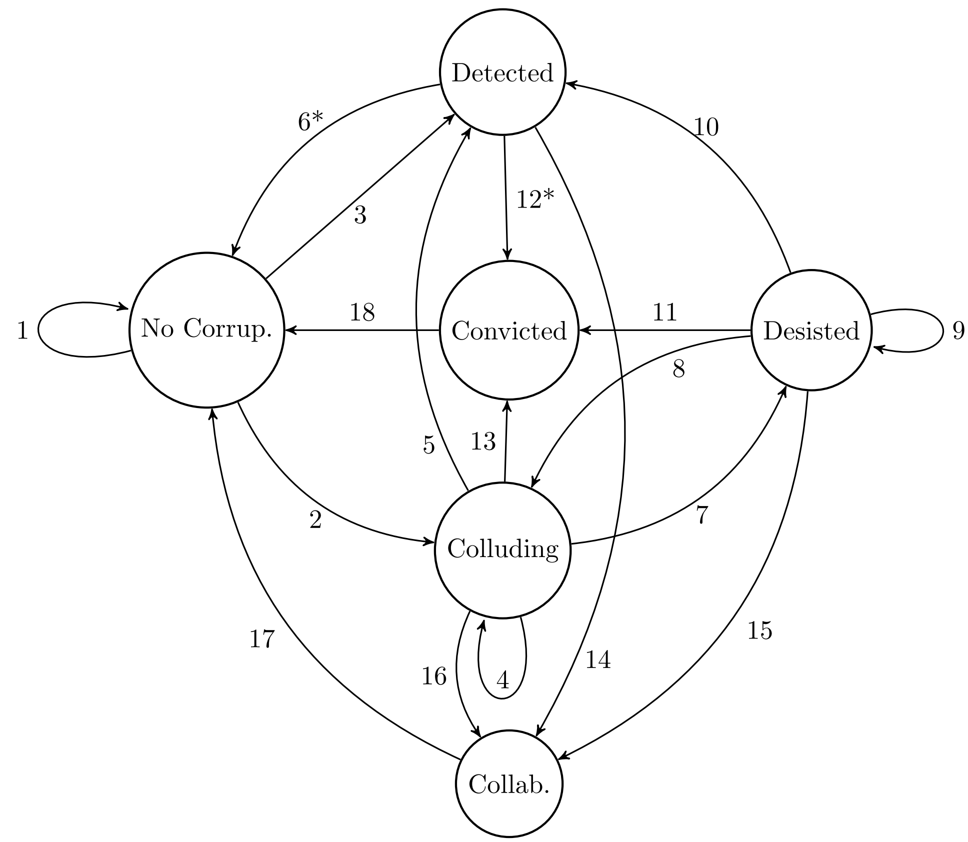
3.1.1 States and Actions

In order to fully describe the new set of possible states of the world, a series of new states in S must be introduced. The Figure 3 shows the new flow of states in the game.

[17] This is an oversimplification of the institution. NTRs can have a variety of distinct features depending on the jurisdiction.

[18] This branch of literature is better explored in Chapter 3.

Figure 3: States of the World with Non-Trial Resolutions



** The flow is represented by two different sets of states and actions.*

In a few words, if agents initially decide to pay a bribe, they can be detected (3) and go to the *detected* state or succeed (2) and go to the *colluding* state. From the latter, if they bribe again they can be detected (5) or go back to the same state (4). However, if they decide not to pay a bribe, they might be detected by their past crimes (5) or, if not detected (7), go to the state *desisted*[19]. The state *collab* stands for the states where the agents self-report (14, 15 and 16) and are collaborating with authorities. The *convicted* state is where agents pay the full fine. Admittedly, agents are convicted if they are reported (11, 12* and 13). Lastly, agents that are the detected but do not report are either convicted (12) or acquitted (6*) in trials. After, this, they go back to the beginning of the game (6*, 17 and 18).

Importantly, differently from the past example, agents do not go together to all states, they can be in distinct states at the same time. Therefore, there is a

[19]The *desisted* state is more meaningful in the next examples. For now, it is just a state where agents did not bribe in the past state, but have entered in bribery in any state before.

state $S_i \in S$ for each player i.

Note that the transition law is still a function of the present state S_t, the decision d_t and a stochastic probability. However, now, there are two distinct stochastic events which can happen with different probabilities[20] depending on the state S_t that agents are in.

$$S_{t+1} = S((S_t, \epsilon(S_{t+1}))|d_t) \tag{6}$$

Where $\epsilon(S_t)$ follows $\sim$ *Bernoulli* (α) or $\sim$ *Bernoulli* (β) depending on the states and actions.

3.1.2 Sanctions and Pay-offs

Given the new set of states[21], there are a new set of rewards $y(S_t)$ from each state in S, such that

[20] The transition rules are exhausted bellow, where the superscript $'$ represents the next period $(t+1)$:

1. $p(x' = s_{nc}|x = s_{nc}, d_i = 0 \; or \; d_j = 0) = 1$
2. $p(x' = s_{cor}|x = s_{nc}, d_i = 1 \; and \; dj_1) = 1 - \alpha$
3. $p(x' = s_{det}|x = s_{nc}, d_i = 1 \; and \; d_j = 1) = \alpha$
4. $p(x' = s_{cor}|x = s_{corr}, d_i = 1 \; and \; d_j = 1) = 1 - \alpha$
5. $p(x' = s_{det}|x = s_{nc}, d_i = 1 \; and \; d_j = 1) = \alpha$
6. $p(x' = s_{acq}|x = s_{cor}, d_i = 0 \; and \; d_j = 0) = 1 - \beta$ and $p(x' = s_{nc}|x = s_{acq}) = 1$
7. $p(x' = s_{des}|x = s_{cor}, d_i = 0 \; and \; d_j = 0) = 1 - \alpha$
8. $p(x' = s_{des}|x = s_{cor}, d_i = 1 \; and \; d_j = 1) = 1 - \alpha$
9. $p(x' = s_{des}|x = s_{des}, d_i = 0 \; and \; d_j = 0) = 1 - \alpha$
10. $p(x' = s_{det}|x = s_{cor}, d_i = 1 \; and \; dj_{=}1) = \alpha$
11. $p(x' = s_{des}|x = s_{con}, d_i \neq 2 \; and \; d_j = 2) = \alpha$
12. $p(x' = s_{con}|x = s_{det}, d_i \neq 2 \; and \; d_j \neq 2) = \beta$ and $p(x' = s_{con}|x = s_{det}, d_i \neq 2 \; and \; d_j = 2) = 1$
13. $p(x' = s_{con}|x = s_{cor}, d_i \neq 2 \; and \; d_j = 2) = 1$
14. $p(x' = s_{col}|x = s_{det}, d_i = 2) = 1$
15. $p(x' = s_{col}|x = s_{des}, d_i = 2) = 1$
16. $p(x' = s_{col}|x = s_{cor}, d_i = 2) = 1$
17. $p(x' = s_{nc}|x = s_{col}) = 1$
18. $p(x' = s_{nc}|x = s_{con}) = 1$

[21] There are hidden states within the state *collab*. They will tell if the agents reported unilaterally or simultaneously.

$$y_{i,t}(S_t) = \begin{cases} 0 \text{ if not colluding} \\ \pi_i \text{ if colluding} \\ 0 \text{ if desisted} \\ f \text{ if convicted} \\ 0 \text{ if acquitted} \\ Rf \text{ if reported alone before detection} \\ rf \text{ if reported simultaneously before detection} \\ Pf \text{ if reported alone after detection} \\ pf \text{ if reported simultaneously after detection} \end{cases} . \quad (7)$$

For instance, if a *payer* reports a bribery to the authorities and the *receiver* does nothing, then the *payer* is convicted and pay a sanction $s = RF$ while the *receiver* is convicted with the full fine $s = f$. However, if the *receiver* simultaneously report, both agents receive a sanctions of $s = rf$. Note that the sanction reduction in the first case is smaller than in the latter. Likewise, agents can report after detection. In this case, the relation between P and p follows the same logic.

3.1.3 Expected Returns

It is possible to decompose the pay-offs by summing up all distinct possible pay-offs. First the agent pays the cost ϕ_i of entering in a bribery scheme. In the next period, there is a chance of α of being detected by the authorities, if detected a chance of β of being convicted. Therefore, the expected return in case of being fined is

$$\gamma(\alpha\left[\gamma(\beta f)\right],$$

or simply

$$\gamma^2\alpha\beta f.$$

There are now two possibilities of earning the corruption gains π_i. Either by not being detected in the next period

$$\gamma(1-\alpha)\pi_i.$$

Or by being detected and not convicted

$$\gamma\alpha\left[\gamma(1-\beta)\pi_i\right],$$

or rearranging,

$$\gamma^2(\alpha(1-\beta)\pi).$$

It is possible to sum the possibilities in which the agents succeed in corruption and earn π_i at the end of two periods as:

$$\gamma^2(1-\alpha)\pi_i + \gamma^2(\alpha(1-\beta)\pi),$$

or simply,

$$\gamma^2(1-\alpha\beta)\pi_i.$$

Therefore, it is possible to write the expected value as:

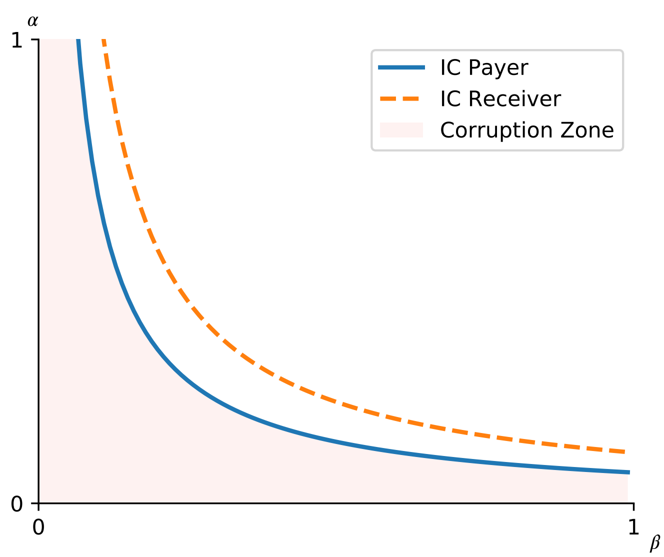
$$E[y_i] = -\phi_i + \gamma^2 \left[(1-\alpha\beta)\pi_i - \alpha\beta f\right]. \tag{8}$$

In the absence of non-trial resolutions, it is possible to calculate the domain in which bribes are profitable, ore else

$$E[y_i] = -\phi_i + \gamma^2 \left[(1-\alpha\beta)\pi_i - \alpha\beta f\right] > 0. \tag{9}$$

The Figure 4 show all the combinations in which $E[y_i]$ is positive for both players.

Figure 4: Indifference Curves for Profitable Bribes

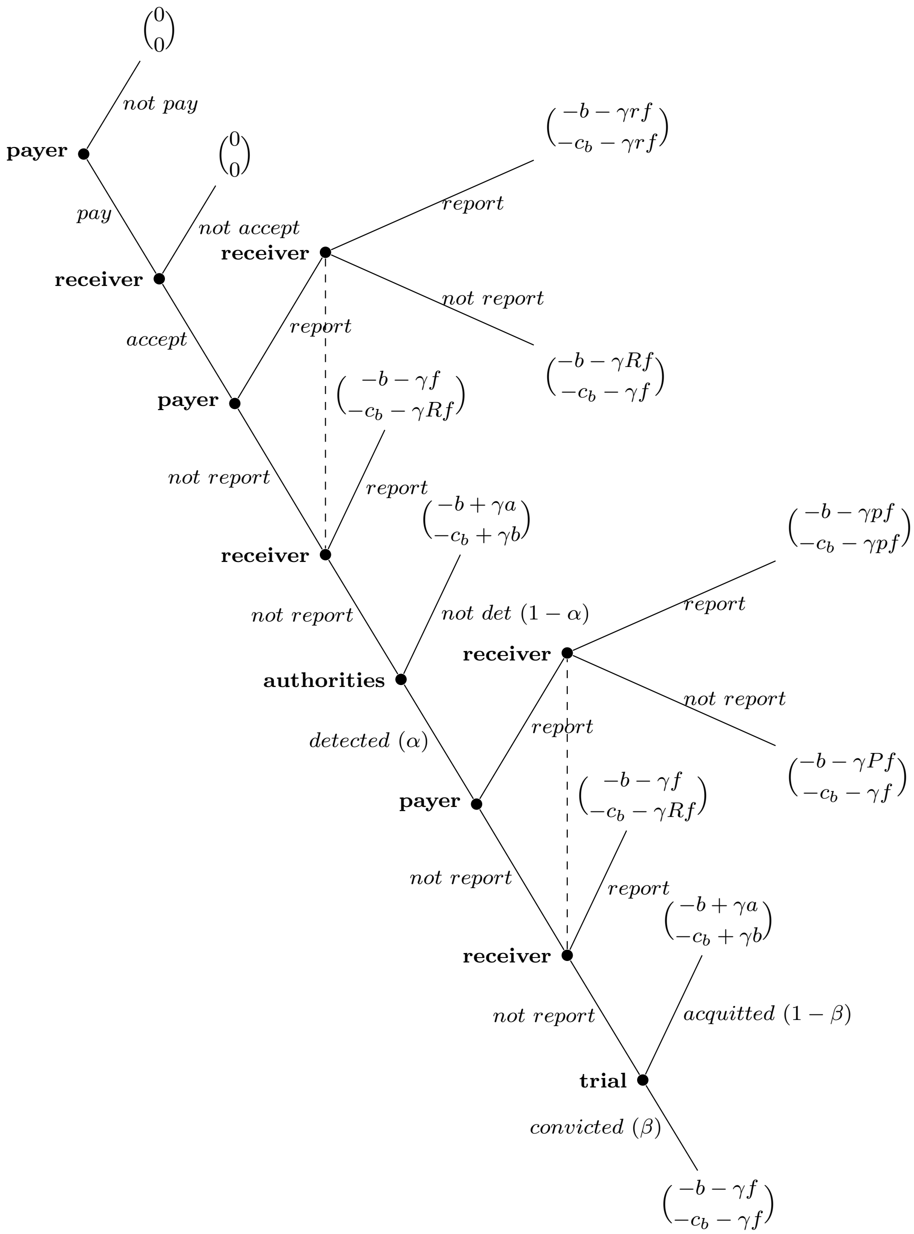


Note that corruption will occur only if corruption is profitable for both players. Given the above conditions, bribes are profitable if $\alpha < \frac{\pi_i - \phi_i}{\beta(\pi_i + f)}$ or $\beta < \frac{\pi_i - \phi_i}{\alpha(\pi_i + f)}$. Therefore, an increase in f moves both curves to the left, increasing the deterrent effect. Meanwhile, a decrease in $(\pi_i - \phi_i)$ moves the indifference curves towards a more deterrent set.

3.2 Solution and Equilibria

In this example, the solution to the problem is more complex. However, it can be solved analytically through backward induction. The Figure 5 below shows the extensive game-tree:

Figure 5: Corruption Game with Non-Trial Resolutions



Notes: The tree shows the players in bold; the actions at the edges of the tree's children; the pay-offs are in the parenthesis where the ones above are for the payer and below for the receiver. Lastly, the dashed line indicates that the players in those nodes and their parent are playing simultaneously.

In this setting there are two stages of simultaneous decisions. Therefore, it is necessary to calculate the distinct subgame perfect Nash equilibria (SPNE)[22] that orients the players' strategy.

Lastly, note that if the bribery conditions lie in the corruption zone from Figure 4, it does not mean that there will be bribery, because players still have take the possibility of being reported into consideration. In this sense, differently from the past example. This game has different outputs if played in one-shot (one time) or repeatedly. This happens because, if players stop collaborating after a defection, then, agents have to account the benefit of a one time defection against repeated gains from cooperation.

3.2.1 One Shot

Figure 5 shows that there are two nodes in which the players play simultaneously. In the first one, before detection, agents have to decide if they are going to report after paying the bribe. The matrix below shows the simultaneous decision payoffs for both players.

	Report	Not Report
Report	$-rf; -rf$	$-Rf; -f$
Not Report	$-f; -Rf$	$\gamma^2 \left[(1-\alpha\beta)a - \alpha\beta f\right]; \gamma^2 \left[(1-\alpha\beta)b - \alpha\beta f\right]$

Assuming that the agents do not want to be criminals, or else, not be in corruption. Then they would prefer a reduced fine Rf at least as good as the expected return from corruption at that node. Therefore, agents would only report if

$$-Rf \geq \gamma^2 \left[(1-\alpha\beta)\pi_j - \alpha\beta f\right],$$

where the subscrpit j represents the other player. Or else, rearranging

$$-R^* \geq \frac{f - \left[(1-\alpha\beta)\pi_j - \alpha\beta f\right]}{f}. \tag{10}$$

where the $*$ means the R in which agents prefer to report.

If the enforcement conditions α and β stay constant after the bribery and knowing that agents would have paid the costs of the bribe in the past node only if it was profitable for them. It would mean that reporting only happens if $R < 0$, or else, if agents gain a bonus from reporting. This is not common in most jurisdictions and this result is in line with past results using leniency to avoid cartels (Spagnolo, 2005).

If agents are detected, they need to decide upon another set of actions. In this case,

[22]The subgames are Nash equilibrium since players build their strategies non-cooperatively.

	Report	Not Report
Report	$-pf; -pf$	$-Pf; -f$
Not Report	$-f; -Pf$	$\gamma\left[(1-\beta)a - \beta f\right]; \gamma\left[(1-\beta)b - \beta f\right]$

Notably, the agents plea guilty if,

$$-Pf \geq \gamma\left[(1-\beta)\pi_j - \beta f\right]$$

or rearranging,

$$-P^* \geq \frac{\gamma\left[(1-\beta)\pi_j - \beta f\right]}{f} \tag{11}$$

Differently from the case of reporting before detection, the sanction reduction here does not strictly requires a bonus. Since the expected return after detection is smaller than before detection and it is probably negative. In this sense, it is possible to have plea-bargains with a $1 > P^* > 0$.

Note that, the first restriction (9) is bigger than the subsequent (10) and (11). Or else, there will be reporting or plea-bargain only if the probabilities α and β, or the rules of leniency change over time. Otherwise, the players would not enter in corruption at first[23].

Note that it is possible understand term on the right side of (10) and (11) as the discounted expected return from corruption[24]. So, let $\left(\frac{f+\pi}{f^2}\right)$ be constant of the return k, and setting the time discount $\gamma = 1$, then it is possible to rewrite (10) as,

$$(1 - R^*) \geq k(1 - \alpha\beta), \tag{12}$$

and, letting $k' \equiv \left(\frac{k-1}{k}\right)$, then

$$-P^* \geq k'\beta, \tag{13}$$

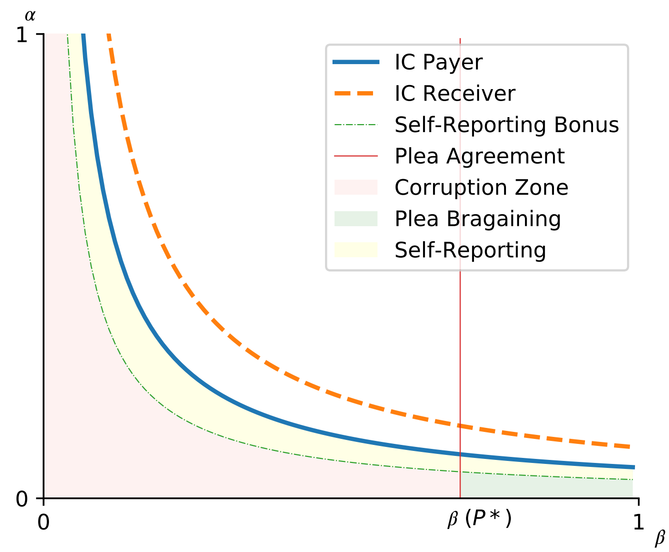
Now, it is possible to draw the separated equilibrium given the new boundaries (12) and (13). Therefore, for any $\alpha > \alpha(R*) = \frac{(R^*+k-1)}{\beta k}$ agents will not enter in corruption because of the fear of self-reporting [25]. Furthermore, for any $\beta > \beta(P*) = \frac{P^*}{k'}$, agents would plea guilty. The Figure 6 shows the indifference curves and the equilibrium areas.

[23] However, there are situations in which agents can enter a game just to explore the payoffs. This would happen if the bonus are big enough, so agents enter in corruption just to receive it. However, this is ratter an unrealistic situation

[24] In the first case for two periods and in the latter for one period.

[25] Once again, in an extreme situation, agents enter the game to self-report,Motta and Polo (2003) refer to this strategy as, CR, or collude and report. Here, this hypothesis is not considered.

Figure 6: Indifference Curves for Profitable Bribes with Non-Trial Resolution



The interpretation from the self-reporting area and the plea bargaining area is straightforward. In the first case, for a given leniency policy R^*, if the combined probability of detection and conviction is higher than $\alpha(R^*)$ and $\beta(R^*)$, it means that it is more profitable (in expected terms) to self-report and get the bonus than it is to stay in the game. Likewise, if the agents are detected, then they will always plea guilty if the observed probability of conviction β is higher then the threshold $\beta(P^*)$. Lastly, the thresholds $\alpha\beta(R^*)$ and $\beta(P^*)$ move towards the origin as R and P are smaller (more lenient). Therefore, more lenient sanction reduction rules from NTR decrease the non-observable corruption.

Lastly, it must be noted that the unilateral reductions rules R and P are the variables that disrupt the equilibrium[26]. However, the simultaneous reduction rule r and p are the ones actually paid by the reporting agents[27]. Therefore a policy recommendation that focus on deterrence from lowering the pay-offs or even if it targets revenue from prosecution would aim for a smaller (even negative) fine reduction for unilateral reporters. As well as a less lenient fine reduction for simultaneous reporters[28].

[26]This is a prisoner's dilemma.

[27]As long as $r < 1$ and $p < 1$, defection is going to be a Nash equilibrium. However, if $r > 1$ and $p > 1$, then there is no equilibrium to the game. Nonetheless, it does not mean that corruption is going to happen

[28]Spagnolo (2005) found that the first best solution would imply that the reporters receive all the gais from the other player as bonus. It would lead to complete deterrence. However, no revenues.

3.2.2 Repeated

In the repeated example, some assumptions are necessary to find the game's set of equilibria. Firstly, suppose that the game is infinitely repeated[29].Secondly, suppose that agents, if they are reported, they learn that the other player is not trustworthy, then they never engage in corruption again (grim trigger)[30][31].Therefore, the ex-ante decision to engage in corruption have to account for trading one-time defection against an infinitely repeated return from corruption[32]. This trade-off is expressed in the following matrix,

	Report	Not Report
Report	$rf; rf$	$Rf; f$
Not Report	$f; Rf$	$\frac{-b+\gamma^2\left[(1-\alpha\beta)a-\alpha\beta f\right]}{(1-\gamma)}; \frac{-c_b+\gamma^2\left[(1-\alpha\beta)b-\alpha\beta f\right]}{(1-\gamma)}$

This time, the agents will enter in corruption if

$$-RF \leq \frac{-\phi_i+\gamma^2\left[(1-\alpha\beta)\pi_i-\alpha\beta f\right]}{(1-\gamma)}. \tag{14}$$

Note that, the infinitely repeated return from corruption is always bigger then the one-time expected pay-off if the time discount $\gamma > 0$, or else

$$\frac{-\phi_i+\gamma^2\left[(1-\alpha\beta)\pi_i-\alpha\beta f\right]}{(1-\gamma)} > -\phi_i+\gamma^2\left[(1-\alpha\beta)\pi_i-\alpha\beta f\right].$$

Therefore, the fine reduction R^* must be more lenient[33] than it need to be in a one-shot game. Consequently, it is more difficult to combat corruption with leniencies in a repeated game. This results holds also when players are detected. The one-time plea bargain is punished by the grim-trigger, making the agents less prone to plea guilty.

Lastly, the difference between repeated corruption and one shot game is given by the size o γ. As $\gamma \to 0$, the difference between the two sets decrease.

3.3 Final Remarks

This section shows that leniency or fine reductions to corruption offenders who report their crimes may deter the crime from happening. The results found here replicate past findings from anti-trust literature (Motta and Polo, 2003; Spagnolo, 2005; Harrington, 2008; Chen and Harrington, 2007).

[29]Finite game would imply that the last game is played as a one-shot. By backward induction, it would make all sub-games played like a one-shot game.

[30]Grim trigger can also be understood as a punishment for defection.

[31]Here agents only learn to punish defectiong agents,they do not learn or update their preferences after each game. However, if players learn and update their preferences then, the the result of the game changes.This setting is explored in the next sections

[32]XXXXXXXXXXXX TO DO – Show the derivation XXXXXXXX.

[33]more positive in this case.

One notable outcome from this specific setting is that it shows that it is preferable to have lower fine reductions for players that unilaterally report but less lenient reductions if they do it simultaneously.

Note that the results hold under the assumption that agents are not learning or even updating any other relevant variable in their choices. In the following sections, some other variables that may evolve after repeated games are introduced. This setting make the model more realistic. However, the solutions are also more complex, and require different techniques to solve the agent's problem.

4 Additional States and Actions

Truly repeated games are very rare in realistic settings. As agents repeatedly interact, they gain (and accumulate) information. Therefore, it is possible to describe this additional information as states which change over time. Notably, the set of new states allows for a set of new actions.

4.1 Wealth and Consumption

Perhaps the most important feature missing in the previous example is the consumption. For individuals, a realist objective must account for the consumption decision c and not the pay-off itself. Note that, the decision of consumption requires a notion of a finite budget or wealth W, otherwise there would be nothing to decide upon.

4.1.1 The Consumption Utility Function

Agents may present distinct propensities towards risk. Therefore it is necessary to introduce the utility operator $u(.)$ which extracts the value of the consumed goods from players given their relative risk aversion η. In the next examples, it is assumed that the utility function $u(.)$ is an isoelastic utility function. Also called constant relative risk aversion (CRRA) function, such that:

$$u(c) = \begin{cases} \frac{c^{(1-\eta)}-1}{1-\eta} \text{ if } \eta \neq 1 \\ ln(y) \text{ if } \eta = 1 \end{cases}$$

Where η is the risk aversion parameter and $\eta > 0$ represents some degree of risk aversion.

In this setting, agents naturally search for a smoother consumption decision $c_{i,t}$.

Lastly, as discussed in chapter two, corporations and individuals might have distinct objectives. Notably, individuals are trying to maximize their utilities and corporations maximize profits. Therefore, the results from the following examples are more suitable to individuals. Admittedly, in the previous sections, the pay-off maximization strategy is more likely to portray a corporative decision. Nonetheless, corporative decisions are all made by individuals ultimately. Consequently, it is hard to attribute a set of decisions to any specific entity. Therefore, the differences between corporative and individual decision, whenever clearly spotted, are going to be pointed out.

4.1.2 Budget and the Asset Market

After each interaction between players, the agents either earn or lose some money from corruption. This may change the players willingness to enter again in collusion. Furthermore, even if they do not enter in a bribery, there are substitute options for player's budget[34]. Therefore, a fundamental state in the players decision making is the budget set.

Notably, the decision to consume $c_{i,t}$ is restricted to the agents budget constraint. It is possible to introduce it as the total wealth of players at each time $W_{i,t}$. Naturally, there must be a trade-off between consuming today and consuming tomorrow. This is given by the time discount γ. Consequentially, if agents choose not to consume today they need to earn some interest rates i_r[35] over their savings.

Note that agents might have multiple sources of income. Admittedly, there is an income from corruption y_c, a financial income y_f and a non-financial income or wage y_0[36]. So that, y_c is given by the expression in (7), y_0 is a constant autonomous income such as a wage and y_f is given by the following expression:

$$y_{fi,t+1} = (W_{i,t} - c_{i,t} - \phi_{i_t})(1 + i_r). \tag{15}$$

Note that 15 implies that the player will save all amount not consumed $c_{i,t}$ or spent bribery costs $\phi_{i,t}$ at time t and invest in an asset market.

4.1.3 Actions, States and Laws of Motion

It is possible to describe the state wealth $W_{i,t}$ as the sum of all player's incomes,

$$W_{i,t} = y_{fi,t} + y_{ci,t} + y_{0i,t} \tag{16}$$

Notably, in this setting, agents have a new allowed action to choose. They can now decide on how much to consume $c_{i,t}$[37]. This action determines how the state $W_{i,t}$ is going to evolve. Therefore

$$W_{i,t+1} = (W_{i,t} - c_{i,t} - \phi_{i_t})(1 + i_r) + y_{ci,t+1} + y_{0i,t+1} \tag{17}$$

[34] Without an asset market, there is no income and substitutions effects in play.

[35] Interest rates ir and time discount γ do not need to be equal.

[36] In this example, there is a living wage, so that agents can get enough funds to pay a bribe. This is a way to overcome the reducibility of the Markov chain. Without this artifice, agents who are convicted and lose everything would never be able to play again. Consequently, they were going to be stuck in that set of states.

[37] The decision can be on how much to save $(W_{i,t} - c_{i,t} - \phi_{i_t})$, since one decision is the complement of the other.

4.1.4 The Agent's Problem

Agents want to make the best set of decision d_i which maximize their utilities. Additionally, they are constrained to the other player's similar utility maximization problem. In this first simple example it is enough to present a simple maximization equation. However, in the following examples, as the complexity increases, it is convenient to introduce the value function equation.

Let γ be a time discount for both agents and E_0 is the expected value at $t = 0$, then the objective function can be written as:

$$\max_{c,d} E_0 \sum_{t=0}^{\infty} \gamma^t u(c_t) \tag{18}$$

$$\begin{aligned} \text{s.t.} \qquad S_{t+1} &= S((S, \epsilon)|d_t) \\ W_{i,t+1} &= (W_{i,t} - c_{i,t} - \phi_{i_t})(1 + i_r) + y_{ci,t+1} + y_{0i,t+1} \end{aligned}$$

Where the notation E_0 means the expected value at $t = 0$.

Note that, in this specific example, the decision of the player i depends on the decision of j[38] through the state transition rule. Therefore there is a strategic interdependence in the players' best decision, or else the decision of one player depends on the other $d_i = d(d_j)$.

4.1.5 Solution and Equilibria

Note that in the previous example, the expected return from corruption needs only to be greater than zero to be preferred in a subgame. Admittedly, now there are a set of new conditions for acceptance. First, agents can choose upon financial gains with certainty or stochastic returns from corruption. Notably here, the risk aversion plays a role. Secondly, agents can trade some instant pay-off for a better position in the game regarding all states. Given that, the current states are sufficient for agents to make their decision[39], it is possible to find a set of equilibria known as the Markov Perfect Equilibria (MPEs).

Before solving the problem in (18), it is convenient to introduce a new set of states that account for the judicial liability.

4.2 Judicial Liability

In the previous example and for most of the literature in leniency it is assumed that players are not liable for crimes committed in the past (Marvão and Spagnolo, 2016). In other words, if the crime is not detected immediately after it is committed, then agents are not liable for it in future stages. This is a convenient but unrealistic assumption about the game[40]. In order to overcome this

[38] Where j represents the other player.

[39] A stochastic process x_t is said to have the Markov property if for all $k \geq 1$ and all t , $Prob(x_{t+1}|x_t, x_{t-1}, ..., x_{t-k}) = Prob(x_{t+1}|x_t)$ (Ljungqvist and Sargent, 2012) p. 24.

[40] Also, without liability from past crimes, once in corruption, agents would never desist from entering in corruption again in a next period.

shortcoming, this section introduces the notion of judicial liability. Notably, it evolves as the agents play the game of corruption repeatedly.

4.2.1 Criminal Liability

Agents are liable for each crime they have committed. Each bribe paid generates a criminal liability l in the next period. Nonetheless, the liability from crimes tend to depreciate at a rate δ over time. Such that,

$$l_{t+1}(d) = \begin{cases} 1 \text{ if players agree on the bribe in } t \\ 0 \text{ if at least one player does not enter in corruption in } t \end{cases}$$

4.2.2 Liability Transition Rule

$$L_{i,t+1} = (1-\delta)L_{i,t} + l_{i,t+1} \tag{19}$$

There are a number of reasons to assume that the judicial liability decays over time. It may happen because information gets lost over time. Furthermore, prescription rules eventually cease the agents from culpability for old enough offences.

5 Complete Game

In this section, the corruption game with non-trial resolutions is solved in a more complex environment. Here, agents need to solve their consumption maximization problem given their choices to enter in corruption and face all implications from criminal law enforcements or make a living out of wages and financial markets.

At this point it is important to reorganize the notations so it is possible to write the problems more clearly. Following the tradition in the dynamic programming literature, the subscript $(t+1)$ is substituted by a *prime* $(')$ superscript. So, all the variables without a subscript (t) are referring to the present period. Additionally, lets define the vector containing all possible states as $\mathbf{x}$ such that $\mathbf{x}_i \supset [W_i, L_i, S_i]$. Similarly, the set of all actions as $\mathbf{d}$, such that $\mathbf{d}_i \supset [d_i, c_i]$. Lastly, θ is the vector containing all constant parameters. In summary, there are two vectors that contain all the states and actions,

$$\mathbf{d} \equiv \begin{bmatrix} d_i \\ c_i \\ d_j \\ c_j \end{bmatrix}, \text{ and } \mathbf{x} \equiv \begin{bmatrix} W_i \\ L_i \\ S_i \\ W_j \\ L_j \\ S_j \end{bmatrix}$$

Where, j is the other player.

5.1 The Agent's Problem

Both agents i face the same economic problem[41]. They want to maximize their utilities. Therefore, it is possible to rewrite (18) as

$$\max_{\mathbf{d}} E_0 \sum_{t=0}^{\infty} \gamma^t u(c_t(\mathbf{x})) \tag{21}$$

$$\text{s.t. } \mathbf{x}' = f(\theta, \mathbf{x}, \mathbf{d}, \epsilon')$$

The complete description of the constraint in (21), is given by the laws of motion (6), (17) and (19). Or else,

$$\begin{aligned} \text{s.t. } W' &= (W_i - c_i - \phi_i)(1 + i_r) + y'_{ci} + y'_{0i} \\ L' &= (1 - \delta)L_i + l'_i \\ S' &= S((S, \epsilon')|d) \end{aligned}$$

5.2 The Value Function

The agent's problem (21) has a recursive nature due its first-order difference equations constraints. Importantly, some future states are stochastically determined. Notably, this kind of stochastic processes are known as Markov decision processes[42]. In this sense, dynamic programming is the best methodological framework to solve this kind of problem (Putterman, 2005). In order to turn the problem into a dynamic programming problem, it is necessary to build the value function, which consists in the value of state vector to the agent $V(\mathbf{x})$.

Lets define a value function from the initial states $V(\mathbf{x_0})$ as being the the optimum value for the initial states W_0, L_0, and S_0. So that the value function $V(\mathbf{x})$ of the states can be defined by the following *Bellman Equation*,

[41]It is possible to look to the problem from another perspective. The government wants to maximize society's welfare. Assuming that there is no 'greasing the wheels' hypothesis from corruption (The hypothesis states that bribes work as shadow prices that increase efficiency of agents' transactions. Notably, it is more realistic when dealing with extortive bribes, when lower wages from civil servants are compensated by a bribe. Since, this work deals with collusive bribery, the hypothesis is less applicable.) and that the welfare gains from the criminals are neglectful. It is possible to infer that welfare is monotonically decreasing with the corruption level. Or else, governments want to minimize the level of corruption, or the number of agents in the state 'colluding' (lets call it S_c) over time, constrained to the agents maximization problem (23). For this, the government has the power to change the fines F and the leniency policies R and P. Therefore it is possible to write the governments problem as:

$$\min_{R,P,F} E_0 \sum_{t=0}^{\infty} \gamma^t S_{c_t}(R, P, F) \tag{20}$$

$$\text{s.t. } V(W, L, S) = \max\{u(c) + \gamma E[V(W', L', S')|W, L, S, \mathbf{d}, c]\}.$$

[42]It is also referred in engineering as stochastic control problems.

$$V(\mathbf{x}) = \max_{\mathbf{d}}\{u(\mathbf{x}) + \gamma E[V(\mathbf{x}')|\mathbf{x},\mathbf{d}]\} \tag{22}$$
$$\text{s.t. } \mathbf{x}' = f(\theta, \mathbf{x}, \mathbf{d}, \epsilon')$$

or,

$$V(W,L,S) = \max_{d,c}\{u(c) + \gamma E[V(W',L',S')|W,L,S,d,c]\} \tag{23}$$

$$\begin{aligned} \text{s.t. } W' &= (W - c - \phi)(1 + i_r) + y_c' + y_0' \\ L' &= (1-\delta)L + l' \\ S' &= S((S,\epsilon)|d) \end{aligned}$$

Therefore, the value function $V(.)$ takes the local maximum of the utility from consumption $u(c)$ today given all the possible consumption from all possible next states. This is an exhaustive exercise of finding the path that leads to the best expected results given all possible states.

5.3 Dynamic Programming Solution

The objective here is to find the strategies that are optimal for both players in an equilibrium. In order to solve dynamic programming problems the agents have to choose the best actions $\mathbf{d}$ that lead to a maximum value of current state $\mathbf{x}$ and future ones $\mathbf{x}'$ given a stochastic transition law Ω. Notably, the transition matrix $\Omega(\mathbf{x}', \mathbf{d}, \mathbf{x})$ can be understood as the probability density of the next state $\mathbf{x}'$, given the current state action pair $(\mathbf{x}, \mathbf{d})$[43]. Hence, the dynamic programming solution to the agent's problem, or the solution of the *Bellman Equation* in this case, can be achieved by simply iterating the value function given the transition matrix Ω up to a fixed point[44]. If the Ω matrix is stochastic, then the solution is given by maximizing each value function $V(\mathbf{x})$ given all states for each period, or else:

$$V(\mathbf{x}) = \max_{\mathbf{d}}\{u(\mathbf{x}) + \gamma \sum_{\mathbf{x}} V(\mathbf{x}')\Omega(\mathbf{x}', \mathbf{d}, \mathbf{x})\} \tag{24}$$

[43] It is possible to calculate the matrix $\Omega(\mathbf{x}', \mathbf{d}, \mathbf{x})$ for a given number of states n and actions m as:

$$\Omega(\mathbf{x}, \mathbf{u}, \mathbf{x}') = \begin{bmatrix} \begin{bmatrix} p(\mathbf{x}_0'|\mathbf{x}_0, \mathbf{d}_0) & \cdots & p(\mathbf{x}_0'|\mathbf{x}_0, \mathbf{d}_m) \\ \vdots & \ddots & \vdots \\ p(\mathbf{x}_0'|\mathbf{x}_n, \mathbf{d}_0) & \cdots & p(\mathbf{x}_0'|\mathbf{x}_n, \mathbf{d}_m) \end{bmatrix} \\ \vdots \\ \begin{bmatrix} p(\mathbf{x}_n'|\mathbf{x}_0, \mathbf{d}_0) & \cdots & p(\mathbf{x}_n'|\mathbf{x}_0, \mathbf{d}_m) \\ \vdots & \ddots & \vdots \\ p(\mathbf{x}_n'|\mathbf{x}_n, \mathbf{d}_0) & \cdots & p(\mathbf{x}_n'|\mathbf{x}_n, \mathbf{d}_m) \end{bmatrix} \end{bmatrix}.$$

[44] The existence of a fixed point is guaranteed for a $\gamma < 1$, which implies that the system is a Banach contraction (Putterman, 2005) (p. 149-151).

If V^* is a unique solution to the Bellman Equation (24)., then it is possible to define an optimal policy $\sigma(\mathbf{x})$ such that:

$$\sigma(\mathbf{x}) \in \text{argmax}_{\mathbf{d}}\{u(\mathbf{x}) + \gamma \sum_{\mathbf{x}} V^*(\mathbf{x}')\Omega(\mathbf{x}', \mathbf{d}, \mathbf{x})\} \tag{25}$$

The optimal policy σ_i is a vector with dimensions $(n \times 2)$, where n is the number of all possible combination of states. Consequently, for each row in σ_i there is an optimal pair of actions (d_i and c_i) that maximizes the value of all possible outcomes, not only for the next period, but for all other periods. In other words, the best optimum policy σ_i is a rule for finding not only the best outcome for the next period, but also to find the best position in the state space which leads to better outcomes.

5.3.1 Strategic Interdependence

Note that the individual decision function c and d, depends on S that depends on d again. This shows the strategic interdependence of the game, where some strategy of the player i depends on the other player j, or else substituting (6) in d_i, then

$$d_i = d(W_i, L_i, S(S_i', d_j(W_j, L_j, S_j, \theta)), \theta), \tag{26}$$

where d_i and $d_j \in [d_{payer}, d_{receiver}]$. Similarly, using the same logic c_i depends on c_j.

From this dependence it is clear to identify the strategic game from the problem. Therefore, it is possible to rewrite the policy function (24) in terms of the separate decisions of i and j,

$$\sigma_i(\mathbf{x}) \in \text{argmax}_{\mathbf{d_i},\mathbf{d_j}}\{u(\mathbf{x}) + \gamma \sum_{\mathbf{x}} V^*(\mathbf{x}')\Omega(\mathbf{x}', \mathbf{d_i}, \mathbf{d_j}, \mathbf{x})\} \tag{27}$$

In this way, it is clear to see that each player has a σ_i strategy that depends on the other player's strategy σ_j and vice versa. The equilibrium (MPE) arises when both players identify the other players optimal decision, given their own ones and still hold to their decisions (do not want to change their policy σ). Therefore, the solution to the agent's problem is a pair of optimal policy vectors (σ_i and σ_j) which maximizes the value function $V(\mathbf{x})$ for both players.

5.3.2 Discretization

Solutions to the continuous-states dynamic programming problems are known to be tricky[45]. One way to deal with this shortcoming is to discretize the problem, or else, turn the states and action variables into finite grids or lattices which agents can access[46].

[45] There are a few known solutions, most of them use a linear quadratic constraints. They are known to lead to Euler Equations and have nice analytical properties.

[46] The discretized variables are presented in the Appendix **??**.

It is important to point out that the discretization of the problem sometimes sacrifices the reality of the model. This happens because it is necessary to impose boundaries to the grids. Consequently, this finite lattice creates conditions to strange decisions near the boundaries. For instance, in this setting, the minimum wealth W that an agent can have is zero. So, if the agents see themselves in situations where the sum of the bad outcomes are greater than their wealth, they may act recklessly. This is a situation known as limited liability, and in a small grid it would occur more than in reality. The same thing can happen on the other extreme of the grid. If agents reach the maximum wealth, they have no incentive to invest in growing their assets, which also leads to strange decisions.

It is very clear that these grid limitations affect also the steady state from the game. Therefore, it is prudent to analyse the results excluding these extreme conditions. In the Section 5.4, the results for the society are sampled from the middle of the grid to avoid these problems.

5.3.3 Curse of Dimensionality

Discretization implies that we abandon traditional analytical mathematics for numerical computing. Notably, each discrete point in the state-action-space represents one point in the lattice of possibilities. Consequently, the number of possible points is obtained by the multiplication of all elements of each state and action. Therefore, it grows exponentially in the number of states or actions. This problem was called by Bellman and Dreyfus (1962) as the '*curse of dimensionality*'. Consequently, if the game is played in its full complexity, it is expected that the number of different combination in the state-action lattice to be excessively large.

For instance, in the complete game with only natural numbers as allowed states and action, with maximum wealth of $\overline{W} = 40$ and maximum liability of $\overline{L} = 5$, them the total number of point is equal to $(\overline{W}^4 \times \overline{L}^2 \times 8^2 \times 4^2 = 36.864 \times 10^9)$[47]. Therefore, it would require a lot of computational power to operate with matrices and vectors of length bigger than 36 billion elements.

5.3.4 Dimension Reduction

There are a couple of ways to avoid the course of dimensionality. Most of them are based on approximation of the necessary integrals (Pakes and McGuire, 2001).However, here, the solution to overcome this problem is to simply shorten the vector of states and action. The strategy is based in two main ingredients.

First, it is necessary to work only with feasible states. In other words, states that are impossible to reach do not need to be on the state-space[48]. For

[47]There are 4 spaces with length 40, the wealth space and the consumption space for both players. Then there are the liability spaces for both players. Also, the states of the world for both players, and lastly, the actions for both players.

[48]Importantly, the excluded spaces are only the ones that are impossible to be reached. The states that can be reached, but are eventually never be chosen still need to be on the space.

instance, if agents are not allowed to pay a bribe if they are detected, than the combination of state '*detected*' and the action '*pay the bribe*' is not accounted for[49].

Lastly, it is possible to work with incomplete informations. Meaning that, in a realistic setting, players are not able to fully observe each others states. For instance, it is unlikely that players can perfectly observe the other players budget set. Therefore, it is possible to use their own wealth as a parameter of the other player's budget.

There are a couple of additional good practices to eliminate some more excessive degrees of freedom in the model. For instance, find clever ways to structure the game such that the number of dimensions decrease by using additional feasibility rules[50]. Additionally, at each iteration the feasible states are recalculated. Unfortunately, it implies in distinct dimensions of the policy vector σ for each iteration[51].

5.3.5 Solution

It is possible to solve the agent's problem by iterating the policy function[52]. However, the optimum solution for one player, may not be the best for the other. In this sense, it is necessary to find the pair of strategies that solve the problem for both players[53].

The solution used here starts by finding the optimal policy σ_i given that j plays always a greed strategy. After, the problem is solved for σ_j using the σ_i from the previous result. By the definition of a MPE given by Maskin and Tirole (1988), the equilibrium is settled when both players observe the other player's best strategy σ but they do not change their strategies. Most importantly, if strategy start alternating, then the solution does not converge and the game has no equilibrium. The complete code containing the solution method and the results used in this work is available at: https://github.com/caxaxa/Corruption_Game.

5.4 Results and Equilibria XX PRELIMINARY XX

The results can be summarized in the best strategies or policy vectors σ_i that both players play to maximize their own objectives. These, policies depend on every parameter in the game. Therefore, there is one equilibrium (if it converges)

[49] I also conveniently implies that the Ω matrix is not sparse. In other words, all rows in the Ω sum up to one. Which makes the Markov chain easier to analyse. Consequently, it is easier to avoid problems such as reducibility in the Markov chain

[50] The set of all feasibility rules is available in the code

[51] If a decision rule is not in *sigma*, then the algorithm chooses a greedy decision.

[52] Putterman (2005) provides four methods to find the solution to discrete dynamic problems. Namely, value iteration, policy iteration, modified policy iteration and linear programming. Nonetheless, the routines used in this article can be found in Quantecon.org (Available at: https://python-advanced.quantecon.org/discrete_dp.html).

[53] Importantly, the agents must separately maximize their own objective functions given their own rewards. So, it is not enough to maximize a global objective function, since this solution would resemble a monopolistic solution rather then a Markov Perfect equilibrium.

to each combination of parameters. So, in order to systematically analyse the numerical results. First, it is necessary to choose a benchmark and analyse it. Then, it is possible to analyse the vicinity of the benchmark in order to observe how changes in variables affect the equilibrium.

5.4.1 The Benchmark

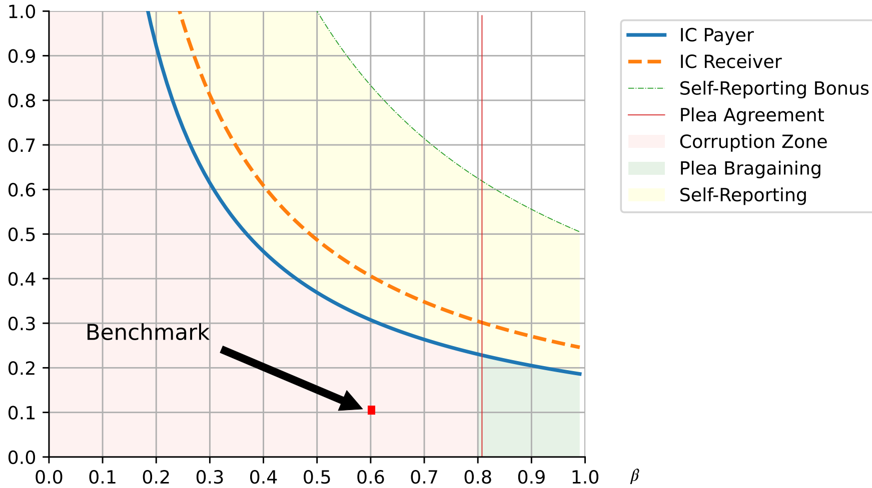
Let the benchmark be an arbitrary point in which players are expected to choose corruption, and not self-report or plea guilty. The Table 1 shows the chosen parameters for this case.

Table 1: Game Parameters

Var	Value	Meaning
		Bribery Parameters
a	5	Benefit from corruption
b	3	Bribe
c_b	1	Cost of corruption
f	5	Fine
		Time Discounts
γ	0.975	Time discount
i_r	0.05	Interest Rate
		Leniency Rules
R	0	Unilateral reporting before detection
r	0.5	Simultaneous reporting before detection
P	0.6	Unilateral reporting after detection
p	0.9	Simultaneous reporting after detection
		Leniency Rules
α		Probability of detection
β	0.5	Probability of conviction

The Figure 7 shows the profitable areas for corruption in and the indifference curves.

Figure 7: The Benchmark Indifference Curves

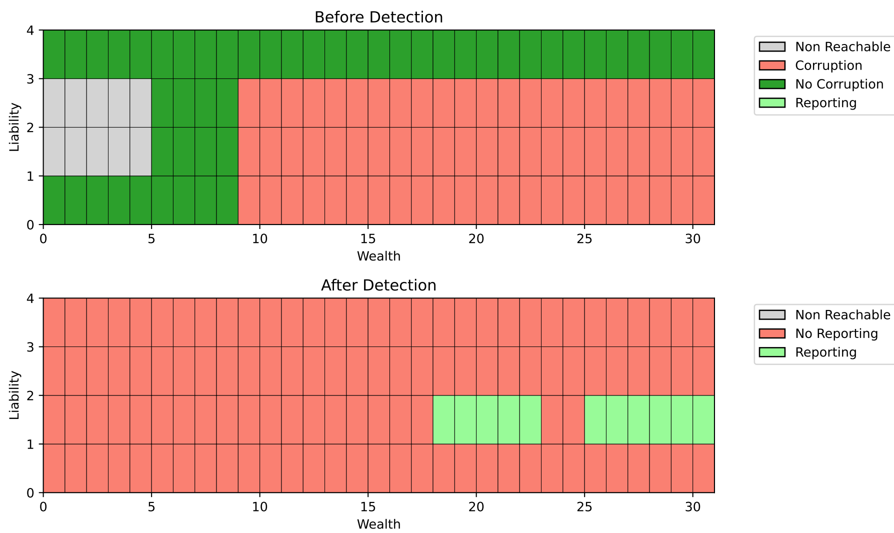


Individual Strategies

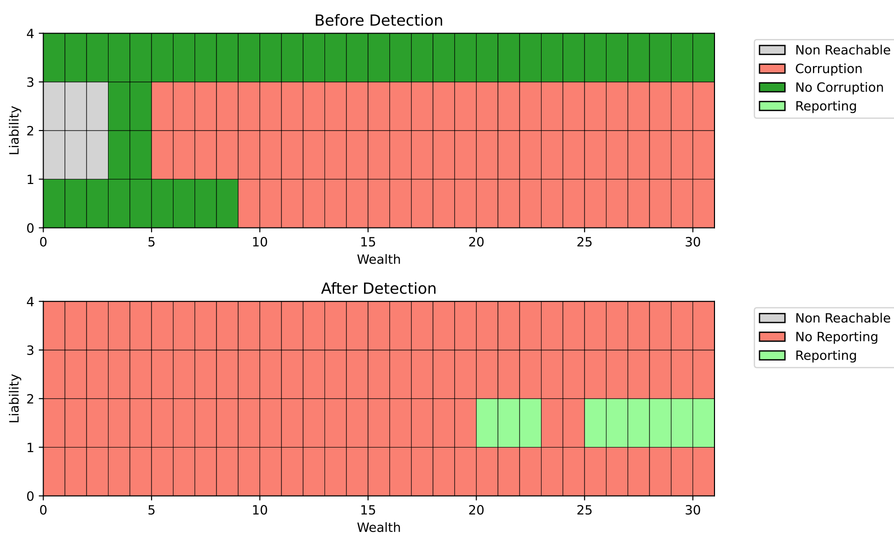
The players' optimal policy consists in a rule that tells the players how to play at each different point in the game. It is possible to visualize the best policy given the agent's current wealth and judicial liability. The Figure 8 shows all the combinations of wealth and liability and the player's decisions before and after detection for the parameters in Table 1.

Figure 8: Players' Optimal Policy

(a) Payer's Policy



(b) Receiver's Policy



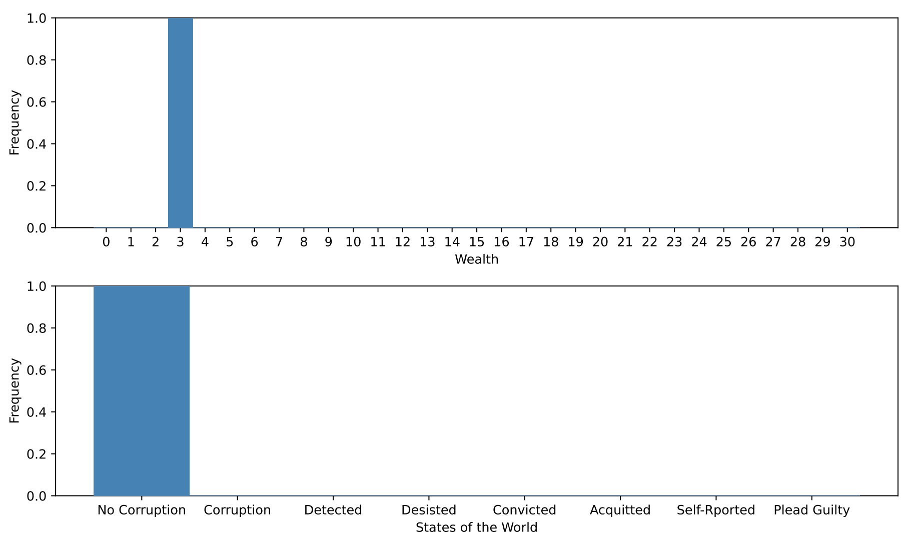
The first takeaway from the players' optimal policy is that the they are different between agents. Note that, even that Figure 7 shows that corruption would be preferred at that point, there are combinations of wealth and liability that would make the agents prefer not to enter in corruption. Moreover, it shows that agents, in that point would never prefer to plea guilty if detected. Therefore, since the players do not observe the other player's budget[54], they might be reported if they fail to assert the other player's wealth level. However, there are combinations of liability and wealth that would make them do. **Notably, there are no combinations that would make players self-report before being detected in this setting.**

Notably, it is necessary to carefully analyse the action choices near the beginning and the end of the grid. In these extreme points, agents may see themselves in special situations as pointed out in Section 5.3.2. The interesting thing is that, players at the beginning of the grid were expected to be more risky, since the fines can go only up to the maximum player's wealth. However, agents prefer not to enter in corruption in that region.

The end of the grid is more complex. Agents, cannot have more wealth than $\overline{W}$, so they might choose to save less, ore consume more. It is hard to say if it would change their decision towards entering in corruption. In this case, agents still enter in corruption.

One of the important characteristics from the Markov Perfect Equilibrium is that it is stationary. In other words, the set of optimal strategies drive the system to an stable equilibrium. The Figure 9 shows the stationary equilibrium.

Figure 9: Steady State Equilibrium



For the benchmark, the steady state equilibrium is the same for both play-

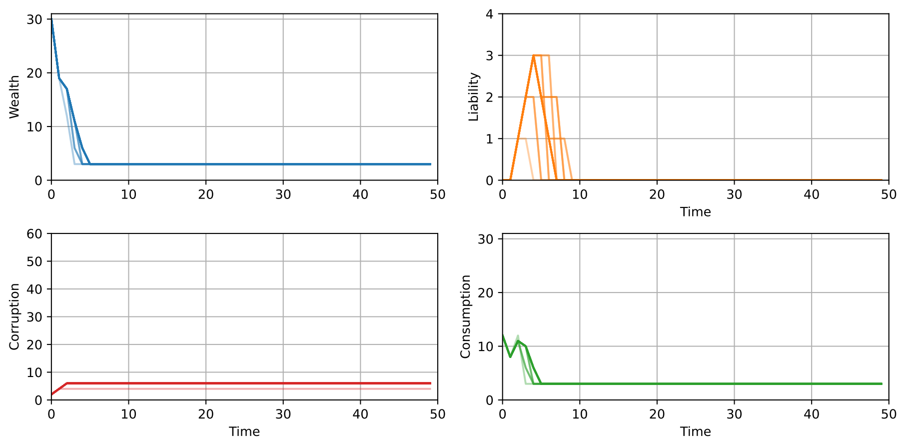
[54]By construction, they estimate the other player's wealth based on their own. In a complete game, it is possible to make the decision with players perfectly observing the other players budget set. However, it takes an entire new set of states, with length $\overline{W}$. Consequently, it requires a lot more computational power.

ers[55]. Note that, by construction, in order to avoid reducibility of the Markov Chain in a situation where agents would have no funds to pay a bribe or to save. The model gives the players a '*living wage*', so they will always have at least an income equal to the value of the bribe. In this case, the agents opted to live with the living wage, and consume everything at each state.

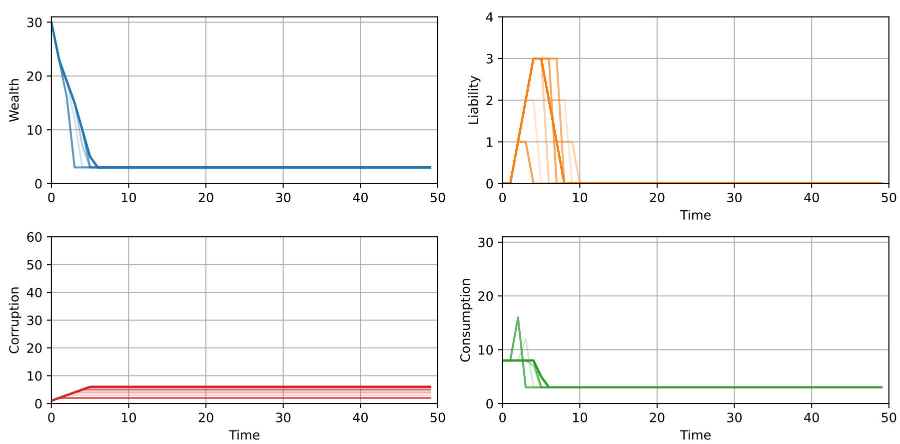
Notably, even if there is no corruption in the steady state, it does not mean that there wouldn't be corruption in the path to the steady state. It is possible to observe how the strategies lead to the steady state over time. The Figure 10 shows the path to the steady state.

Figure 10: Path to Steady State

(a) Payer's Path



(b) Receiver's Path



[55]It is not always the case

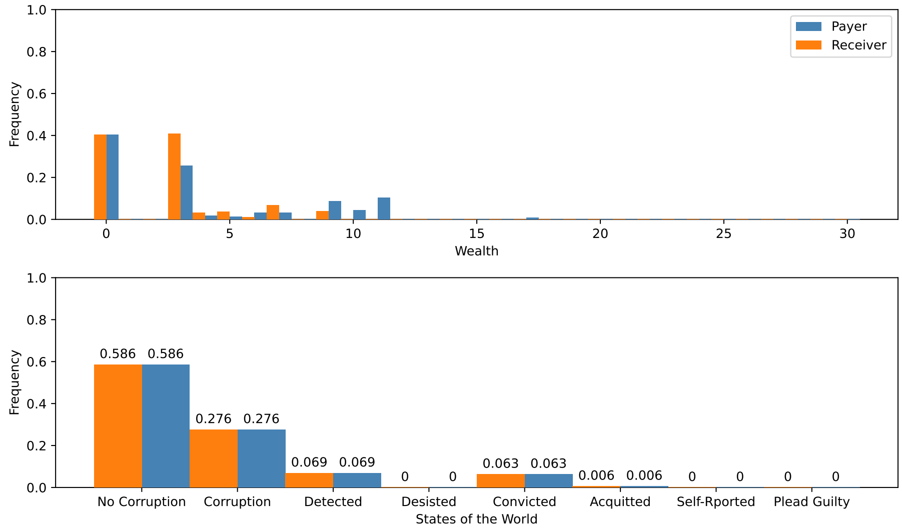
The Figure 10 shows how the agents alone would play this game for 50 periods. It also shows 50 stochastically different outcomes from the decisions. Therefore, in some cases, agents are detected and in others they succeed. The multiple paths are superposed so the thicker the line, the more probable is the outcome. It clearly shows that, when agents start with the maximum wealth $\overline{W}$ and then play the game, they choose some corruption in the way, but eventually they consume all the wealth and stay in the steady state indefinitely.

Note that Figure 10 shows the strategy of the players if they play alone and they accurately predict the other player's decision. Once again, this game has incomplete information about the other players budget. Therefore, the results observed if we put players to play against each other might differ substantially. Below, the results of the interaction between players is exposed as the distribution of decisions in the society.

Corruption Distribution in the Society

One of the goals from this chapter is to understand the distribution of the unobservable corruption in the society as the parameters from the game change. In order to obtain such distribution, it is possible to make both players play their best strategy against each other and sample from this interaction. To avoid the unwanted weird decisions on the extremes of the grids. Let the players' wealth be normally sampled from the wealth distribution, with mean 15 and standard deviation 3. So the majority of the sample is going to be centred in the middle of the grid. Also, the players are paired from a situation of '*no corruption*' and then interact for 5 rounds. After this, we keep the results and proceed with another sampling for 200 rounds. The Figure 11 shows the distribution of states on the society.

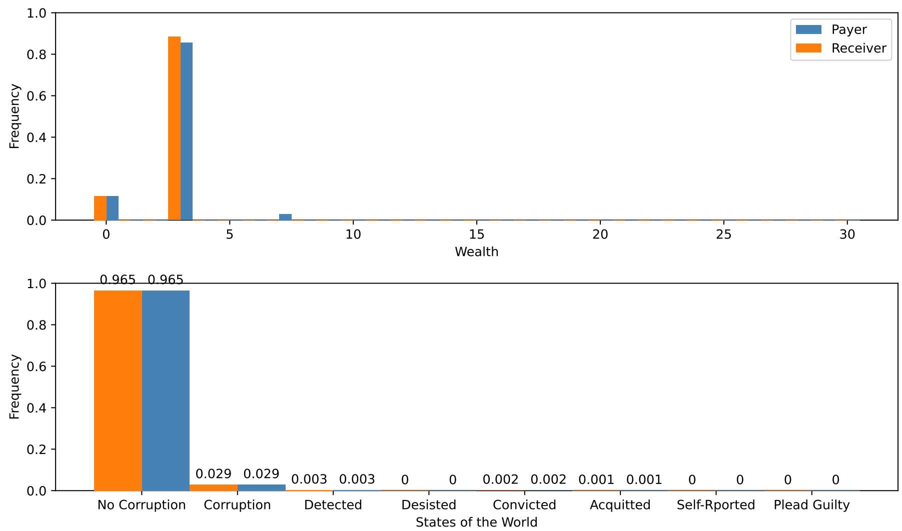
Figure 11: Distribution Sampling



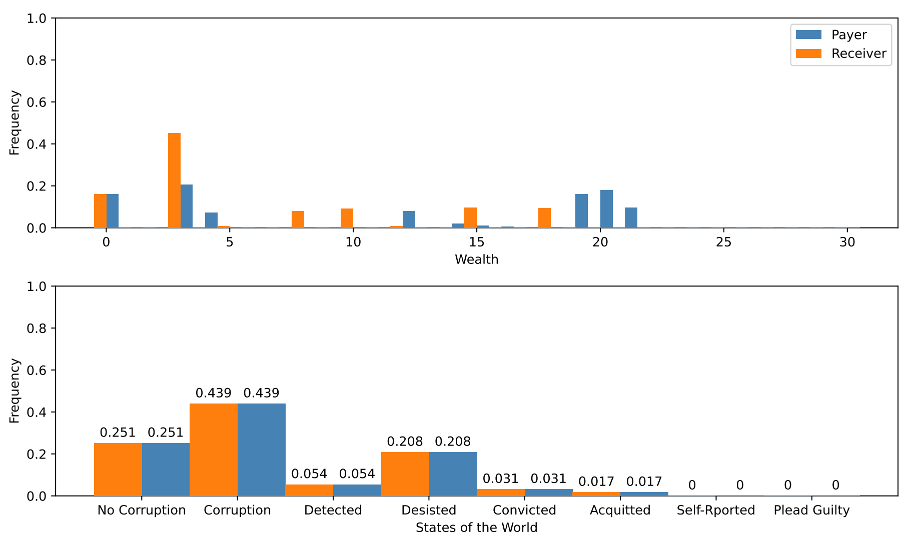
Note that, differently from the steady states, corruption is observed. Another interesting result may arise from testing a slightly different level of initial wealth. In other words, we can test the perseverance of corruption in a poorer or richer environment.

Figure 12: Distribution Sampling from Unevenly Wealthy Societies

(a) Poorer Society



(b) Richer Society



The Figure 12 shows that less wealthy society, tend not to choose corruption. While richer societies would. This results seems to be counter-intuitive. But it does shows that collusive corruption is a type of corruption that affects richer countries, while more extortive corruptions tend to be bigger in smaller countries(Søreide, 2018). The results follow directly from the optimal policies shown in Figure 8.

5.5 Conclusions XX TO DO XX

References

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

Bellman, R. and Dreyfus, S. (1962). *Applied Dynamic Programming*. Princeton University Press.

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

Chen, J. and Harrington, J. J. E. (2007). The impact of the corporate leniency program on cartel formation and the cartel price path. *Contribution to Economic Analysis*.

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

Ljungqvist, L. and Sargent, T. J. (2012). *Recursive Macroeconomic Theory*. The MIT Press Ser.

Marvão, C. and Spagnolo, G. (2016). Cartels and leniency : Taking stock of what we learnt. *SITE Working Paper, No. 39, Stockholm School of Economics, Stockholm Institute of Transition Economics (SITE)*.

Maskin, E. and Tirole, J. (1988). A theory of dynamic oligopoly, i: Overview and quantity competition with large fixed costs. *Econometrica*.

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

Pakes, A. and McGuire, P. (2001). Stochastic algorithms, symmetric markov perfect equilitbrium, and the curse of dimensionality. *Econometrica*.

Putterman, M. L. (2005). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. University of British Columbia, A JOHN WILEY AND SONS, INC., PUBLICATION.

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

Søreide, T. (2018). Corruption and criminal justice: Bridging economic and legal perspective. *The Research Handbook on Corporate Crime and Financial Misdealing - Edited by Jennifer Arlen*.

A Appendix

This appendix contains a short description of all variables, states and actions that contained in the model.

A.1 Constants θ

Given the following constants θ for all periods t and all agents $i \in (payer, receiver)$:
* Let b be the fixed price of the bribes;
* Let a be the gains from corruption to the payer;
* The cost from performing the corruption favour for the receiver is c_b ;
* Let f be the monetary fines for agents caught in corruption;
* i_r is the interest rate;
* δ is the liability depreciation or prescription;
* γ is the time discount;
* η is the agents' risk aversion;
* y_0 is an autonomous non-financial income;
* Probability of detection is α; and
* Probability of conviction is β.

A.1.1 Leniency rules

Regarding the rules of leniency, they can be:
* R sanction reduction for unilateral self-reporting before detection;
* r sanction reduction for simultaneously self-reporting before detection;
* P sanction reduction for unilateral self-reporting after detection; and
* p sanction reduction for simultaneously self-reporting after detection.

A.1.2 Leniency rules

A set of higher level notation to identify cost ϕ_i and benefits π_i of corruption to each agent i:
* $\pi_{payer} = a$;
* $\pi_{receiver} = b$;
* $\phi_{payer} = b$; and
* $\phi_{receiver} = c_b$.

A.2 A set of discrete actions or decisions (controls) $\mathbf{d}_{i,t}$

A.2.1 Action of Entering in Corruption d

The decision regarding corruption di, t can assume 3 different values:
* $not_{i,t}$ agents can choose to do nothing;
* $cor_{i,t}$ is the decision from player i to pay/receive a bribe in time t; and
* $rep_{i,t}$ is the decision from player i to self-report.

Agents decide how much they are going to consume at each time c_t. A complementary way to represent the consumpion decision, is the decision of how much to save w, or buy financial assets.

A.2.2 Action of Consuming or Saving c and w

* $w_{i,t}$ is the decision of how much to save, can go from 0 to (W_t); and

* $c_{i,t}$ is the decision of how much to consume, can go from 0 to (W_t).

Note that de decisions are complementary and equal to $W_{i_t} = c_{i,t} + w_{i,t} + \phi(d_{i,t})$.

A.3 A set of discrete states $\mathbf{x}_{i,t}$

A.3.1 State of the world S

* $S_{i,t}$ is the state of the world at time t;

* s_{nc} is the state of the world where neither players are in corruption;

* s_{cor} is the state of the world where both agents succeded in corruption;

* $s_{det,i}$ is the state of the world where the agent i is detected for corruption;

* $s_{con,i}$ is the state of the world where the agent i is convicted;

* $s_{acq,i}$ is the state of the world where the agent i is acquitted;

* $s_{sr,i}$ is the state of the world where the agent i self-reported before being detected;

* $s_{pg,i}$ is the state of the world where the agent i plead guilty;

and

* The states of the world $(s_{nc}, s_{cor}, s_{det,i}, s_{col,i}, s_{con,i}, s_{acq,i}, s_{sr,i}, s_{pg,i}) \in S$;

A.3.2 Wealth from Players W

* $W_{i,t}$ are the assets from player i at time t. It goes from 0 to a maximum of $\overline{W}$;

A.3.3 Liability from Players L

* $L_{i,t}$ is the wealth from player i at time t. It goes from 0 to a maximum of $\overline{L}$;

A.4 Discretization, State and Action Spaces

Let $\mathbf{x}$ be the entire state-space. If,

$W \in [0, ..., \bar{W}]$,

$L \in [0, ..., \bar{L}]$, and

$S \in [s_{nc}, s_{cor}, s_{des}, s_{det}, s_{con}, s_{acq}, s_{sr}, s_{pg}]$,

then,

$\mathbf{x} = [W, L, S]$ and has dimention $(3 \times (\bar{W} * \bar{L} * 8))$.

* Let $\mathbf{d}$ be the entire action or control-space. If,

$d \in [not_{i,t}, cor_{i,t}, rep_{i,t}]$ and

$w = [0, ..., \bar{W}]$, then

$\mathbf{d} = [w, d]$ and has dimension $(2 \times (\bar{W} * 3))$.