# A Dynamic Programming Approach for the Game of Corruption with Non-Trial Resolutions: The Brazilian Example

July 12, 2023

**Abstract**

Sanction reduction policies for self-reporters and collaborators may be effective in deterring collusive corruption. In order to test this hypothesis, this paper builds a theoretical model that solves the game of corruption with and without sanction reduction policies, using a dynamic programming setting to find equilibrium strategies. Results show that these policies may effectively deter bribery. Additionally, it provides a way to measure policy effectiveness over non-observable corruption crimes by utilizing the number of corruption crimes. The model predicts that a successful anti-corruption policy shock will lead to an increase in corruption detections, even if the probability of detection remains unchanged. The paper tests the introduction of the Brazilian anti-corruption policy of 2013, finding significant signs of immediate anti-corruption shock and evidence of further deterrence.

## 1 Introduction

Collusive corruption happens when two players, a bribe payer and a bribe receiver, agree to exploit some rent or contract. Naturally, when the bribe receiver is a public officer, society loses from this agreement. Notably, previous works on corruption represent this relation as a three-tier principal-agent problem (Burguet et al., 2016). Here, the principal (society) loses from an agreement between the agent (bribe payer) and the supervisor (bribe recipient).

Anti-bribery policies rely on many distinct features. One particularly well-defended strategy against corruption is to reduce the sanctions for criminals who collaborate with authorities to unveil bribery schemes (Nations, 2004)[1]. The idea is to break the agreement between payer and receiver. However, granting sanction reduction may reduce the expected fines and increase criminal activities. Previous literature points to the effectiveness of the policy (Buccirossi

[1]This work focuses on the case where the players are the collaborators. Some whistle-blowing policies rely on third parties to come up and denounce bribery schemes. This relationship implies a distinct game

and Spagnolo, 2006; Dufwenberg and Spagnolo, 2014). However, most studies fail to measure its impact on the corruption level. This work aims to answer the following open questions: Does a sanction reduction policy for collaborators deter corruption? Is it possible to measure the effect of the policy on the overall corruption level?

In order to address these questions, there are two main challenges. Firstly, it is necessary to understand the decision-making process of agents that leads them to engage in corruption in both collaborative and non-collaborative environments. This can be achieved by developing a corruption game model. Secondly, it is important to find a way to measure the level of corruption and determine the extent to which it changes when anti-bribery policies are implemented. If the model accurately predicts the level of corruption, it can be tested against real-world data from a specific case study.

The model used to predict the decision-making of players in entering into corruption is a game, where players must decide whether to enter into bribery or not and whether to defect on the agreement. The profitability of bribery is determined by the terms of the agreement and the advantage gained from it versus the costs. Effective anti-corruption policies can turn the game into a prisoner's dilemma, where defection becomes the dominant strategy, deterring corruption even when it is profitable. Sanction reduction policies can incentivize players to report their misconduct, resulting in both parties avoiding bribery. Admitting guilt can also lead to benefits in line with anti-bribery legislation. This type of analysis identifies potential areas of corruption based on game parameters.

Notably, any realistic model that tries to predict decisions in society must account for a utility maximization strategy. Sometimes, it is not only necessary to know if the game is profitable or not. It is required to know the rewards of choosing between corruption and a substitute investment. From this tradeoff, it is possible to infer the economic decision of entering into corruption. Notably, this kind of answer comes from solving a utility maximization problem. Furthermore, this type of game provides the analyst with more variables to test within the model, such as interest rates and risk aversion. These are necessary features to solve this kind of game, and they are also handy for empirically testing the predictions.

The specific objective of predicting changes in corruption due to a policy shock requires a dynamic model. Specifically, if the environment changes midstream, past decisions must be re-evaluated. Therefore, the model must carry information through time, which shows the consequences of past decisions. Consequently, the wealth, assets, and liabilities accumulated over time are crucial in the players' decision-making process. Therefore, the game is solved in a dynamic setting for a Markov Perfect equilibrium[2].

The last necessary characteristic of the model concerns the information set. Under the reasonable assumption that agents cannot perfectly observe the other

[2]Static games have the assumption of not allowing things to change over time. However, it is obvious that agents update their states as the game is repeated, making their decision more consistent with repetition.

player's wealth or criminal record, players infer the other party's states based on their own state. This strategy is crucial to economize state dimensions. Given that the game requires numerical solutions, this is very convenient to overcome the problem of the curse of dimensionality. Furthermore, it provides enough volatility and uncertainty to agents, making it possible to observe unilateral defections even under optimal strategies.

Solving the game with all necessary features is challenging due to multiple equilibria arising from too many degrees of freedom. Therefore, the model is solved for a benchmark configuration, such as a balanced power relation between the players, similar bribery profitability, and equal risk aversion. Results show that if bribery is profitable, it might happen, but reporting and plea agreements are also possible. The presence of sanction reduction policies decreases the expected corruption level, and lenient policies lead to less corruption in agents' decisions and society.

Once the model is specified, it is possible to make predictions about corruption that can be tested. However, it is first necessary to find a reliable way to measure corruption in a society. Measuring corruption is challenging as it is difficult to quantify the amount of money spent on bribes, the losses in public revenue, or the number of bribes paid. Additionally, only a portion of corruption is detected. Therefore, the detection of corruption is an essential parameter for measuring corruption in the economy. Fortunately, the proportion of detected corruption crimes is an output from the model.

The analysis of the detection of corruption over time can be misleading. Notably, an increase in corruption detection may happen because there was an increase in the number of crimes, or because there was an enhancement of the prosecution productivity. The contrary is also true. So, how is it possible to identify if changes in the detection of corruption come from higher criminal activities or from higher anti-bribery enforcement? Once again, the answer relies on the chosen model. It separates enforcement variables from the decision of entering into corruption. This happens because the probability of detection is an independent input of the model, and the quantity of detection a dependent output of the model. This relation provides the necessary ingredients to predict the path of the unobservable corruption given the observable detection of corruption under distinct scenarios[3].

Once the dynamic of the corruption detection is provided. It is possible to infer the causes of its behaviour. Notably, a successful policy intervention is any that lowers the corruption level, predictions show that a successful policy shock must increase the detections of corruption, at least in the first periods. The more successful the policy intervention is, the higher the detection of corruption and the shorter the period. If the policy is successful to deter corruption, corruption detection should decrease to levels lower than before the introduction of the policy, like in the study of Miller (2009)[4]. Therefore, in order to test this

[3] Such as, low enforcement low corruption; high enforcement high corruption; high corruption low enforcement; and low enforcement high corruption.

[4] Miller (2009) made his model considering anti-trust offences. However, as Berlin et al. (2018) point out, the very same structure might be useful to detect effectiveness of other types

hypotheses, it is necessary a real world example, where there was an unpredicted introduction of a sanction reduction policy. In this sense, the Brazilian anti-corruption policy of 2013 might be the perfect candidate.

The Brazilian laws No. 12.846/13 and No. 12.850/13 were enacted in August 2013 as a response to the Brazilian protests of June 2013. Protesters demanded (among other things) stronger measures against corruption from the government. The laws were inspired by the guidelines of the United Nations Convention Against Corruption (UNCAC)[5]. They introduced the possibility for individuals and corporations involved in corruption to disclose their activities in exchange for sanction reductions. Notably, they also provide other measures to inhibit corruption besides the self-reporting mechanisms. In this sense, the law No. 12.846/13 (Anti-Corruption Law) also sets strict liability for firms on corruption crimes conducted by their employees. It institutes compliance measures as possible punishments and states other sanctions in case of corporate corruption. Furthermore, the law No. 12.850/13 (Law against organized crimes) also instituted new types of investigation procedures for Brazilian authorities. All these provisions have features that may somehow deter corruption.

After the laws were enacted in 2013, a series of large-scale corruption investigations were observed nationwide in Brazil starting in 2014 and continuing in subsequent years. This raises the question of whether these investigations are indicative of widespread corruption or represent a shift towards more effective prosecution standards. Specifically, it is important to investigate whether the anti-corruption policy implemented in 2013 helped to reduce incidents of domestic bribery in Brazil.

In summary, this paper has two objectives. The first objective is to predict and to simulate what happens with the distribution of the crime of corruption in society under distinct scenarios. The second is to test the hypothesis of the predictions over the Brazilian case.

Section 2 establishes the initial conditions and parameters of the game, while Section 3 encompasses the process of solving the game to generate macro-level observables that serve as input for the empirical strategy outlined in Section 4. Finally, Section 5 provides a comprehensive summary of the key findings and conclusions derived from the paper.

## 2 Setting the Game

This section presents the necessary ingredients for setting up the game: players, stages, game parameters, states and corresponding actions or decisions for each state. Notably, each state has a pay-off that depends on the game's constant parameters. Also, each state-action pair on the current stage dictates the state in the next stage. Additionally, the game requires a timing protocol to rules

of crimes, such as corruption.

[5]Brazil signed the Agreement of the Convention on the 9th December 2003 and ratified on 15th June 2005. However, until 2013 little enforcement was observed.

when and how players can make their moves. These essential features are further explained in the following subsections.

## 2.1 Players and Timing

The game is played by two players: a bribe payer and a bribe receiver, denoted as $i \in \{\text{payer, receiver}\}$.

The game has discrete time $t$. It is assumed that in each period $t$, there is one corruption opportunity. Alternatively, $t$ can be understood as a constant time between each corruption opportunity. The players play simultaneously, which seems more realistic than a sequential game [6]. Thus, each player observes the current state at time $t$ and makes their action. The other player observes the same state but does not see the action taken by the other player at time $t$.

## 2.2 Parameters

The relevant parameters for this first example are the price of the bribe $b$ [7], 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 authorities, and the probability of being detected by the authorities $\alpha$. Note that, for this first example, the sanction $s$ is given by a fine $f$ [8], as well as everything that the agents have gained from corruption ($a$ for the payer and $b$ for the receiver).

## 2.3 States, Actions and Laws of Motion

In a game (or system), any position can be summarized as a state variable[9]. In this example, important states for players are wealth $W$, accumulated criminal history $L$, and 'states of the world' $S$, representing where agents are in a corruption game, e.g. if they are currently in a bribery or not or if they are detected or not by the authorities. The full scope of game states is explored ahead.

The actions for the players are to $pay\ a\ bribe$, to $self-report$, to ($collaborate$ or $admit\ guilt$) if detected and to $do\ nothing$[10].Each set of states has a corresponding set of allowed player actions. In this game, players with a certain level of wealth $W_t$, criminal history $L_t$, and state pf the world $S_t$ can take an appropriate action. For example, undetected players with sufficient funds can

[6] This would logically imply a take-it-or-leave-it proposition of the bribe, which is also unrealistic.

[7] It is possible to make the bribe $b$ an endogenous decision. For now, let us assume that the players are price takers regarding the bribe $b$.

[8] In many jurisdictions, there are non-monetary sanctions for both individuals (e.g., imprisonment) and corporations (e.g., license and activity impediments). As Becker (1968) points out, they are complementary and necessary for optimal deterrence. For now, assume that agents are able to translate the non-monetary fines into the value of $f$. Nonetheless, the argument on non-monetary fines is further developed in the following sections.

[9] States are subjective and depend on the analyst's game design. Finding the right state is an art (Ljungqvist and Sargent, 2012).

[10] Which can also mean not accepting a bribe.

*payabribe* and enter corruption, players who previously paid bribes can self-report and send the other player to jail, and detected players can choose to collaborate or not.

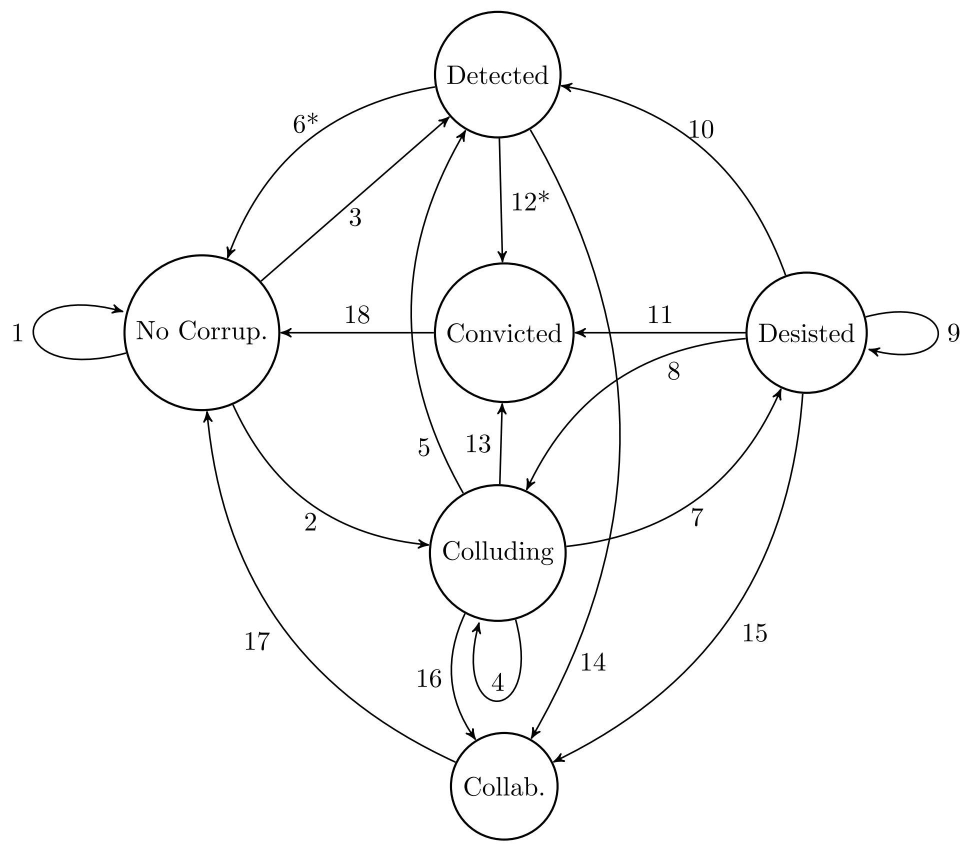
Each pair of state and action from both players follows a law of motion that sets rules for the next state. Therefore, the laws of motions are functions that predict the state in $t+1$ given the pair 'state-action' in $t$. For example, if players have never paid bribes and decide not to pay, they will be in a 'no corruption' state. If they choose to pay and are not detected, they will be in a 'corruption' state; if they are detected, they will be in a 'detected' state.

The next subsections present the set of states, some actions related to each state and their laws of motion.

### 2.3.1 States of the World

To smoothly present all the possible states of the world, and how each action leads to other states, the Figure 1 shows the a diagram relating all the states.

Figure 1: States of the World with Non-Trial Resolutions




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

The arrows represent the transition from the states[11]. 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.

### Law of Motion

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 *pay a bribe* again they can be detected (5) or go back to the same state (4). However, if they decide to *do nothing*, they might be detected by their past crimes (5) or, if not detected (7), go to the state *desisted*[12]. The state *collab* stands for the states where the agents *collaborate*(14), or $self - report$ (15 and 16)[13] and are collaborating with authorities. The *convicted* state is where agents pay the full fine. Admittedly, agents are convicted if they are reported (11 and 13) or when the other party *collaborates* (12*). 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).

It is important to note that agents may not be in the same state simultaneously. Thus, there is a state $S_i \in S$ for each player $i$.

As mentioned before, law of motion is a function of the present state $S_t$, the decision $d_t$ and a stochastic probability. There are two distinct stochastic events which can happen with different probabilities[14] depending on the state

[11] 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 $t$.

[12] 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.

[13] For simplicity, the actions of *collaborate* and $self - report$ are represented as the same, but in different states. It is only possible to *collaborate* if detected and to $self - report$ if not detected.

[14] 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. $p(x' = s_{con}|x = s_{cor}, d_i \neq 2 \ and \ d_j = 2) = 1$

$S_t$ that agents are in.

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

Here, $\epsilon(S_t) \sim$ Bernoulli($\alpha$) or $\epsilon(S_t) \sim$ Bernoulli($\beta$) depending on the states and actions. It is important to note that the current states contain all past information (Markovian Property)[15].

### 2.3.2 Wealth and 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.

#### 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 $\eta$. In the next examples, it is assumed that the utility function $u(.)$ is an isoelastic utility function, also known as a constant relative risk aversion (CRRA) function, defined as:

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

Where $\eta$ 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 $c_{i,t}$.

#### Budget and the Asset Market

Following each interaction between players, the agents either gain or lose some money due to corruption, which can influence their willingness to engage in collusion again. Moreover, even if they choose not to participate in bribery, there are alternative options available to the player's budget[16]. Hence, a crucial factor in the players' decision-making process is their budget set.

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. $p(x' = s_{nc}|x = s_{con}) = 1$

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

[16] Note that without an asset market, income and substitution effects are not applicable.

It should be noted that the decision to consume $c_{i,t}$ is subject to the agent's budget constraint, which can be represented as the total wealth of players at each time $W_{i,t}$. Naturally, there exists a trade-off between consuming today and saving for tomorrow, which is determined by the time discount factor $\gamma$. As a result, if agents choose to save instead of consuming today, they must earn some interest rate $i_r$[17] on their savings.

It is important to note that agents may have multiple sources of income. Specifically, there is an income from corruption denoted by $y_c$, a financial income denoted by $y_f$, and a non-financial income or wage denoted by $y_0$[18]. The income from corruption $y_c$ can be expressed as shown in (6), the non-financial income $y_0$ is a constant autonomous income such as a wage, and the financial income $y_f$ can be expressed as:

$$y_{fi,t+1} = (W_{i,t} - c_{i,t} - \phi_{i_t})(1 + i_r). \tag{2}$$

Note that 2 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.

**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,

$$W_{i,t} = y_{fi,t} + y_{ci,t} + y_{0i,t} \tag{3}$$

Notably, in this setting, agents have a new allowed action to choose. They can now decide on how much to consume $c_{i,t}$[19]. This action determines how the state $W_{i,t}$ is going to evolve. Therefore

$$W_{i,t+1} = (W_{i,t} - c_{i,t} - \phi_{i_t})(1 + i_r) + y_{ci,t+1} + y_{0i,t+1} \tag{4}$$

### 2.3.3 Liability or Criminal History

In the previous example, as well as in most of the literature on leniency, it is assumed that players are not held liable for crimes committed in the past (Marvao and Spagnolo, 2016). In other words, if a crime is not detected immediately after it is committed, then agents cannot be held accountable for it in future stages. This assumption is convenient but unrealistic, as it implies that once agents engage in corruption, they would never abandon it in future periods[20]. To overcome this limitation, this section introduces the concept of judicial liability, which evolves as the agents repeatedly play the corruption game.

[17] Note that the interest rate $i_r$ and the time discount factor $\gamma$ may not necessarily be equal.

[18] In this example, a living wage is assumed to be available to enable agents to obtain sufficient funds to pay bribes. This is done to prevent the Markov chain from becoming reducible. Without this assumption, agents who are convicted and lose all their resources would be unable to play again and would remain trapped in that set of states.

[19] 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.

[20] Additionally, without liability for past crimes, there is no incentive for agents to report corruption committed by others, as they would face no consequences for their own actions.

#### 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 $\delta$ over time. Such that,

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

#### Liability Transition Rule

$$L_{i,t+1} = (1-\delta)L_{i,t} + l_{i,t+1} \tag{5}$$

There are several reasons to assume that judicial liability decays over time. One reason is that information may become lost or less reliable over time, making it difficult to prove that a crime was committed. Additionally, prescription rules may eventually absolve agents of liability for offenses committed long ago.

### 2.3.4 Information

The game design determines how information is made available to players, and in this case, the game is considered to be complete and imperfect. The agents have complete information because they know the payoffs associated with each state and how each pair of actions and states leads to future states. However, there is uncertainty about future states because players are uncertain about whether authorities will detect or convict them for their actions, which makes the game imperfectly informed.

### 2.3.5 Sanctions and Pay-offs

In the leniency program, there are rules for reducing fines. If agents unilaterally self-report before detection, they receive a reduced fine of $Rf$, where $R < 1$. If both the 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 plead guilty, in which case the fine reduction is lower, with $Pf$ for unilateral pleading guilty and $pf$ when both players plead guilty simultaneously, where $1 > p > P$. In summary, the fine reductions are more lenient for unilateral self-reporting than for simultaneous reporting, and more lenient for self-reporting before detection than after. Admission of guilt is considered an act of reporting after detection. Thus, a collaborator is a player who either reports before detection or admits guilt after detection.

Given the new set of states[21], there are corresponding rewards $y(S_t)$ from each state in $S$. Lastly, the criminal history indicates the number of crimes for which a player is liable..

[21] Note that there are hidden states within the state *collab*, which indicate whether the agents collaborated unilaterally or simultaneously.

$$
y_{c\ i,t}(S_{i,t}, L_{i,t}) = \begin{cases} 0 \text{ if not colluding} \\ \pi_i \text{ if colluding} \\ 0 \text{ if desisted} \\ 0 \text{ if acquitted} \\ -(L_{i,t} \times f) - \pi_i \text{ if convicted} \\ -(L_{i,t} \times Rf) - \pi_i \text{ if reported alone before detection} \\ -(L_{i,t} \times rf) - \pi_i \text{ if reported simultaneously before detection} \\ -(L_{i,t} \times Pf) - \pi_i \text{ if the other party admits guilty} \\ -(L_{i,t} \times pf) - \pi_i \text{ if both parties admit guilty} \end{cases} . \tag{6}
$$

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 consistent of the number of times that the crime was performed $L_t$[22] times the reduced fine plus any benefit received from the corruption crime $Rf$, or else, $spayer, t = L_t Rf + a$. Whereas, the *receiver* is convicted with the full sanction times the number of times the crime was performed and their gains from corruption $sreceiver, t = L_t f + b$. However, if the *receiver* simultaneously report, both agents receive a sanctions of $s_{i,t} = L_t rf$. 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 Solving the Game

In this section, the corruption game with non-trial resolutions is solved. 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 notation for clarity in problem formulation. Following the tradition in dynamic programming literature, the subscript $(t+1)$ will be replaced by a *prime* $(')$ superscript. Thus, all variables without a subscript $(t)$ refer to the present period. Let us define the state vector $\mathbf{x}$ as a collection of all possible states, where $\mathbf{x}_i \supset \{Wi, L_i, S_i\}$. Similarly, let the action set $\mathbf{d}$ be defined as a collection of all possible actions, where $\mathbf{d}_i \supset \{di, c_i\}$. Finally, $\theta$ is the vector containing all constant parameters. In summary, there are two vectors that contain all the states and actions, respectively.

[22] Accounting for the liability decay rule discussed before.

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

Where, $j$ denotes the other player.

## 3.1 The Agent's Problem

Both agents $i$ face the same economic problem[23]. They want to maximize their utilities. Therefore, it is possible to write their objective function as:

$$\max_{\mathbf{d}} E_0 \sum_{t=0}^{\infty} \gamma^t u(c_t(\mathbf{x})) \tag{8}$$

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

The complete description of the constraint in (8), is given by the laws of motion (1), (4) and (5). Or else,

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

[23] 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 (10). 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:

$$\min_{R,P,F} E_0 \sum_{t=0}^{\infty} \gamma^t S_{c_t}(R, P, F) \tag{7}$$

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

### 3.2 The Value Function

The agent's problem (8) is recursive due to its first-order difference equation constraints, and some future states are stochastically determined. Such stochastic processes are known as Markov decision processes[24]. Dynamic programming is the most appropriate methodological framework to solve this type of problem (Puterman, 2005). To transform the problem into a dynamic programming problem, it is necessary to construct the value function, which represents the value of the state vector to the agent $V(\mathbf{x})$.

Let us define the value function $V(\mathbf{x})$ for the initial states $W_0$, $L_0$, and $S_0$ as $V(\mathbf{x_0})$, which represents the optimal value for the initial states. The value function $V(\mathbf{x})$ of the states can be defined by the following *Bellman Equation*,

$$V(\mathbf{x}) = \max_{\mathbf{d}}\{u(\mathbf{x}) + \gamma E[V(\mathbf{x}')|\mathbf{x}, \mathbf{d}]\} \tag{9}$$
$$\text{s.t. } \mathbf{x}' = f(\theta, \mathbf{x}, \mathbf{d}, \epsilon')$$

or,

$$V(W, L, S) = \max_{d,c}\{u(c) + \gamma E[V(W', L', S')|W, L, S, d, c]\} \tag{10}$$

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

The value function $V(.)$ computes the local maximum of the current utility from consumption $u(c)$, considering all possible future states and their corresponding consumption. This exhaustive exercise aims to identify the path that yields the best expected outcomes, given all possible states.

### 3.3 Dynamic Programming Solution

The objective is to determine the optimal strategies for both players in equilibrium. To solve dynamic programming problems, agents must select the best actions $\mathbf{d}$ that maximize the current state value $\mathbf{x}$ and future ones $\mathbf{x}'$, given a stochastic transition law $\Omega$. The transition matrix $\Omega(\mathbf{x}', \mathbf{d}, \mathbf{x})$ represents the probability density of the next state $\mathbf{x}'$ given the current state-action pair $(\mathbf{x}, \mathbf{d})$[25]. 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

[24] Also referred to as stochastic control problems in engineering.

[25] The matrix $\Omega(\mathbf{x}', \mathbf{d}, \mathbf{x})$ can be calculated for a given number of states $n$ and actions $m$ as:

iterating the value function given the transition matrix $\Omega$ up to a fixed point[26]. If the $\Omega$ matrix is stochastic, then the solution is given by maximizing each value function $V(\mathbf{x})$ given all states for each period, or else:

$$V(\mathbf{x}) = \max_{\mathbf{d}}\{u(\mathbf{x}) + \gamma \sum_{\mathbf{x}} V(\mathbf{x}')\Omega(\mathbf{x}', \mathbf{d}, \mathbf{x})\} \tag{11}$$

If $V^*$ is a unique solution to the Bellman Equation (11)., then it is possible to define an optimal policy $\sigma(\mathbf{x})$ such that:

$$\sigma(\mathbf{x}) \in \text{argmax}_{\mathbf{d}}\{u(\mathbf{x}) + \gamma \sum_{\mathbf{x}} V^*(\mathbf{x}')\Omega(\mathbf{x}', \mathbf{d}, \mathbf{x})\} \tag{12}$$

The optimal policy $\sigma_i$ is a vector of dimensions $(n \times 2)$, where $n$ represents all possible state combinations. For each row in $\sigma_i$, 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 also for all subsequent periods. Therefore, the optimal policy $\sigma_i$ is a rule for identifying the best position in the state space that leads to better outcomes, not just for the next period, but for all periods.

### 3.3.1 Strategic Interdependence and Game Equilibrium

Note that the individual decision functions $c$ and $d$ depend on $S$, which in turn depends on $d$ again. This demonstrates the strategic interdependence of the game, where a player's strategy depends on the other player's strategy. Substituting (1) in $d_i$ yields,

$$d_i = d(W_i, L_i, S(S_i', d_j(W_j, L_j, S_j, \theta)), \theta), \tag{13}$$

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 (11) in terms of the separate decisions of $i$ and $j$,

$$\sigma_i(\mathbf{x}) \in \text{argmax}_{\mathbf{d_i},\mathbf{d_j}}\{u(\mathbf{x}) + \gamma \sum_{\mathbf{x}} V^*(\mathbf{x}')\Omega(\mathbf{x}', \mathbf{d_i}, \mathbf{d_j}, \mathbf{x})\} \tag{14}$$

---

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

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

Each player's strategy $\sigma_i$ depends on the other player's strategy $\sigma_j$, and vice versa. The equilibrium (MPE) arises when both players identify the other player's optimal decision, given their own, and still maintain their decisions (i.e., they do not want to change their policy $\sigma$). Therefore, the solution to the agent's problem is a pair of optimal policy vectors ($\sigma_i$ and $\sigma_j$) that maximizes the value function $V(\mathbf{x})$ for both players.

#### The Markov Perfect Equilibrium

The equilibrium found in this game is a Markov perfect equilibrium (MPE). This is because both players are selecting actions based only on the current state and the actions taken by the other player in the previous period, which satisfies the Markov property. Moreover, both players are maximizing their expected utility given the stochastic transition law $\Omega$, which is also a requirement for MPE. In this way, the equilibrium policy vectors ($\sigma_i$ and $\sigma_j$) represent the best responses to each other's optimal strategy and form a Nash equilibrium. Additionally, the MPE strategy guarantees that each player is optimizing their utility given the other player's strategy, which leads to a socially optimal outcome. Therefore, the MPE solution is the best solution for this game, given the strategic interdependence and stochasticity of the decision-making process.

It is important to note that some games may not have an equilibrium. In other words, the MPE equilibrium occurs when both agents accept their current strategy, given the predicted next best strategy from the other player. Consequently, if players' strategies keep changing after they predict the other player's next move, the policy vector does not converge, and there is no solution to this game.

### 3.3.2 Discretization

Solving continuous-state dynamic programming problems is often challenging[27]. One approach to address this challenge is to discretize the problem by converting the state and action variables into finite grids or lattices that agents can access[28].

It is important to note that discretizing the problem may sacrifice the model's realism. This is because boundaries must be imposed on the grids, which can lead to strange decisions near the boundaries. For example, in this setting, the minimum wealth $W$ that an agent can have is zero. Therefore, if the sum of bad outcomes exceeds an agent's wealth, they may act recklessly, a situation known as limited liability, which occurs more frequently in a small grid than in reality. The same can occur at the other end of the grid, where agents have no incentive to invest in growing their assets when they reach the maximum wealth.

It is evident that these grid limitations also affect the steady state of the game. Therefore, to avoid these problems, we sample the results for society from the middle of the grid in Section 3.4.1.

[27]While there are a few known solutions, most of them use linear quadratic constraints, which yield Euler Equations with desirable analytical properties.

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

### Curse of Dimensionality

Discretization implies abandoning traditional analytical mathematics for numerical computing. 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 multiplying all the elements of each state and action. As a result, the number of possible combinations grows exponentially with the number of states or actions. This problem, called the '*curse of dimensionality*' by Bellman and Dreyfus (1962), means that the number of different combinations in the state-action lattice becomes excessively large if the game is played in its full complexity.

For example, in the complete game where only natural numbers are allowed for states and actions, with a maximum wealth of $\overline{W} = 40$ and a maximum liability of $\overline{L} = 5$, the total number of points is equal to ($\overline{W}^4 \times \overline{L}^2 \times 8^2 \times 4^2 = 36.864 \times 10^9$)[29]. Therefore, operating with matrices and vectors with more than 36 billion elements would require a significant amount of computational power.

### 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[30]. For 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[31].

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[32]. Additionally, at each iteration the feasible states are recalculated. Unfortunately, it implies in distinct dimensions of the policy vector $\sigma$ for each

[29] There are four spaces with a length of 40: the wealth space and the consumption space for both players, the liability spaces for both players, the states of the world for both players, and the actions for both players.

[30] 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.

[31] I also conveniently implies that the $\Omega$ matrix is not sparse. In other words, all rows in the $\Omega$ 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

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

iteration[33].

## 3.4 Solution

The agent's problem can be solved by iterating the policy function using methods such as value iteration, policy iteration, modified policy iteration, or linear programming. However, finding the optimum solution for one player may not be the best for the other. Therefore, it is necessary to find the pair of strategies that solve the problem for both players, separately maximizing their own objective functions given their own rewards.

The solution approach used here involves finding the optimal policy $\sigma_i$ assuming that player $j$ always plays a greedy strategy, followed by solving for $\sigma_j$ using the previously determined $\sigma_i$. By definition, a Markov Perfect Equilibrium (MPE) occurs when both players observe the other player's best strategy $\sigma$ but do not change their own strategies. If strategies start alternating, the solution does not converge, and the game has no equilibrium[34].

### 3.4.1 Results and Equilibria

The findings can be condensed into optimal strategies or policy vectors, denoted as $\sigma_i$, which each player utilizes to maximize their objectives. These policies are contingent on all game parameters, resulting in a unique equilibrium (if convergence occurs) for each parameter combination. To systematically examine the numerical results, it is crucial to first select and analyze a benchmark. Subsequently, the surrounding area of the benchmark can be evaluated to determine how alterations in variables impact the equilibrium.

#### The Benchmark

Let the benchmark be an arbitrary point in which players are expected to choose corruption, and not self-report or plea guilty. Importantly, the bargain power of the parties is set to be equal. This is done by choosing a bribe $b$ which is the median point in the interval between the lowest bribe accepted for the payer and the receiver[35]. The Table 1 shows the chosen parameters for this case.

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

[34] The complete code, including the solution method and results, is available at: https://github.com/caxaxa/Corruption_Game/blob/main/Corruption_game_leniency_and_non_trial_resolution.ipynb.

[35] It is possible to make the bribe endogenous, and calculated ate each point by the game parameters and states. However, it would imply in increasing the computation complexity of the model. Therefore here, agents are price takers for a current 'market bribe price'.

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 |
| $\gamma$ | 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 pleading guilty |
| $p$ | 0.8 | Simultaneous pleading guilty |
| | | Leniency Rules |
| $\alpha$ | 0.1 | Probability of detection |
| $\beta$ | 0.6 | Probability of conviction |

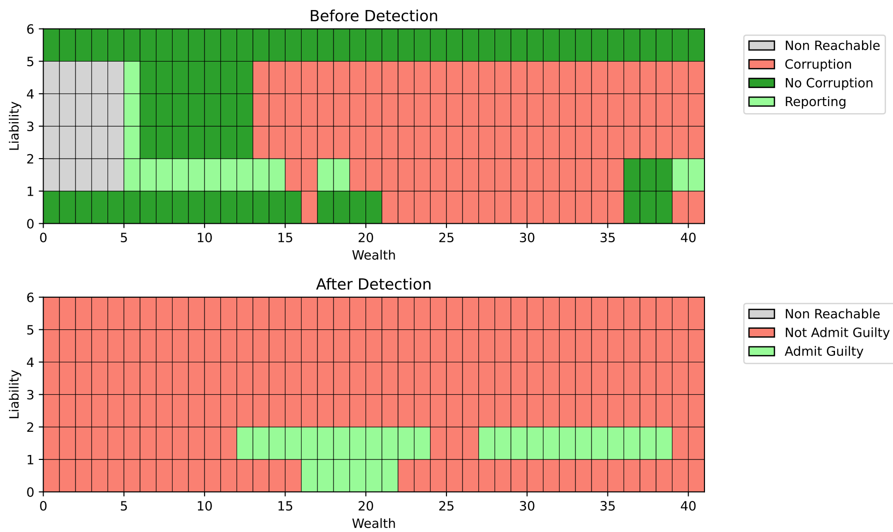
The selection of parameters for the game is an arbitrary process, aimed at approximating real-world conditions to enable a better understanding of player decision-making. It is not necessary to achieve a perfect match with reality, as the main objective is to comprehend how changes in game conditions affect player behavior. The chosen variables are intended to represent real-world conditions as closely as possible, but they may not be perfectly tuned to reality. Nonetheless, they are useful in facilitating analysis of player decisions when conditions in the game change midstream. By systematically analyzing different combinations of variables, one can gain insights into how players may behave in various scenarios. Therefore, the selection of variables is a critical step in designing a game, as it influences the understanding of players' behavior and the interpretation of results.

**Individual Strategies**

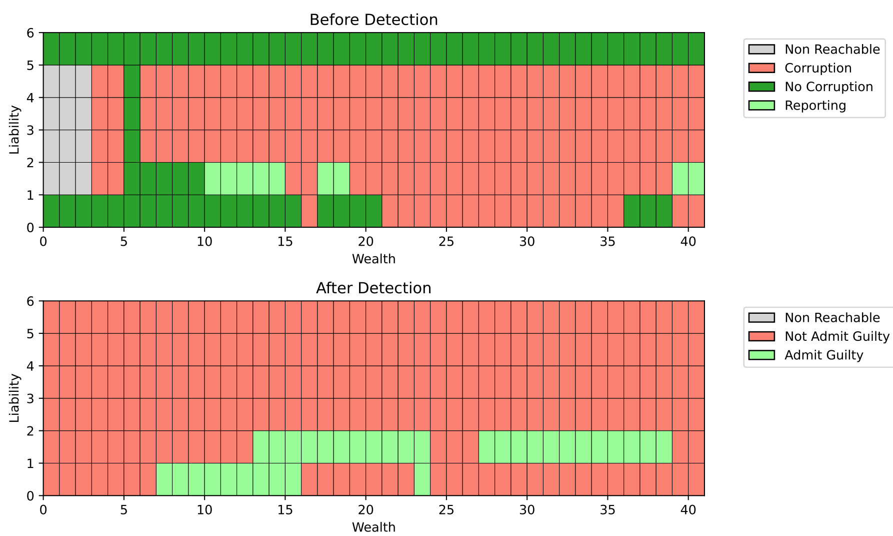
The optimal policy for each player consists of a set of rules that determine their actions at every possible state of the game. By considering the parameters listed in Table 1 along with the player's current wealth and judicial liability, it is possible to visualize the best policy for each player. Figure 2 depicts the player's decisions before and after detection for all possible combinations of wealth and liability, based on the parameters in Table 1. This visualization aids in understanding how game parameters affect player behavior and can reveal patterns and trends in decision-making. These insights can be valuable for policy-making and regulation.

Figure 2: Players' Optimal Policy

(a) Payer's Policy




(b) Receiver's Policy




Note that, if the level of liability is above zero, it means that they have entered in corruption in the past. So, if the players decide for 'no corruption' at that point, it means that they desisted from the criminal act.

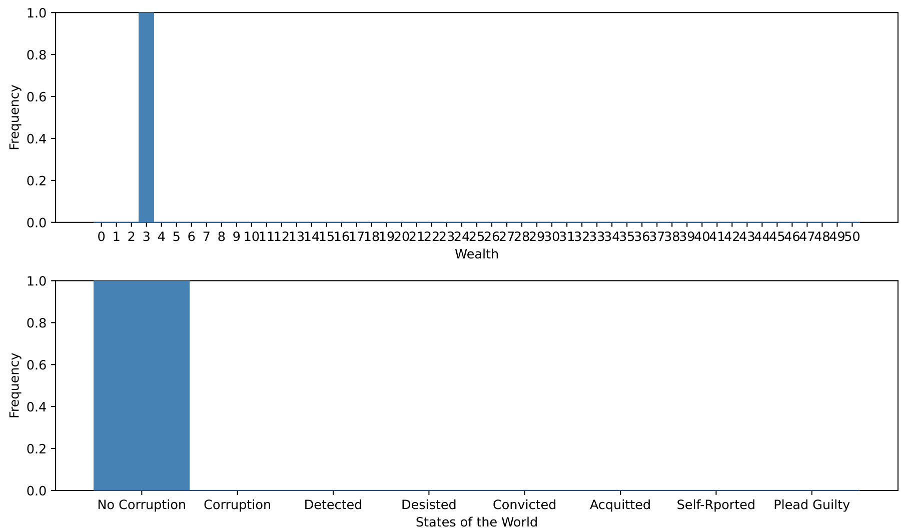
The first takeaway from the players' optimum policy is that the they are different between agents. Note that, even if one calculates the return from corruption at that point, it would be positive and be preferred at that point. However, there are combinations of wealth and liability that would make the agents prefer not to enter in corruption when it is profitable or to defect and report even when it is not profitable. However, since the players do not observe the other player's budget[36], they might be reported if they fail to assert the other player's wealth level. Lastly, there are some combinations that would make players self-report before being detected, even when it is not profitable.

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 3.3.2. Notably, player's choices at the beginning of the grid were expected to be more risky, since the fines can go only up to the player's maximum wealth. However, agents prefer not to enter in corruption in that region. On the other hand, the end of the grid is more complex to analyse. In these points agents cannot have more wealth than $\overline{W}$, so they might choose to save less, or consume more. It is hard to say if it would change their decision towards entering in corruption. In this case, for most points, if agents are not in corruption (liability $= 0$) they do not enter in corruption. While, if they are liable for one crime in the past, they choose to collaborate and if they are liable for more crimes, than, they choose not to desist.

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 3 shows the stationary equilibrium.

[36]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.

Figure 3: Players' Optimal Policy



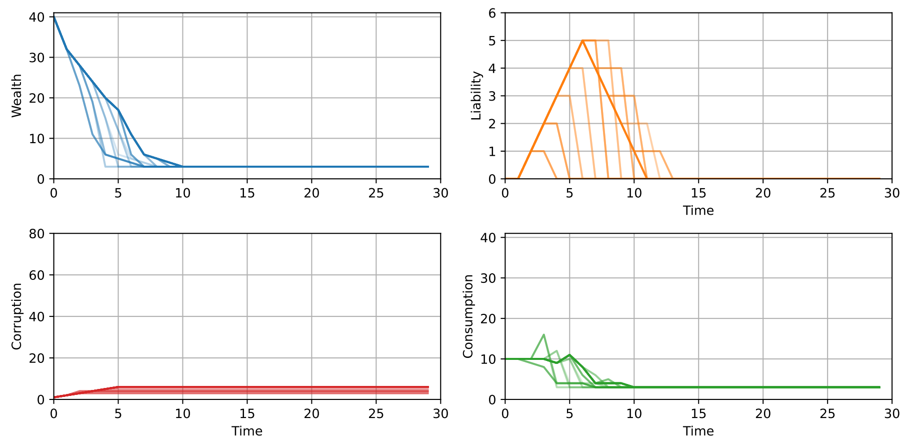

For the benchmark, the steady state equilibrium is the same for both players[37]. Note that, by construction, in order to avoid reducibility of the Markov Chain in a situation where agents would have no funds to pay a bribe or to save. The model gives the players a '*living wage*', so they will always 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 4 shows the path to the steady state.

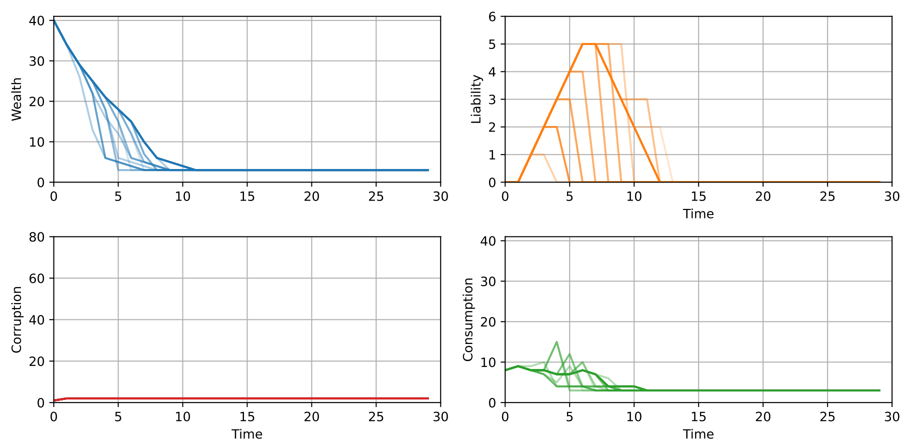
[37]It is not always the case

Figure 4: Path to Steady State

(a) Payer's Path




(b) Receiver's Path



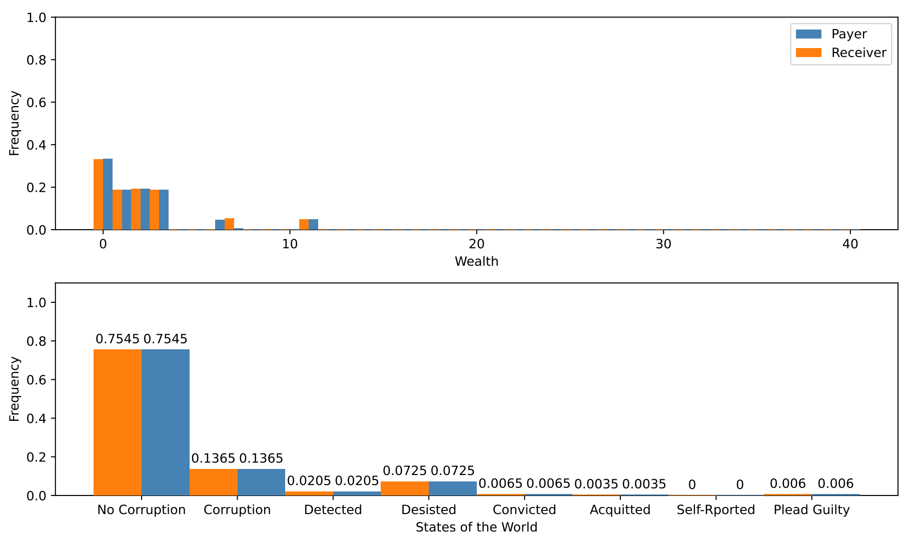

The Figure 4 shows how the agents alone would play this game for 30 periods. It also shows 50 stochastic outcomes from the decisions. Therefore, in some cases, agents are detected and in others they succeed. The multiple paths are superposed so the thicker the line, the more probable is the outcome. 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 4 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 objectives of this paper is to determine the distribution of unobserved corruption in society as game parameters vary. To achieve this, both players play their optimal strategy against each other and the resulting interaction is sampled. To avoid unusual decisions at the extremes of the grid, player wealth is sampled from a normal distribution with a mean of 20 and a standard deviation of 4. This sampling ensures that the majority of the sample is centered around the middle of the grid. The players are paired in a scenario with no corruption and play for five rounds. Afterward, the results are recorded, and another sampling is conducted for 200 rounds. The distribution of states in society is shown in Figure 5.

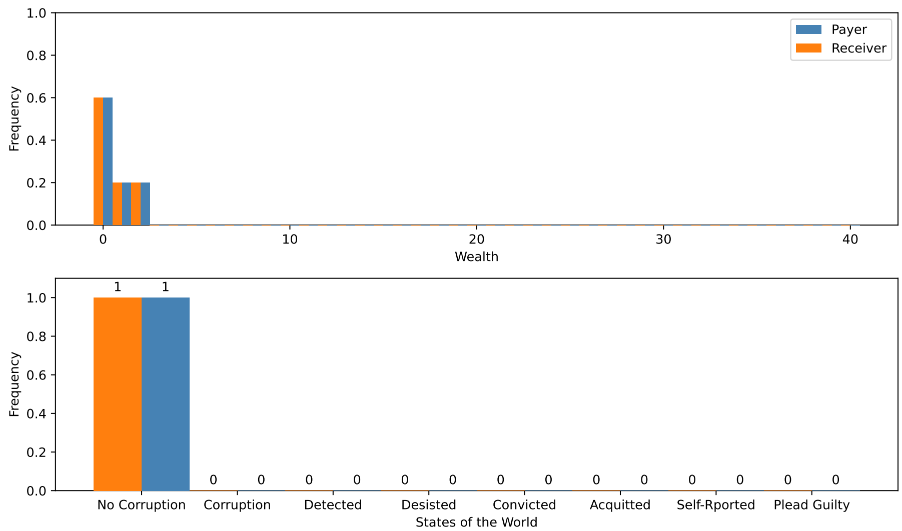
Figure 5: Distribution Sampling



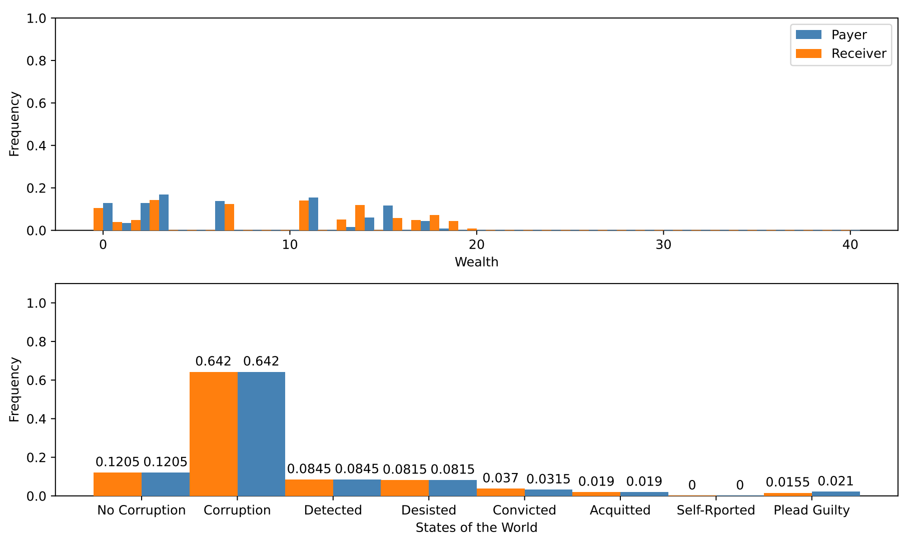

It is worth noting that, in contrast to the steady states, corruption is observable in this context. Another intriguing avenue for exploration involves testing various levels of initial wealth to assess the persistence of corruption in different socioeconomic environments. For instance, one could examine whether corruption is more or less likely to persist in a wealthier or poorer setting.

Figure 6: Distribution Sampling from Unevenly Wealthy Societies

(a) Poorer Society




(b) Richer Society

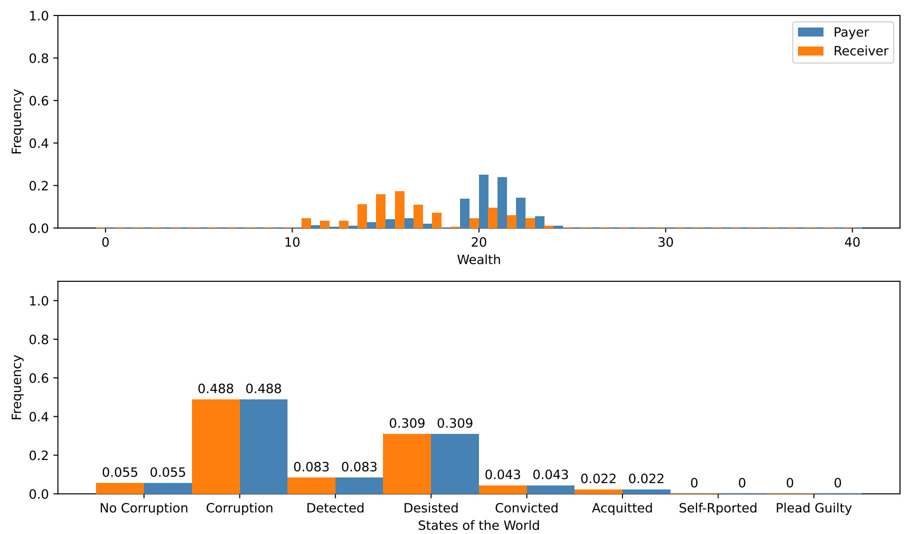



As depicted in Figure 6, poorer societies tend to avoid engaging in corruption, whereas wealthier societies tend to embrace it. Although counterintuitive, this result aligns with the notion that collusive corruption is more prevalent in wealthier countries, while more coercive forms of corruption tend to be more prevalent in smaller nations (Soreide, 2018). These findings are directly attributable to the optimal policies illustrated in Figure 2.

### 3.4.2 The Effectiveness of The Sanction Reductions

In the benchmark scenario, the players are subject to lenient fine reductions. Specifically, if either player unilaterally reports before detection, they pay no fines ($R = 0$). If both players simultaneously report, they pay only half of the fine ($r = 0.5$). However, if one player unilaterally pleads guilty, they receive a 50% reduction in the fine ($P = 0.5$), whereas if both players plead guilty simultaneously, they each receive only a 20% discount ($p = 0.8$). Figure 7 illustrates the distribution of corruption in society when there are no fine reductions ($R = 1$, $r = 1$, $P = 1$, and $p = 1$).

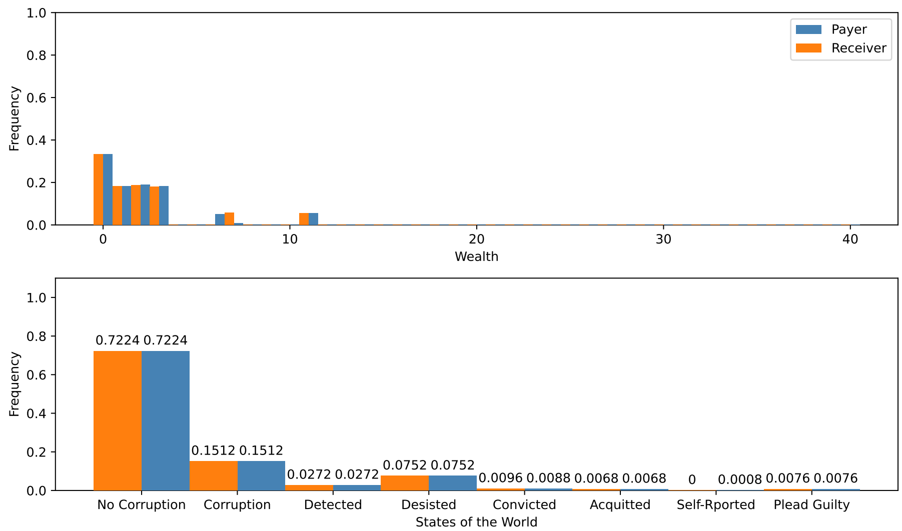
Figure 7: Distribution of States without Sanction Reductions




We can now examine whether lenient sanction reductions serve as a deterrent against corruption. To assess this, we can compare the lenient sanction reduction scenario depicted in Figure 5 against the case without sanction reductions shown in Figure 7. Interestingly, corruption falls by 72% when lenient reductions are applied in this context. Moreover, the case with moderate sanction reductions ($R = 0.5$, $r = 0.75$, $P = 0.75$, and $p = 0.9$) also demonstrates a reduction in corruption of approximately 10% from the scenario without such reductions [38].

[38] This finding is consistent with Spagnolo (2005), where the author discusses the extensions of results from Motta and Polo (2003), which argue that "first come first serve" is the only optimal policy.

Figure 8: Distribution of States with Moderate Sanction Reductions




Note that the difference between the more lenient and the moderate sanction reduction is small. Therefore, any policy focused in efficiency, or in revenue from combating corruption should consider the intensity of the corruption reduction. However, this result is sufficient to show that sanction reductions are effective against corruption.

#### 3.4.3 The Effects of a Policy Shock

The analysis of the model provides insights into agents' decision-making processes, and it can also be used to predict the trajectory of the unobservable level of corruption based on observable inputs. Specifically, the model links the number of corruption cases to the frequency of detection, whether through reporting by agents or random discovery by authorities. This connection enables the prediction of changes in the actual unobservable level of corruption based on changes in the observed level of corruption.

To illustrate this, consider a hypothetical scenario where *Regime 0* has no sanction reductions from collaborators. In contrast, *Regime 1* features more lenient sanction reductions, and *Regime 2* represents a scenario where sanction reductions do not change, but corruption becomes more challenging to detect and convict. Additionally, *Regime 3* depicts a situation where corruption is more frequently detected, even though sanction reductions are not more lenient. Lastly, *Regime 4* portrays a scenario where corruption is even harder to detect than in *Regime 2*. Table 2 summarizes the key features of each regime.

Table 2: Policy Shocks

| Regime | Sanction Reductions | Probability of Detection | Corruption Change |
| --- | --- | --- | --- |
| 0 | No Reductions | Average Detection | - |
| 1 | Lenient Reductions | Average Detection | Decreases |
| 2 | No Reductions | High Detection | Increases |
| 3 | Lenient Reductions | Low Detection | Decreases |
| 4 | No Reductions | Very High of Detection | Increases |

The Figure 9 shows path of the corruption detection after four distinct policy shocks.

Figure 9: Policy Shocks over Corruption and Detections

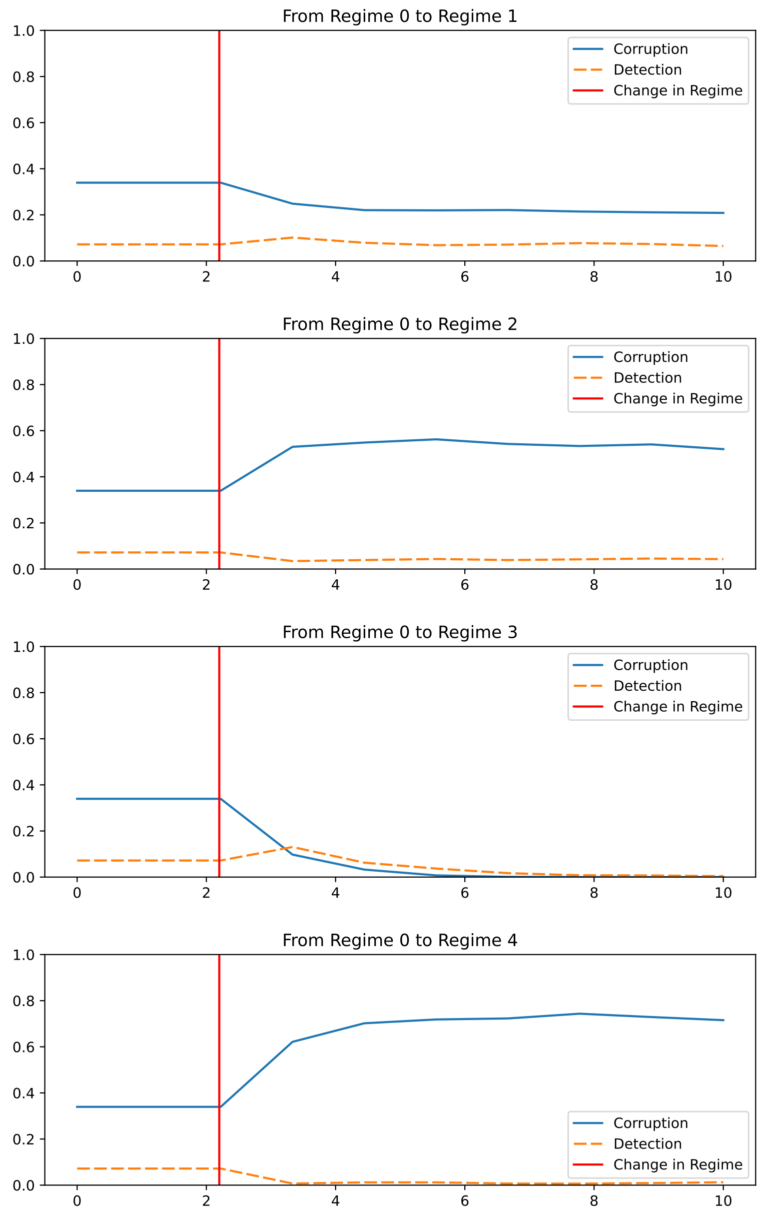



Figure 9 indicates that observable detections tend to increase as corruption decreases and vice versa, which may appear counterintuitive. One might expect more corruption detections in more corrupt countries. However, this line of thinking does not apply to observable corruption detections over time within the same place. Notably, if two different locations are compared, the one with more corruption detections is likely more corrupt.

This finding aligns with the work of Miller (2009), where the author employs a simpler model to predict the number of cartels detected in an economy. This study offers a robust response to criticisms that Miller's work oversimplifies criminal interactions, by incorporating more states and a real economic problem of utility maximization. Importantly, this research introduces a key analytical difference by separating the effects of sanction reductions from detection probability impacts. It quantifies corruption deterrence due to sanction reduction policies versus increased prosecution capacity. This distinction is vital for analyzing anti-corruption policies, like the 2013 Brazilian reforms, which included sanction reduction, collaboration mechanisms, and administrative measures such as task-forces[39].

## 3.5 Game Predictions

This study investigates the corruption game, examining how agents decide to engage in corruption and potentially report misconduct in exchange for judicial benefits. The model considers the tradeoff between bribery and investing in the asset market, suggesting that profitable bribes are necessary but insufficient for entering corruption. Additionally, the model analyzes the game's dynamics, where agents make optimal decisions based on their current state and its impact on their future. Consequently, wealth and criminal liability are critical factors in agents' decision-making.

The model suggests that sanction reductions for collaborators decrease corruption, with more lenient reductions resulting in less corruption. Successful anti-corruption policy shocks can initially increase corruption detection, due to new reporting or enhanced prosecution productivity.

#### Caveats

The curse of dimensionality necessitates using lower-dimensional grids, which can oversimplify the space of possibilities for players. To avoid this issue, players estimate others' wealth as equal to their own. However, if the game continues too long with the same players, they may systematically mispredict each other's wealth levels, requiring the game not to be repeated excessively with the same participants.

#### Further Studies

The model presented has significant flexibility, allowing for various scenario predictions, such as the impact of interest rate changes, bribery terms, or agents' risk aversion. Reinforcement learning techniques like DPG, DQN, and DDPG

[39]The Brazilian anti-corruption experience is explored in the empirical extension of this work.

can provide policy vectors for agents' decisions and potentially expand the state space without memory constraints.

From the estimated corruption and detection path (Figure 9), real-world data can be used to calibrate the model, inferring changes in unobservable corruption. Notably, Brazil's 2013 anti-corruption policies, which introduced leniency agreements and new plea bargaining conditions, offer an ideal case study to infer corruption levels based on the country's data.

# 4 Empirical Strategy

In natural sciences, when researchers seek to infer something about an unobservable phenomenon, they often rely on matching a predicted sign to noisy data. Corruption is not directly observable. Nevertheless, some outputs from bribery can be observed, such as the number of inquiries and convictions. Fortunately, the preceding sections provide a good theory about the mechanics of bribery, enabling us to predict the sign of the expected relationship between observable variables and the unobservable level of corruption. In other words, it is possible to anticipate how observed variables should behave in response to changes in the unobservable level of corruption.

The findings in this study resemble those of Miller (2009), making it prudent to adopt the author's approach for empirically testing if an antitrust policy can deter cartel formation. Miller innovatively constructs a semi-structural model of cartel formation, depicting the transition from collusion to non-collusion[40]. Although useful for guiding empirical strategy, this model is less effective in explaining the impact of proposed enforcement changes.

Similar to Miller (2009), this study measures criminal levels by observing the number of cartels detected by an Antitrust Authority before and after a policy shock. Successful anti-crime enforcement should initially raise crime detection due to improved efficiency, followed by a decrease below pre-enforcement levels, indicating deterrence, as Figure 9 predicts.

Though focused on cartel activity, the same methodology can analyze policy interventions for other misconducts. Various studies have applied Miller's model to different criminal variables. Brenner (2009) investigated cartel activities in the European Union, while Acconcia et al. (2014) examined mafia-related crimes in Italy, and Berlin et al. (2018) assessed corruption crimes in China. These studies searched for detection spikes as evidence of enhanced prosecutorial efficiency, with subsequent downtrends indicating deterrent effects.

Brenner (2009) studied the European Union Leniency program of 1996, finding a significant spike in cartel detections but no evidence of increased deterrence. Acconcia et al. (2014) reported positive effects on prosecution and deterrence of mafia crimes, with increased whistleblowers linked to greater judicial efficiency. Berlin et al. (2018) found that China's 1997 criminal law reform failed to improve detection and deterrence due to insufficient asymmetry.

[40] Extensive literature in this field includes (Harrington, 2008; Spagnolo, 2005; Aubert et al., 2006; Motta and Polo, 2003), which inspired the theoretical model in Part II of my thesis

Other empirical studies, like Amir et al. (2018), explore similar approaches for tax evasion crimes, using tax revenue as the relevant variable. If tax evasion is deterred, tax revenue should remain high after policy introduction.

Using the logic from Miller (2009) and subsequent studies, this methodology can be applied to Brazilian corruption detection data[41] to analyze the country's anti-corruption policy.

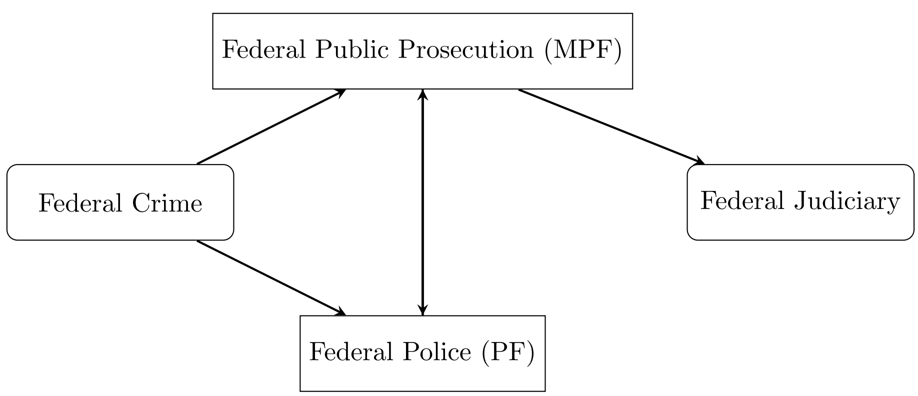
## 4.1 Institutional Background

To apply Miller's method to corruption in Brazil, it is essential to understand the country's criminal procedural processes.

Under Federal Law No. 3.689/41, both the Brazilian General Attorney's Office (Public Prosecution) and the Federal Police are competent to independently investigate public interest corruption crimes. However, the standard (*de facto*) procedure involves both institutions conducting investigations jointly. Typically, the Federal Police initiate the corruption investigation (Art. 5, I), conclude it, and submit it to the Public Prosecution for accusation (Art. 24). Alternatively, the investigation may begin with the Prosecution and then involve the police for parallel joint investigations (Art. 5, II). Only in exceptional cases does the Prosecution solely conduct an investigation[42].

Figure 10 depicts the process flow within the Brazilian system.

Figure 10: Brazilian Criminal Investigation Procedures Flowchart




It is reasonable to conclude that the Federal Police is generally the first to detect potential corruption crimes and initiate an inquire in Brazil. Even though the Public Prosecution can sometimes also be the first to detect. However, in the standard investigation cases, the prosecution will only know about an eventual corruption crime after the police complete its investigations.

[41]Candidate variables are discussed in Section 4.2

[42]In 2015, the Brazilian Supreme Court formalized the understanding that the Brazilian Prosecution could investigate a case alone ((HC) 89837/STF). Prior to this decision, criminal investigation was a Police monopoly.

## 4.2 The Data

The data used in this study was obtained from the Brazilian Public Prosecution's (MPF) online processual search engine[43]. It comprises 885,675 investigations conducted by the Brazilian Federal Police, organized by the start date of the investigation from January 2009 to January 2020.

Data is available for periods before 2009 and after 2020 in the database. However, data before 2009 may be biased for minor states that did not utilize the information system at that time. The period from 2020 onwards was excluded due to the COVID-19 health crisis, which led to exceptional emergency procurements and contracts, causing a rent-seeking supply shock that could have impacted the number of corruption denunciations and investigations in Brazil[44].

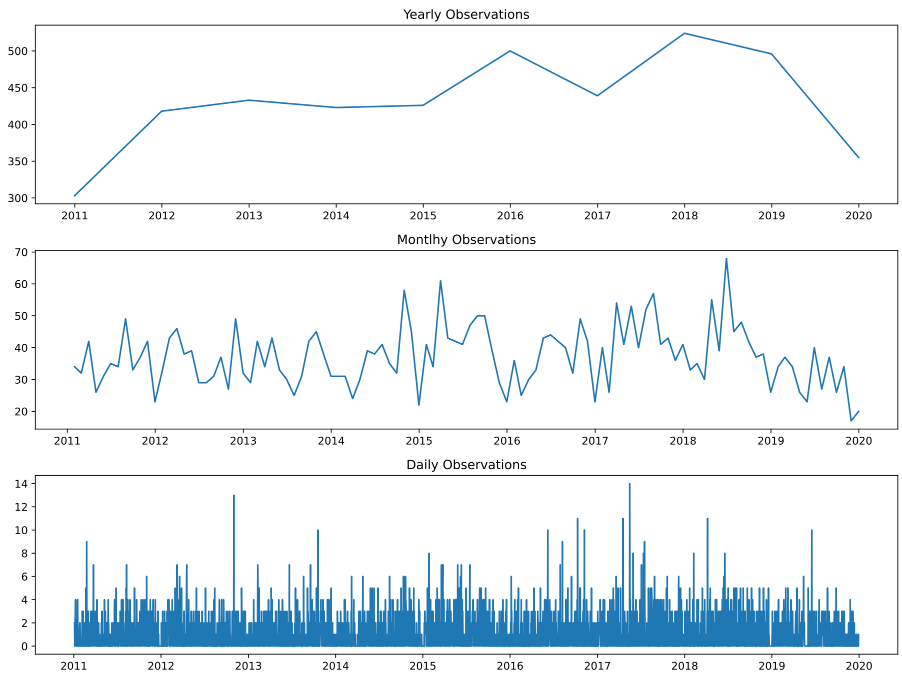
Figure 11 displays the number of inquiries regarding passive and active corruption crimes[45] from 2010 to the end of 2019.

[43] Available at: http://apps.mpf.mp.br/aptusmpf/portal?servidor=portal and downloaded between February and March 2020. The code I created for scraping the data is available at: https://github.com/caxaxa/MPF_Crime_Scraper.

[44] The corruption trend in Brazil after 2020 is a relevant topic for further study. It is primarily attributed to notable setbacks in the anti-corruption efforts, which stem from a series of Supreme Court decisions favoring defendants and halting the enforcement of sentences following second-instance convictions. These setbacks have created a situation where several civilians convicted in the Lava Jato case are not incarcerated any longer. Notably, no convicted politicians, including the current president of Brazil, are in prison.

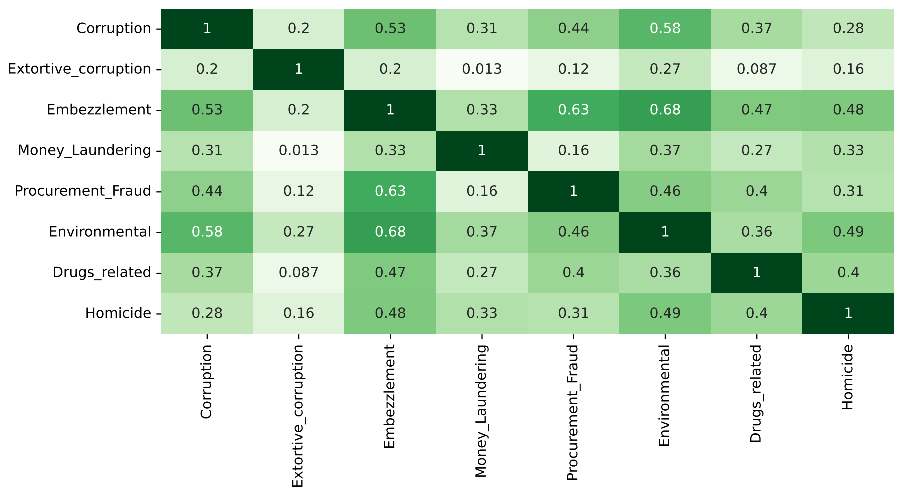
[45] Articles 317 and 333 from the Brazilian Penal Code, Federal Decree-Law No. 2.848/40.

Figure 11: Corruption Inquiries by Starting Dates




The database also provides the crime, and the state in which the investigation is conducted. The Tables 5 and 6 show some aggregate values from the dataset, also Figure 12 shows the correlation between multiple different crimes in Brazil.

Figure 12: Crime Correlation Matrix




From Figure 12, it is possible to observe some level of correlation between crimes, particularly between embezzlement and environmental crimes. This may be because corruption occurs when these crimes happen, or they might even be part of the same inquiry. Alternatively, it could be due to law enforcement as a whole responding to other exogenous variables. As a result, this paper also controls for a set of external variables. The chosen ones are the quarterly Brazilian real GDP growth[46], the monthly Brazilian unemployment rate[47], the monthly average of the main national treasury bonds' real interest rate (SELIC)[48], and all other crime investigations except corruption crimes.

## 4.3 Estimation and Hypothesis Testing

The main empirical strategy is similar to the approaches used in previous works such as Miller (2009), Brenner (2009), and Spagnolo (2018). While the first two studies focused on cartel detection, the last one focused on corruption, which is precisely the objective of this work. The authors' methodologies can be adapted to this case and formalized as follows:

[46]Source: Instituto de Pesquisa Economica Aplicada (IPEA), available at: http://ipeadata.gov.br.

[47]Before 2012, the source is Fundacao Sistema Estadual de Analise de Dados (SEAD), available at: https://www.seade.gov.br/. After 2012, the data is from Instituto Brasileiro de Geografia e Estatistica (IBGE), available at: https://www.ibge.gov.br/estatisticas/sociais/educacao/9127-pesquisa-nacional-por-amostra-de-domicilios.html?=&t=series-historicas. The data was corrected using the global monthly weighted average.

[48]The nominal rate was gathered from the Brazilian Central Bank, available at: https://www.bcb.gov.br/estatisticas/txjuros, and deflated by the IPCA inflation index, available at: http://ipeadata.gov.br.

$$Y_t = \beta_0 + \beta_1 D_t + \beta_2 T1_t^n + \beta_3 T2_t^n + \beta_4 X_t + \varepsilon_t \quad (15)$$

for,
$t$= Observation period (month or day); and
$n$= Order of the polynomial.

Where
$Y_t$= Number of new corruption inquiries. Or, detected corruption crimes;
$D_t$= Dummy for the impact of the enforcements, being 0 before August 2013 and 1 after;
$T1$= Time effect of all sample. Being 1 at the first observation 2 at the next and so on;
$T2$= Time effect from the beginning of the policy shock. Being 1 after august 2013, 2 in the next month and so on;
$X_t$ = Vector of control variables (GDP, unemployment rate and interest rates)[49]; and
$\varepsilon_t$= Is the error term from functional predictions, it is expected to be normally distributed and i.i.d.

In order to explore various curve shapes that could fit the expected detection curve, the matrices $T1$ and $T2$ are assigned values of different-order polynomials to assess the goodness of fit. If the regression reveals a polynomial with a spike after the policy shock, followed by a decrease in the estimated mean, this may indicate evidence of an effective policy in terms of both detection and deterrence.

Equation (15) is presented in a linear form and regressed using OLS. However, it can also be regressed using Poisson and Negative Binomial estimation methods, which are better suited for countable variables like the one of interest. Although the estimated Poisson coefficients are not linear with the parameters, they can be linearized through simple transformations.

This paper tests two hypotheses:

**First Hypothesis:** There is an immediate spike in the detection of corruption crimes.

$$H_0 : \beta_1 \leq 0, \text{ against } H_1 : \beta_1 > 0$$

**Second Hypothesis:** After the spike, the mean predicted detections ($\hat{Y}$) are lower than in the pre-policy period.

$$H_0 : \hat{Y}_{2013} \geq \hat{Y}_{2020}, \text{ against } H_1 : \hat{Y}_{2013} < \hat{Y}_{2020}$$

If both hypotheses can be proven true by rejecting $H_0$ in both cases, it is possible to conclude that the policy was effective in enhancing prosecution and further deterring new corruption crimes.

[49]Note that $\beta_0$ and $\beta_1$ are constants, while $\beta_2$, $\beta_3$ and $\beta_4$ are vectors. Where, $\beta_2$, $\beta_3$ have dimensions (1 x $n$), and $\beta_3$ has dimension (1 x 3).

## 4.4 Results

This section discusses the methods used to analyze the relationship between observed data and the unobservable corruption behavior. The data analyzed consists of monthly new corruption inquiries in Brazil from 2014 to 2020. The analysis is divided into two parts: regression analysis without controls and regression analysis with controls. In the first part, Poisson regression and locally weighted scatterplot smoothing are used to investigate any trends or spikes in corruption inquiries before and after a policy intervention. In the second part, the regression is controlled for other factors that may affect corruption detection, such as unemployment, GDP, and real interest rates. Overall, the section provides a detailed overview of the analysis methods used to study the causal relationship between observed data and corruption behavior.

### 4.4.1 Poisson Regression

The first set of regressions show how different order polynomial fit the data of crimes of corruption[50]. The Table 3 shows the regressions (1) to (6), containing the coefficients for the intercept and the policy dummy for different order polynomials.

The results show that there is a consistent upward shift after the policy shock, as evidenced by the consistently positive and significant value of the Dummy variable in all estimations. Therefore, it is possible to successfully reject $H_0$ from the first hypothesis. Consequently, it means that the predicted detections of corruption increase immediately after the policy shock. If the second hypothesis is also true, then this is consistent with an increase in prosecution efficiency. Figure 13 shows the predicted values and a 95% confidence interval for the mean predicted parameters.

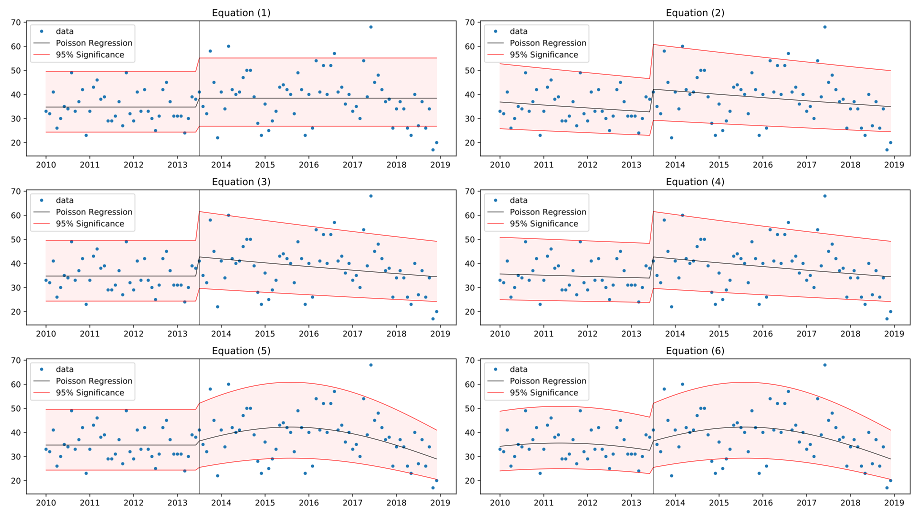
[50] Here the chosen method uses global polynomials, i.e. there is a polynomial function that predicts the entire series. Most of the time these functions tend to be very unstable. Therefore, one possible improvement is to try locally weighted scatter-plot smoothers (LOWESS). This is done in the next subsection.

Table 3: Poisson Regression without Controls

| | *Dependent variable: Number of New Corruption Inquiries* | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Constant | 3.549*** | 3.607*** | 3.549*** | 3.573*** | 3.549*** | 3.533*** |
| | (0.026) | (0.032) | (0.026) | (0.051) | (0.026) | (0.075) |
| Dummy | 0.101*** | 0.256*** | 0.206*** | 0.232*** | 0.047 | 0.119 |
| | (0.033) | (0.060) | (0.046) | (0.066) | (0.064) | (0.102) |
| Full Seri Polynomial | None | 1st Order | None | 1st Order | None | 2nd Order |
| Post-policy Polynomial | None | None | 1st Order | 1st Order | 2nd Order | 2nd Order |
| Observations | 108 | 108 | 108 | 108 | 108 | 108 |
| Pseudo $R^2$ | 0.011 | 0.0229 | 0.0235 | 0.0238 | 0.0408 | 0.0418 |
| Residual Std. Error | 1.000(df = 106) | 1.000(df = 105) | 1.000(df = 105) | 1.000(df = 104) | 1.000(df = 104) | 1.000(df = 102) |
| F Statistic | (df = 1.0; 106.0) | (df = 2.0; 105.0) | (df = 2.0; 105.0) | (df = 3.0; 104.0) | (df = 3.0; 104.0) | (df = 5.0; 102.0) |

*Note:* $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Figure 13: New Corruption Inquiries Poisson Regression




The mean predicted values for the last periods are lower than the pre-policy level for models (3), (5), and (6). However, they are not statistically significant. Nonetheless, in equation (3), the downward slope[51] is significant at the 99% confidence level, indicating a clear sign of an upward spike and further downward trend. This finding is consistent with Miller's hypothesis of increased crime detection and deterrence. Similarly, in the quadratic equation (5), the quadratic coefficient[52] is also significant at the 99% confidence level and negative, indicating that the function is concave downwards on period $T2$. Therefore, the curve of corruption detections has a maximum point after the policy shock, and then it decreases, as shown in the left plot in the last row of Figure 13. Although the analysis did not rule out the second hypothesis, it points to a spike, showing that there is no upward trend after the policy shock.

The above result holds true for other regression methods such as OLS and Negative Binomial regression. Notably, the Negative Binomial regression has more significant coefficients, as shown in Table 7. Additionally, OLS regression allows for the model to be regressed with higher order polynomials, which was not possible with Poisson and Negative Binomial regressions. However, parameters appear to be perfectly correlated for higher order parameters, which may suggest that OLS models could be overfitting the data[53].

[51] Poisson coefficient = -0.003

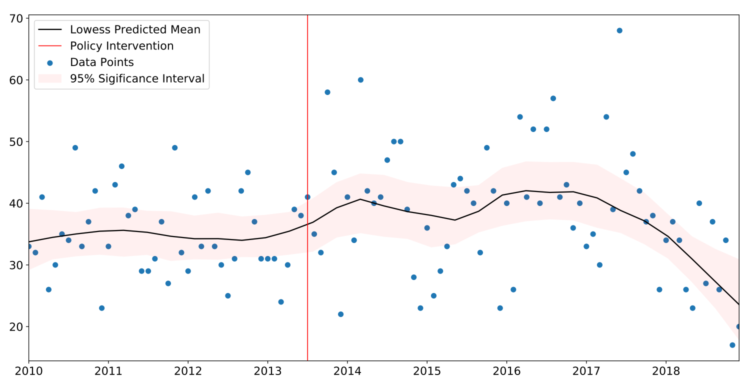
[52] Poisson coefficient = -0.000236

[53] All regressions and further methods are available at https://github.com/caxaxa/PhD_Empirics/blob/main/Data_Analysis.ipynb.

### 4.4.2 Locally Weighted Scatterplot Smoothing

Fitting a polynomial curve to noisy data can present some challenges. One known issue is that the predicted function may misbehave by predicting points outside of the data range, which can result in overfitting or "chasing noise." With the increase in computational power, newer methods such as local regressions are becoming more popular. One such method is locally weighted scatterplot smoothing, first introduced by Cleveland in 1979. Figure 14 displays the predicted LOWESS curve[54] over the Brazilian corruption inquiry data.

Figure 14: New Corruption Inquiries LOWESS




Note that, unlike the previous examples, the mean predicted value at the end of the entire period is statistically significantly lower than the pre-policy period. The immediate pre-policy period has a mean predicted value of 35.46 inquiries per month, with a lower bound[55] of 32.06 new inquiries per month. At the end of the period, the mean drops to 23.57 and the upper bound is 31.61. This represents an average drop of 33.53%. Therefore, it rejects $H_0$ from the second hypothesis, providing evidence of enhanced deterrence after the policy shock.

Conversely, there are no statistically significant peaks, although the mean predicted values increase before decreasing. In other words, none of the higher points is statistically different from the points in the pre-policy period[56].

In summary, the analysis without controls shows that the data is consistent with the hypothesis that the policy intervention was effective in better prosecuting and deterring corruption crimes. In the next subsection, the model is controlled for other factors that may affect corruption detection.

[54] Estimated using a fraction of 0.2 Available at https://github.com/caxaxa/PhD_Empirics/blob/main/Data_Analysis.ipynb

[55] With 95% confidence level.

[56] The results using daily data are similar, as shown in Figure 16

### 4.5 Regression With Controls

Crime literature suggests that potential criminals can be influenced by their environment Becker (1968). As this study uses macro-level data, it may be affected by other macroeconomic variables. Therefore, the regression is controlled for unemployment, which is expected to affect newly detected corruption crimes by the supply side, as rent and opportunities determine the willingness to commit crimes (Becker, 1968; Mauro, 1995). Additionally, GDP is a proxy for government budget, so a higher GDP could lead to more public contracts, corruption opportunities, and better-equipped public prosecution. Finally, as corruption is a rent-seeking activity, it is expected to be more prevalent when real interest rates are low. The results from Table 3 controlled for the above-mentioned variables are shown in Table 4.

The results show more significant coefficients, with models (3), (4), (5), and (6) indicating a lower predicted value at the end of the period. However, they are still not statistically significant. Nonetheless, the positive and significant coefficients for the dummy variable in all models provide further evidence of enhanced detection after the policy intervention.

To address the potential influence of past environments on current crime detection, a small robustness test is conducted by regressing the same equations with one-year lagged control variables. Results, shown in Table 8, reveal little difference, with the coefficients for constant, dummy, and real interest rates remaining similarly significant. This suggests that control variables have a more permanent effect on corruption crimes.

## 5 Conclusion

This paper attempts to address the issue of corruption under non-trial resolution policies, particularly policies that reduce sanctions for agents who collaborate with authorities. Given that corruption is often an unobservable phenomenon, this study proposes to model the dynamics of corruption crimes as a bribery game and predict observable variables such as criminal detection, specifically corruption inquiries.

To simulate predicted corruption detection, the study utilizes dynamic programming. The predicted variables are then compared against real-world data to test hypotheses about corruption behavior. Specifically, this paper reproduces the methodology presented in previous studies such as Miller (2009), Brenner (2009) and Berlin et al. (2018) to test the efficacy of sanction reduction interventions over crimes of corruption in Brazil. The study uses data on the number of monthly and daily inquiries on corruption from 2010 to 2020.

In the first part, the data is analyzed using Poisson regression and locally weighted scatterplot smoothing. Poisson regression is used to fit different order polynomial functions to the data to investigate the presence of any trends or spikes in corruption inquiries before and after the policy intervention.

The results of the Poisson regressions reveal that the policy intervention

[H]

Table 4: Poisson Regression with Controls

| | *Dependent variable:* | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Constant | 3.705*** | 3.619*** | 3.255*** | 3.213*** | 3.721*** | 3.726*** |
| | (0.108) | (0.109) | (0.132) | (0.158) | (0.243) | (0.300) |
| Dummy | 0.137** | 0.378*** | 0.350*** | 0.331*** | 0.243*** | 0.298** |
| | (0.068) | (0.086) | (0.076) | (0.086) | (0.090) | (0.116) |
| Unemployment | -0.012 | 0.014 | 0.041*** | 0.043*** | -0.022 | -0.026 |
| | (0.012) | (0.013) | (0.015) | (0.016) | (0.032) | (0.034) |
| Real Interest | -6.941** | -9.051*** | -7.960** | -7.686** | -9.928*** | -9.897*** |
| | (3.352) | (3.397) | (3.375) | (3.421) | (3.492) | (3.632) |
| GDP | -0.181 | 0.575 | 2.029** | 2.170** | 2.613** | 2.650** |
| | (0.915) | (0.926) | (0.988) | (1.029) | (1.017) | (1.050) |
| Full Seri Polynomial | None | 1st Order | None | 1st Order | None | 2nd Order |
| Post-policy Polynomial | None | None | 1st Order | 1st Order | 2nd Order | 2nd Order |
| Observations | 108 | 108 | 108 | 108 | 108 | 108 |
| $R^2$ | | | | | | |
| Adjusted $R^2$ | | | | | | |
| Residual Std. Error | 1.000(df = 103) | 1.000(df = 102) | 1.000(df = 102) | 1.000(df = 101) | 1.000(df = 101) | 1.000(df = 99) |
| F Statistic | (df = 4.0; 103.0) | (df = 5.0; 102.0) | (df = 5.0; 102.0) | (df = 6.0; 101.0) | (df = 6.0; 101.0) | (df = 8.0; 99.0) |
| *Note:* | | | | | | *p<0.1; **p<0.05; ***p<0.01 |

resulted in a consistent upward shift in the number of new corruption inquiries. This shift is statistically significant and, if accompanied by a subsequent decline in corruption detections, supports Miller's hypothesis of improved prosecutorial efficiency and increased deterrence against criminal activities. However, the findings do not provide sufficient evidence to substantiate the hypothesis of enhanced deterrence. Although the shifts are statistically significant, there is no statistically significant decrease in the level of corruption at the end of the period compared to the pre-intervention period.

Notably, locally weighted scatterplot smoothing is employed to fit a non-parametric curve to the data and examine any significant peaks or drops in the data. The results indicate that the average predicted value of new corruption inquiries at the end of the period is significantly lower than during the pre-policy period, suggesting that the policy intervention has indeed strengthened deterrence.

Taken together, both sets of results demonstrate that the policy was effective in combating corruption in Brazil during the analyzed period.

Furthermore, the regressions were controlled for other factors that may affect corruption detection, such as unemployment, GDP, and real interest rates. The results show that the policy intervention still led to an increase in the number of new corruption inquiries, and this shift is statistically significant. However, the coefficients of other control variables are not significant. The results also show that the dummy variable coefficients are still positive and significant for all models. Moreover, models (3), (4), (5), and (6) show a lower predicted value of new corruption inquiries at the end of the period, but the values are still not significant. Showing that, not only the results are resilient to outside effects, but they are more robust under such conditions.

In summary, the analysis presents evidence indicating that the policy intervention resulted in an increase in crime detection and deterrence, primarily driven by improvements in prosecutorial and judicial efficiency. The findings demonstrate that the enhanced deterrence is statistically significant and remains robust even when controlling for other factors that could potentially influence corruption detection.

# 6 Disclosure of data and computer code availability

The data used in this study was obtained from the Brazilian Public Prosecution's (MPF) online processual search engine, available at `http://apps.mpf.mp.br/aptusmpf/portal?servidor=portal` and downloaded between February and March 2020. The code I created for scraping the data is available at, `https://github.com/caxaxa/MPF_Crime_Scraper`. Other sources of external data are the quarterly Brazilian real GDP growth from '*Instituto de Pesquisa Economica Aplicada (IPEA)*', available at, `http://ipeadata.gov.br`. The monthly Brazilian unemployment rate which Before 2012, the source is' *Fundacao Sistema Es-*

*tadual de Analise de Dados (SEAD)*', available at, https://www.seade.gov.br/, after 2012, the data is from '*Instituto Brasileiro de Geografia e Estatistica (IBGE)*', available at, https://www.ibge.gov.br/estatisticas/sociais/educacao/9127-pesquisa-nacional-por-amostra-de-domicilios.html?=&t=series-historicas. The monthly average of the main national treasury bonds' real interest rate (SELIC), the nominal rate was gathered from the Brazilian Central Bank, available at, https://www.bcb.gov.br/estatisticas/txjuros, and deflated by the '*IPCA*' inflation index, available at, http://ipeadata.gov.br.

The datasets generated by the survey research during and/or analyzed during the current study are available in https://github.com/caxaxa/PhD_Empirics/tree/main/Data.

The application used for producing the theorethical part of the paper is available at the repository, https://github.com/caxaxa/Corruption_Game/tree/main, where the main code for running the simulations is in https://github.com/caxaxa/Corruption_Game/blob/main/Corruption_game_leniency_and_non_trial_resolution.ipynb. Lastly, the code for the empirical part of this study is available at https://github.com/caxaxa/PhD_Empirics/blob/main/Data_Analysis.ipynb.

# 7 Declarations

## Funding and Competing interests

The author declares to have no financial and non-financial interests.

The author was employed by the State Prosecution of Santa Catarina, Brazil, up until November 2022. However, this employment did not involve any activities related to anti-corruption taskforces associated with the Federal Public Prosecution or the Federal Police of Brazil within the time frame analyzed in this study.

This work was supported by the small grants scheme of the research initiative on Rebalancing Public and Private Interests and Erasmus Center of Empirical Legal Studies of Erasmus School of Law and the sector plan for law funding of the Ministry of Education, Culture and Research.

# References

Acconcia, A., Immordino, G., Piccolo, S., and Rey, P. (2014). Accomplice witnesses and organized crime: Theory and evidence from italy. *Scandinavian Journal of Economics*, 116(4):1116–1159.

Amir, E., Lazar, A., and Levi, S. (2018). The deterrent effect of whistleblowing on tax collections. *European Accounting Review*, 27(5):939–954.

Aubert, C., Rey, P., and Kovacic, W. E. (2006). The impact of leniency and whistle-blowing programs on cartels. *International Journal of Industrial Organization*, 24(6):1241–1266.

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 76(2):169–217.

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

Berlin, M. P., Qin, B., and Spagnolo, G. (2018). Leniency, asymmetric punishment and corruption: Evidence from china. *CEIS Tor Vergata Research Paper Series*.

Brenner, S. (2009). An empirical study of the european corporate leniency program. *International Journal of Industrial Organization*, 27(6):639–645.

Buccirossi, P. and Spagnolo, G. (2006). Leniency policies and illegal transactions. *Journal of Public Economics*, 90(6-7):1281–1308.

Burguet, R., Ganuza, J.-J., and Montalvo, J. G. (2016). The microeconomics of corruption: A review of thirty years of research. Working Paper.

Dufwenberg, M. and Spagnolo, G. (2014). Legalize bribe giving. *Economic Inquiry*, 52(4):1595–1607.

Harrington, J. E. (2008). Optimal corporate leniency programs. *The Journal of Industrial Economics*, 56(2):215–246.

Ljungqvist, L. and Sargent, T. J. (2012). *Recursive Macroeconomic Theory*. MIT Press, 3 edition.

Marvao, C. and Spagnolo, G. (2016). Cartels and leniency: Taking stock of what we learnt. Technical Report 39, Stockholm School of Economics.

Mauro, P. (1995). Corruption and growth. *The Quarterly Journal of Economics*, 110(3):681–712.

Miller, N. H. (2009). Leniency and cartel enforcement. *American Economic Review*, 99(3):750–68.

Motta, M. and Polo, M. (2003). Leniency programs and cartel prosecution. *International Journal of Industrial Organization*, 21(3):347–379.

Nations, U. (2004). United nations convention against corruption.

Pakes, A. and McGuire, P. (2001). Stochastic algorithms, symmetric markov perfect equilibrium, and the curse of dimensionality. *Econometrica*, 69(5):1261–1281.

Puterman, M. L. (2005). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons.

Soreide, T. (2018). Corruption and criminal justice: Bridging economic and legal perspective. In Arlen, J., editor, *The Research Handbook on Corporate Crime and Financial Misdealing*. Edward Elgar Publishing.

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

# A Appendix

## A.1 Simulation Model Variables

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

### A.1.1 Constants $\theta$

Given the following constants $\theta$ 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;
* $\delta$ is the liability depreciation or prescription;
* $\gamma$ is the time discount;
* $\eta$ is the agents' risk aversion;
* $y_0$ is an autonomous non-financial income;
* Probability of detection is $\alpha$; and
* Probability of conviction is $\beta$.

### A.1.2 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.1.3 Leniency rules

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

### A.1.4 A set of discrete actions or decisions (controls) $\mathbf{d}_{i,t}$

**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 consumption decision, is the decision of how much to save $w$, or buy financial assets.

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

### 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 succeeded 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 $(s_{nc}, s_{cor}, s_{det,i}, s_{col,i}, s_{con,i}, s_{acq,i}, s_{sr,i}, s_{pg,i}) \in S$;

### Wealth from Players $W$

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

### 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.1.6 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 dimension $(3 \times (\bar{W} * \bar{L} * 8))$.

* 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 $(2 \times (\bar{W} * 3))$.

## A.2 Supporting Plots and Tables

Figure 15: New Corruption Inquiries Multivariate Poisson Regression


Figure 16: LOWESS Daily Frequency

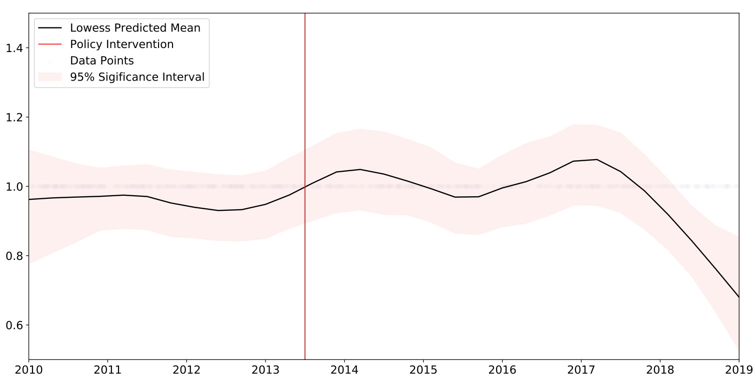


Table 5: Number of Investigations of Corruption Crimes per Brazilian State (Corruption, Embezzlement and Extortive Corruption)

| date<br>region | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018- | 2019 | 2020-12-31 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AC | 31.0 | 19.0 | 23.0 | 11.0 | 18.0 | 21.0 | 20.0 | 30.0 | 21.0 | 22.0 | 20.0 | 0.0 |
| AL | 9.0 | 3.0 | 26.0 | 22.0 | 22.0 | 13.0 | 24.0 | 28.0 | 30.0 | 31.0 | 16.0 | 0.0 |
| AM | 32.0 | 28.0 | 36.0 | 68.0 | 47.0 | 48.0 | 67.0 | 38.0 | 44.0 | 31.0 | 34.0 | 3.0 |
| AP | 34.0 | 31.0 | 47.0 | 53.0 | 48.0 | 23.0 | 30.0 | 35.0 | 35.0 | 28.0 | 16.0 | 0.0 |
| BA | 48.0 | 62.0 | 64.0 | 89.0 | 102.0 | 80.0 | 74.0 | 81.0 | 89.0 | 86.0 | 53.0 | 2.0 |
| CE | 24.0 | 39.0 | 57.0 | 60.0 | 76.0 | 69.0 | 64.0 | 48.0 | 51.0 | 90.0 | 63.0 | 0.0 |
| DF | 50.0 | 57.0 | 97.0 | 128.0 | 162.0 | 155.0 | 171.0 | 162.0 | 188.0 | 192.0 | 157.0 | 2.0 |
| ES | 0.0 | 17.0 | 16.0 | 17.0 | 26.0 | 14.0 | 18.0 | 33.0 | 33.0 | 38.0 | 19.0 | 1.0 |
| GO | 32.0 | 24.0 | 40.0 | 42.0 | 43.0 | 35.0 | 43.0 | 37.0 | 54.0 | 49.0 | 27.0 | 0.0 |
| MA | 22.0 | 31.0 | 24.0 | 41.0 | 37.0 | 53.0 | 57.0 | 42.0 | 65.0 | 77.0 | 58.0 | 1.0 |
| MG | 61.0 | 51.0 | 80.0 | 102.0 | 102.0 | 113.0 | 92.0 | 82.0 | 102.0 | 85.0 | 71.0 | 1.0 |
| MS | 23.0 | 36.0 | 55.0 | 78.0 | 84.0 | 81.0 | 84.0 | 73.0 | 51.0 | 43.0 | 26.0 | 0.0 |
| MT | 49.0 | 39.0 | 50.0 | 62.0 | 41.0 | 35.0 | 53.0 | 44.0 | 28.0 | 29.0 | 15.0 | 0.0 |
| PA | 28.0 | 34.0 | 76.0 | 68.0 | 45.0 | 50.0 | 84.0 | 82.0 | 67.0 | 64.0 | 32.0 | 5.0 |
| PB | 13.0 | 20.0 | 16.0 | 24.0 | 21.0 | 28.0 | 18.0 | 35.0 | 34.0 | 43.0 | 34.0 | 0.0 |
| PE | 32.0 | 29.0 | 43.0 | 69.0 | 65.0 | 51.0 | 46.0 | 56.0 | 57.0 | 39.0 | 23.0 | 1.0 |
| PI | 18.0 | 24.0 | 36.0 | 35.0 | 40.0 | 38.0 | 32.0 | 23.0 | 42.0 | 30.0 | 26.0 | 0.0 |
| PR | 69.0 | 103.0 | 147.0 | 143.0 | 166.0 | 75.0 | 78.0 | 41.0 | 60.0 | 25.0 | 50.0 | 2.0 |
| RJ | 74.0 | 74.0 | 113.0 | 124.0 | 82.0 | 82.0 | 90.0 | 103.0 | 83.0 | 108.0 | 101.0 | 4.0 |
| RN | 13.0 | 14.0 | 36.0 | 24.0 | 43.0 | 52.0 | 43.0 | 32.0 | 52.0 | 43.0 | 29.0 | 0.0 |
| RO | 16.0 | 12.0 | 23.0 | 20.0 | 23.0 | 30.0 | 23.0 | 28.0 | 22.0 | 18.0 | 22.0 | 0.0 |
| RR | 10.0 | 17.0 | 26.0 | 21.0 | 31.0 | 24.0 | 18.0 | 19.0 | 35.0 | 34.0 | 29.0 | 1.0 |
| RS | 12.0 | 35.0 | 50.0 | 56.0 | 54.0 | 65.0 | 76.0 | 69.0 | 75.0 | 54.0 | 51.0 | 2.0 |
| SC | 28.0 | 49.0 | 67.0 | 56.0 | 56.0 | 37.0 | 34.0 | 33.0 | 25.0 | 42.0 | 32.0 | 0.0 |
| SE | 2.0 | 6.0 | 12.0 | 18.0 | 8.0 | 14.0 | 9.0 | 17.0 | 15.0 | 26.0 | 23.0 | 0.0 |
| SP | 107.0 | 103.0 | 131.0 | 132.0 | 129.0 | 151.0 | 183.0 | 182.0 | 189.0 | 162.0 | 117.0 | 2.0 |
| TO | 52.0 | 24.0 | 31.0 | 23.0 | 28.0 | 26.0 | 20.0 | 34.0 | 30.0 | 20.0 | 22.0 | 0.0 |
| Sum | 889.0 | 981.0 | 1422.0 | 1586.0 | 1599.0 | 1463.0 | 1551.0 | 1487.0 | 1577.0 | 1509.0 | 1166.0 | 27.0 |

Table 6: Number of Annual Brazilian Investigations per Crime

| Date | Total | Corruption | Embezz-lement | Environ-mental | Extortive Corruption | Swindle | Theft | Drugs | Against Property | Financial Financial | Authority Abuse | Procure ment Fraud |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2009 | 84612 | 521 | 734 | 16 | 52 | 3217 | 2695 | 1835 | 9674 | 764 | 101 | 775 |
| 2010 | 79977 | 569 | 700 | 35 | 67 | 3923 | 1745 | 1960 | 8773 | 938 | 101 | 612 |
| 2011 | 81835 | 867 | 1088 | 179 | 77 | 9081 | 3969 | 2327 | 4051 | 1267 | 161 | 1026 |
| 2012 | 76483 | 1103 | 1099 | 186 | 74 | 12060 | 4663 | 2036 | 975 | 1089 | 145 | 1080 |
| 2013 | 85489 | 1191 | 1109 | 234 | 50 | 17926 | 4637 | 4907 | 801 | 1155 | 105 | 1323 |
| 2014 | 90158 | 985 | 1108 | 273 | 70 | 23121 | 7106 | 4645 | 505 | 1116 | 105 | 1180 |
| 2015 | 84525 | 1137 | 1177 | 533 | 62 | 18213 | 7249 | 4662 | 287 | 1289 | 73 | 1073 |
| 2016 | 80539 | 1088 | 1191 | 291 | 53 | 19022 | 7060 | 3906 | 254 | 1050 | 86 | 1125 |
| 2017 | 80188 | 1266 | 1241 | 236 | 61 | 18626 | 7346 | 3579 | 193 | 1074 | 75 | 1339 |
| 2018 | 76688 | 1073 | 1263 | 399 | 49 | 17522 | 6625 | 4012 | 177 | 968 | 71 | 1272 |
| 2019 | 62349 | 743 | 890 | 317 | 20 | 11078 | 3593 | 2758 | 63 | 787 | 45 | 894 |

Table 7: Negative Binomial Regression

| | *Dependent variable:* | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| cons | 3.549*** | 3.609*** | 3.549*** | 3.573*** | 3.549*** | 3.533*** |
| | (0.038) | (0.047) | (0.037) | (0.073) | (nan) | (nan) |
| dummy | 0.101** | 0.259*** | 0.210*** | 0.236** | 0.047*** | 0.119*** |
| | (0.048) | (0.087) | (0.068) | (0.095) | (nan) | (nan) |
| t2 | | | -0.003** | -0.002 | 0.012*** | 0.019*** |
| | | | (0.002) | (0.003) | (nan) | (nan) |
| s2 | | | | | -0.000*** | -0.000*** |
| | | | | | (nan) | (nan) |
| Full Seri Polynomial | None | 1st Order | None | 1st Order | None | 2nd Order |
| Post-policy Polynomial | None | None | 1st Order | 1st Order | 2nd Order | 2nd Order |
| Observations | 108 | 108 | 108 | 108 | 108 | 108 |
| $R^2$ | | | | | | |
| Adjusted $R^2$ | | | | | | |
| Residual Std. Error | 1.000(df = 106) | 1.000(df = 105) | 1.000(df = 105) | 1.000(df = 104) | 1.000(df = 104) | 1.000(df = 102) |
| F Statistic | (df = 1.0; 106.0) | (df = 2.0; 105.0) | (df = 2.0; 105.0) | (df = 3.0; 104.0) | (df = 3.0; 104.0) | (df = 5.0; 102.0) |
| *Note:* | | | | | | *p<0.1; **p<0.05; ***p<0.01 |

Table 8: Poisson Regression with Lagged Variables

| | *Dependent variable:* | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Constant | 3.559*** | 3.538*** | 3.187*** | 3.135*** | 3.742*** | 3.745*** |
| | (0.111) | (0.110) | (0.126) | (0.151) | (0.253) | (0.312) |
| Dummy | 0.099 | 0.358*** | 0.350*** | 0.325*** | 0.238*** | 0.295** |
| | (0.067) | (0.088) | (0.078) | (0.087) | (0.090) | (0.116) |
| Unemployment | -0.000 | 0.021 | 0.048*** | 0.050*** | -0.026 | -0.030 |
| | (0.012) | (0.013) | (0.015) | (0.015) | (0.033) | (0.036) |
| Real Interest | -0.785 | -5.463 | -6.214* | -5.921* | -9.771*** | -9.737** |
| | (3.319) | (3.493) | (3.456) | (3.486) | (3.736) | (3.901) |
| GDP | -0.050 | 0.593 | 2.135** | 2.314** | 2.844*** | 2.888*** |
| | (0.919) | (0.927) | (0.989) | (1.028) | (1.023) | (1.052) |
| Full Seri Polynomial | None | 1st Order | None | 1st Order | None | 2nd Order |
| Post-policy Polynomial | None | None | 1st Order | 1st Order | 2nd Order | 2nd Order |
| Observations | 108 | 108 | 108 | 108 | 108 | 108 |
| $R^2$ | | | | | | |
| Adjusted $R^2$ | | | | | | |
| Residual Std. Error | 1.000(df = 103) | 1.000(df = 102) | 1.000(df = 102) | 1.000(df = 101) | 1.000(df = 101) | 1.000(df = 99) |
| F Statistic | (df = 4.0; 103.0) | (df = 5.0; 102.0) | (df = 5.0; 102.0) | (df = 6.0; 101.0) | (df = 6.0; 101.0) | (df = 8.0; 99.0) |

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