

# Corruption and the effects of influence within social networks: An agent-based model of the “Lava Jato” scandal.

Amira Al-Khulaidy and Valentin Vergara

July 24, 2019

## Abstract

Corruption, and more specifically corruption in Latin America, is a complex phenomenon which is affected by politics, social structures and institutions, as well as individual behaviors. The *Lava Jato* is a large-scale example of corruption in Brazil. As of early 2019, there is still an ongoing investigation surrounding what has been lauded as the largest corruption scheme in Latin America. Advances in data analysis, computation, and social networks have allowed progress to be made with these types of investigations. The *Lava Jato* case has been a clear example of how breaking up social networks and understanding the extent of crime and individual corruption has revealed webs of corruption that have influenced politics, as well as hindered economic development in Brazil. The several layers of interactions between individuals and institutions can be difficult to grasp and understanding the patterns and relationships within complex large-scale phenomena, such as corruption can seem impossible. Agent-based models can help with understanding these complex behaviors and systems. By capturing the patterns and gaining a better understanding of how corruption emerges and is manifested, we can help inform policy, as well as create better tools and methods for crime prevention and detection.

## 1 Introduction

Corruption in Latin America has been a topic of interest across many fields. The historical, social and political roots of Latin America are deeply tied to the origins of the socio-political systems that govern many Latin American countries (Weyland, 1998). Latin America has been at the forefront of political and economic discussions for decades, however, more recently, several webs of events have emerged, and the effects are long reaching with more complex ramifications. Corruption tends to focus on financial and political power, but in Latin America, there are several aspects that make

the topic of corruption more complex in terms of the financial and legal intricacies (Goldstein and Drybread, 2018). This paper is focused specifically on corruption in Latin America, and more specifically on the *Lava Jato* corruption scheme and in Brazil.

The reach and intermingled web of the *Lava Jato* scheme has had global implications, with many contributing factors. The complex interactions of individuals, institutions, and transactions can be better explored in an agent-based model (ABM) with a focus on the social networks of the agents. ABMs and social network analysis (SNA) allow us to capture the patterns from data to build a model, and then run the simulation to understand the patterns and results of the phenomenon we are investigating. From the model we can better capture what are the more prominent patterns, to gain a better understanding of how the different processes interact (Gilbert and Troitzsch, 2005).

By building a social network and running the model through time, we can derive a better understanding of how the *Lava Jato* corruption scheme occurs, and how the impact of the various factors give rise to corrupt agents within a network. To better understand these factors, we will look at the different players and how their relationships enable this level of corruption based on whether agents engage in corruption, as well as their risk aversion which represents whether or not they engage in criminal behavior. We aim to look at how the social networks are formed and if different social networks have different outcomes in terms of the frequency of corrupt behavior.

The importance of understanding corruption in a dynamic way is that law enforcement entities can experiment with different ways of targeting corruption. Ultimately, models and experiments that can be done for corruption detection, as well as understanding patterns for prevention can help with the social inequities that arise from misdistribution of funds and abuses of power.

## 2 Background

*Operacao Lava Jato*, or "operation car wash" is an ongoing investigation by the Federal Police of Brazil in the Southern part of Brazil (Federal, 2018). Beginning in 2014 and currently (2019) still ongoing, the scale of corruption has ranged from elected officials, to private companies, as well as cartels and off-shore banks- which has provoked public uproar, protests and political upheavals in Brazil (Goldstein and Drybread, 2018). The initial investigation, which gave origin to the title of the operation was set to focus on *doleiros*, black market money dealers, who used petrol stations and car washes to launder the profits of crime. The interactions of individual agents with criminal activity from the "bottom-up" gave rise to the emergence of one of the largest, if not largest corruption

scandal in Latin America, if not the world (Angelico, 2017).

Several other investigations related to “Lava Jato” came to the forefront, including the beef and poultry industries. Sanctions were placed by other countries on imports from Brazil, which led to a cycle of effect that impacted consumer options and choices in homes around the world, which in return has resulted in an impact on the Brazilian Stock Market (Araújo et al., 2018). Sergio Moro, the federal judge responsible for the persecutions states that these types of crimes “represent a challenge for law-enforcement agencies” and that the level sophistication and the lack of tools and resources to deal with these types of crimes have allowed for years of corruption to go undetected and unpunished (Moro, 2018).

While the investigation has yielded trials and arrests, there has been controversy surrounding the methods used for investigation and non-traditional means of persecuting these crimes (Lopes, 2016). Previous scandals have also come to light, such as the *Mensalao* (also known as Case 470) in 2002, but the results of previous investigations have not resulted in the scale and outcomes of the *Lava Jato* investigation. The reason has been a delay in the justice system which make it harder for people of power to be tried (Moro, 2018). The defendants from this case have been politicians, CEOs, judges, cartels, and large corporations. Politicians who engage in corruption through their individual actions do impact policy, and political outcomes- “corruption can change the alignment of citizen and leader preference” (Besley, 2006).

Corruption in Brazil has its origins in the history and political narrative of the country (Lopes, 2016). More recently, the *Mensalao* or Case 470 had been a controversial investigation where the vulnerabilities of the Brazilian legal system had come to the forefront. Due to the slow and bureaucratic nature of the legal system, *Foro Privilegiado* required all political figures and elected public servants to be trailed exclusively by the Supreme Court (Moro, 2018). This delayed the trials by a decade, which in addition to the corruption and political clout that exempted many from actually being trialed, corruption continued on with impunity. With the *Lava Jato*, there have been changes to the justice system, focusing on speedier trials, repatriation, as well as quicker arrests of those involved in corruption and crimes. The openness and publicity involved in these entails, as well as the open resources online for denouncing crime have allowed us to have more data in terms of who the corrupt people are, as well as their network and how all the people connect to one another (Angelico, 2017).

As we gain a better understanding of the linkages between the major players in the *Lava Jato* case, we can better understand the patterns of corruption in this case, and how individual choices

from the bottom-up have created large scale corruption by looking at individual choices with an understanding of human psychology and sociology (Besley, 2006). Why and how people choose to commit or not commit crimes and engage in corruption is a complex issue, but many psychologists (Milgram, 1963; Raven and French, 1958) have looked at group behavior and power dynamics to better understand how people act and react within a social context, as well as tipping points of when people decide to adopt a behavior or not (Centola et al., 2005). The “infectious” nature of how behaviors are adopted in social settings is the aim of this paper, and we hope to expand some of these thoughts to examine the corruption of individuals involved in the *Lava Jato* scheme through modeling this behavior.

Specifically, an agent-based model can help understand the patterns of these relationships, and theories or how power and influence can have a larger effect on individual choices. The more powerful an individual, the more power it holds in a social web. In the specific case of *Lava Jato*, we have the different agents, and their relationship to other agents. We also know which agents are more connected given the arrests from the investigation and the data from our proceedings (Garay et al., 2018).

The reach of corruption has been extensive, and the effects on political parties and campaigns have also tainted Brazilian politics and the trust of the people in the government. It would be impossible to reflect all of these different aspects in a model, but the focus of this paper is to understand the patterns of the social network and the *influence* (Centola et al., 2005) of people on their peers, and whether they are more or less likely to commit crimes. Categories used includes the public sector, which includes public employees of any government institution; private sectors, referring to those people who carry out licit commercial activities outside the public sphere; full-time criminals, which covers those nodes/agents who regularly carry out activities defined as criminal offenses; and financial institutions and which groups agents/nodes belong to the financial system. These classifications allow us to better understand the organizational and institutional role of each agent, in addition to the particular characteristics of the structure of the network (Garay et al., 2018).

### 3 A model for corruption

#### 3.1 General description

The model we developed tries to recreate how corrupt behavior spreads in a network structure, and how the cumulative effects of corruption cases discovered by the public opinion behave in a reasonable period of time. We based some our assumptions and most of the model from the recent corruption cases in Brazil (Garay et al., 2018).

Past models have looked at corruption within government with an analysis of corruption and the effects on the effectiveness of Rule of Law (Guerrero and Castañeda, 2019). Others have examined the micro and macro-level interactions of corruption with agent-based models within a game theory context (Hammond, 2000). Our model hopes to look at the spread of corruption, specifically through a lens of influence of corruption on the agents within a network. We model the spreading of corruption in a network structure and describe with more detail the different networks that we will create and the processes underlying their changes. Then, we describe what our model does and the reasons behind our assumptions.

#### 3.2 Generation of networks

For this model we used custom code, since we had to create networks where the nodes are the agents in our model. Examples of the networks chosen are in Figure 1. The first network is a Random Network (Gilbert, 1959). In an undirected network with  $n$  nodes, for every one of the possible  $\binom{n}{2}$  links, there is a probability  $p$  of creating that edge. The density of this network approaches  $p$ . We then look at the *Small World* network (Watts and Strogatz, 1998). We started with a ring network where every node has degree 4<sup>1</sup>. Then, for every link  $(i, j)$  we broke it and created another link  $(i, k)$  with probability  $p$ , where  $k$  is a node chosen at random.

The network on the right of Figure 1 is a Homophily Dynamic one, based on Murase et al. (2019). It is assumed that every node has some number of attributes  $f \in \{1, 2, \dots, F\}$ . Those attributes can take  $q$  different values, where  $2 \leq q \leq Q$ . As  $F$  is increased, the social system becomes more complex, and as  $Q$  increases, the individuals become more heterogeneous. At every time step, the edges in the network may change based on the following three processes:

**Global Attachment (GA)** A node selects one of its  $F$  attributes and chooses at random from

---

<sup>1</sup>This means that every node is connected to the adjacent 2 in the "left" and in the "right", relative to its position in the ring.

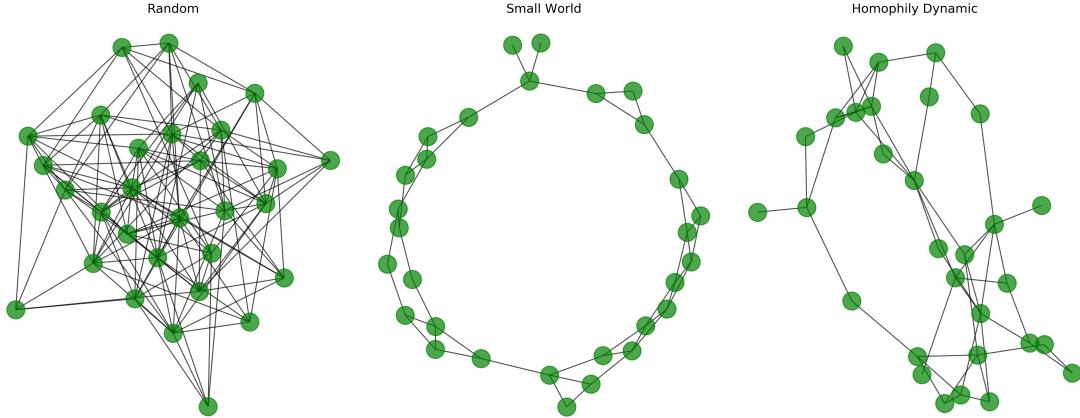


Figure 1: Three network types used in our model.

the subset of nodes with the same  $q$  value in that attribute. They establish a link.

**Local attachment (LA)** A node  $i$  with  $k_i > 0$  chooses one of its neighbors,  $j$ , who in turn, chooses one of its neighbors,  $k$  such that  $f_j = f_k$ . Then, if  $i$  and  $k$  are not connected, a node is created.

**Link deletion (LD)** Every element in the edge set is removed with probability  $p_d$ .

For the purpose of this research, we only used one attribute of the nodes as the *homophily* variable : the political party of every node, choose at random at the initialization of the model with a uniform distribution between  $A$ ,  $B$ , and  $C$ .

## 4 Attributes of the model

Our model places the agents in a network, with the following attributes.

### 4.1 Agents

In this artificial social system, agents can have one of three possible roles (Parsons, 1991): they can belong to the private, public or criminal sector, which represents three of the major aspects of the *Lava Jato* case. We chose the proportion of agents in each of the roles as 0.7, 0.19, and 0.11, respectively, as these proportions are based on empirical data from Garay et al. (2018).

During the model, agents can be **corrupt** or not. This means that they engage in an illicit transaction where they get a payoff. We let the corrupt status of every agent be a random variable resulting on a Bernoulli trial with probability  $c$  of being corrupt.

Since every agent can be, at different parts of the model, in both ends of an illicit transaction, we created two attributes for this. The first one is a **transaction size** value, when the agents proposes the illegal exchange. This is the result of a random discrete variable with uniform distribution with domain [100, 1000]. What this variable represents is how big the transaction is, in terms of an abstract currency unit. The second attribute puts the agent at the other end and corresponds to the **risk aversion**, which is a random variable that draws from different distributions depending on the argument passed. The possible arguments are:

'u' Uniform Risk Aversion. Uniform distribution between 0 and 1.

'l' Low risk aversion.  $\beta(2, 5)$  distribution. Biased to the left.

'm' Medium risk aversion.  $\beta(2, 2)$  distribution. Symmetrical centered around 0.5.

'h' High risk aversion.  $\beta(5, 2)$  distribution. Biased to the right.

Finally, agents keep track of their **political party** (random variable that chooses between *A*, *B*, or *C*, which were abstracted due to the many different political parties in Brazil); whether they are under **investigation** for corruption or not; and the total amount of payoff obtained from illegal transaction since the first time-step.

## 4.2 Initialization of the model.

When the model initializes, we create a **list** of the agents, with the proportions described above. After that, we take that list and make every item in it the node in a **network** using either a random, small world or homophily dynamic structure. At the beginning of every time step, we updated the network using a random rewiring or a homophily based one (with the GA, LA and LD processes in sequential order).

## 4.3 Interactions

Every time step is a week  $w$ , which is chosen due to the nature of the transitions carried out and captures the frequency of the transactions based on all the interacting agents. Before any interactions, we choose a random number of transaction attempts  $\tau$ . Then, for every one of those, we choose a random node  $i$  that will try to make an illegal transaction with node  $j$ . In case  $j$  is corrupt, the transaction happens and  $j$  gets as a payoff 5% of the transaction value (transaction

size). In case  $j$  is not corrupt, it can accept the exchange with probability  $(1 - \text{risk aversion})$ . If the transaction is successful,  $j$  changes its status to corrupt and updates its payoff.

#### 4.4 Corruption punishment.

After all the transaction attempts have been made, the model checks the current value of the payoff for every agent. If it exceeds some value  $p^*$ , the agent gets caught with a probability  $p_c$ . When this happens, the caught agent goes to jail (is removed from the agent set and from the network) and all of its neighbors change their status to "under investigation" and increase their risk aversion.

#### 4.5 New agents enter the model

We replace every agent removed from the model, keeping the same proportion between criminal, public and private sectors. Since at the beginning of each  $w$  we will update the network, these "new agents" get connected with another node depending of the specific network structured used to do the update.

### 5 Model Experiments and Results

#### 5.1 Baseline parameters

To begin testing our abstract model, we established a set of parameters (in Table 1) with enough variability that will produce noticeable results, against which we will compare different variations.

Table 1: Baseline parameters of the model

	Initial values
Number of nodes	500
Number of transaction attempts per week	[20, 30]
Amount of payoff until agent gets investigated	100
Probability of getting caught	0.1
Proportion of corrupt agents	0.2
Distribution of risk aversion	'u'
Network update process	GA, LA and LD

Based on these parameters, we created three networks, whose initial configuration was a random,

small world and homophily network; and then let the model run for 52 weeks to capture the cycle of different transitions that occur within the private, public, and criminal sectors, and with an update process that uses GA, LA and LD (Murase et al., 2019) each week. At the end, the resulting networks are in Figure 2. As expected, the random network looks like a hairball and depicts an unreal scenario where every node has random links.

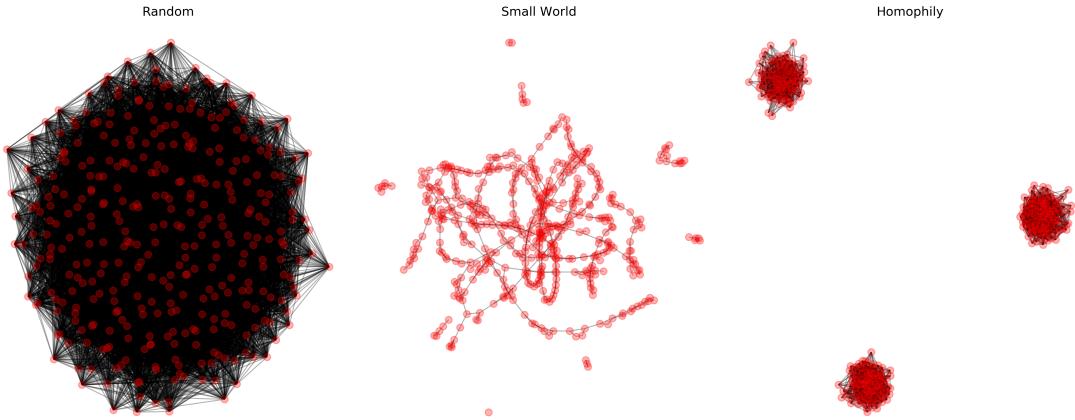


Figure 2: Networks produced with baseline parameters at  $w = 52$ .

The network in the middle, the small world, has some characteristics that could be useful for the purposes of this research. For example, two nodes at random have a shortest path  $L$  that is proportional to  $\ln n$ , where  $n$  is the number of nodes. In our example, this mean that from one node, it is easy to reach any other in the network. However, the fact that the nodes can easily reach the other does not mean that they can easily influence each other and spread corrupt behavior. That is the reason why we included the networks in the right, where every node will influence others through *homophily*, defined here as the political party. The three distinct subgroups in the plot are the three political parties used in our model.

Besides the obvious differences in the shape of the three networks, they have different values in some of the usual measurements in networks. This differences can be closely inspected in Table 2.

To understand what happens inside these networks, over the one-year period, there are a few things to notice. First, the number of illegal transactions is depicted in Figure 3 (top), where clearly we can see that in the small world network it remains stable, which is explained due to the low average degree compared to the other two networks. Since the nodes have less neighbors in average, they can *ask* to do an illegal transaction to less people at every time step.

The second thing to look is the number of corrupt agents that get caught and are sent to jail.

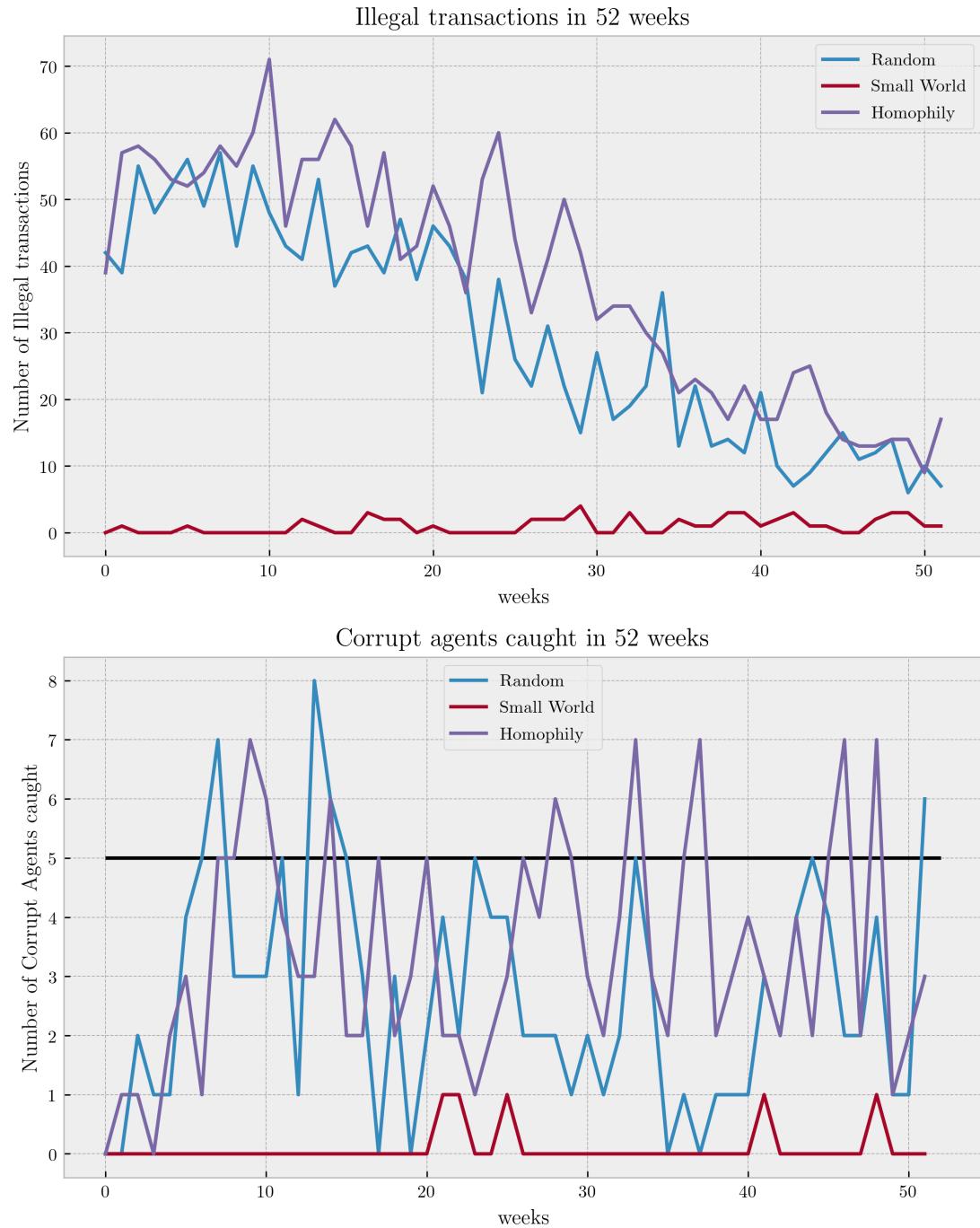


Figure 3: Number of illegal transactions, baseline parameters.

Table 2: Differences in network measurements at  $w = 52$ , baseline parameters.

	Random	Small World	Homophily
Number of nodes	330	491	306
Number of links	9525	741	2467
Average Degree	57.7273	3.0183	16.1242
Average Clustering	0.1741	0.3492	0.1644
Average Path Length	1.8246	6.2859	1.9136

Figure 3 (bottom) shows these results. We included a black horizontal line at 5 people caught. The reason is that in a given week, if more than 5 people get caught and are sent to jail for corruption, the whole social system will be on alert and could mean the beginning of what could resemble a corruption scandal.

Both plots coincide in that the small world network has the fewest people involved in illegal transactions, and therefore, the fewest people caught. In the arbitrarily threshold we set, there is also evidence to support that homophily network will lead to more corruption scandals than the random network. In the homophily network nodes tend to influence other in their political party, and thus, when the corrupt agents are caught, it is easier to uncover the corruption acts of their neighbors.

## 5.2 Modifying the network update mechanism.

In the baseline model, at every time step (every week) the network gets updated by a sequence of the GA, LA and LD processes. We tried running the three models in Figure 3 by changing the network update process while all the other conditions remained the same. Instead of the sequence GA, LA, and LD, we updated the networks through a random process, where at every time step we broke all links and created new ones with probability  $p$ .

The results in Figure 4, as expected, show little difference between the three types of networks. All three are affected by the risk aversion increase at more or less the same rate, because every week new networks are created.

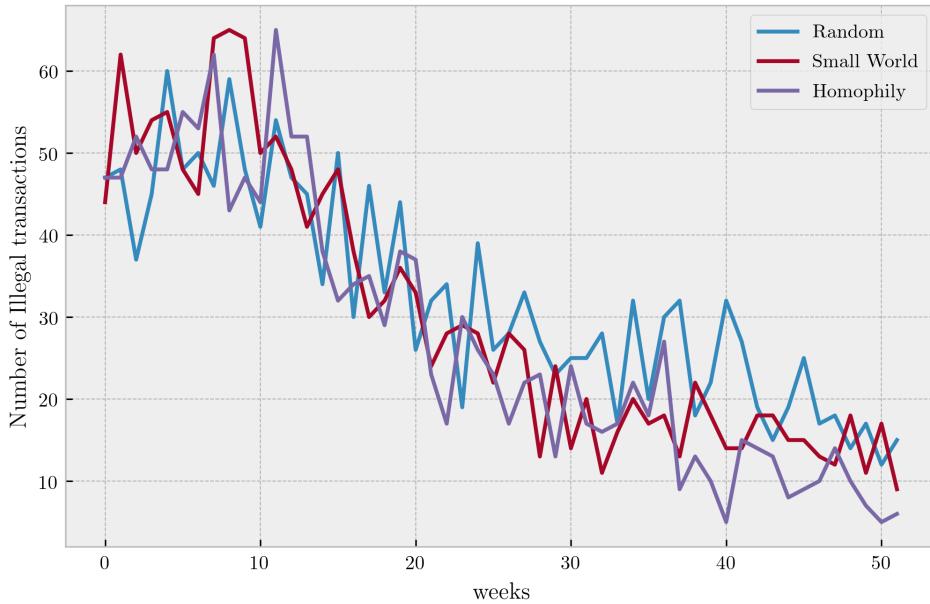


Figure 4: Number of illegal transactions, with a random updating scheme.

### 5.3 Effects of risk aversion

Looking at our results so far, we decided to focus only on the homophily networks, since they capture more closely how people create political contacts. This is the reason why Figure 5 shows the effects of different distributions of risk aversion only for this type of network.

### 5.4 Different starting proportions of corrupt agents

The final test we performed on the parameters was to see how different proportions of corrupt agents affect the outcome of the model. We used the number of illegal transaction per week as our outcome; with 1, 20, 40, 60, 80, and 99% of corrupt agents. The results in Figure 6 reproduce what was found by modifying the risk aversion distribution. However, in case of the proportion of corrupt agents, the extreme case where 99% of agents are corrupt maintains a higher number of illegal transactions, regardless of how much time has passed, and also appears to decrease at the same rate.

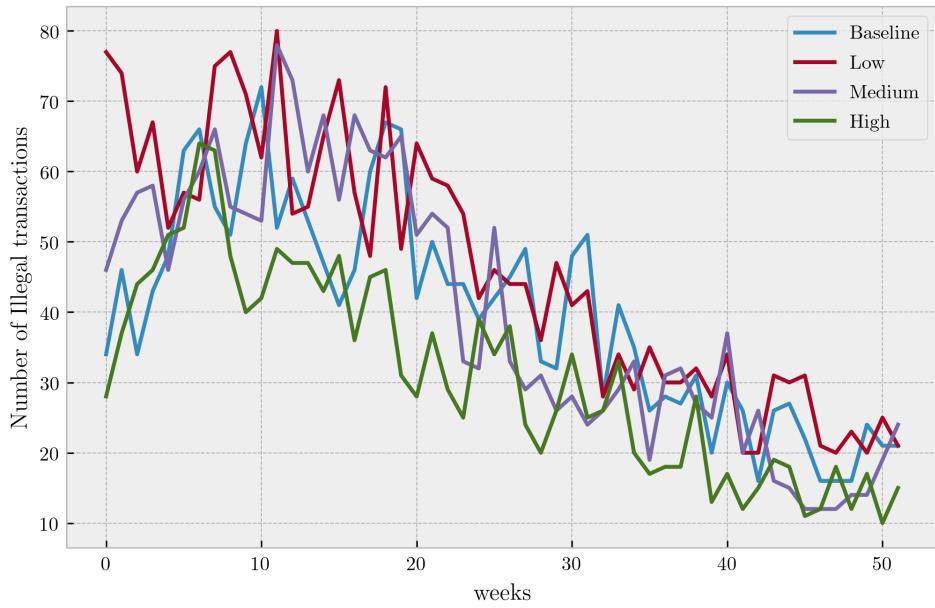


Figure 5: Different settings for risk aversion, homophily network

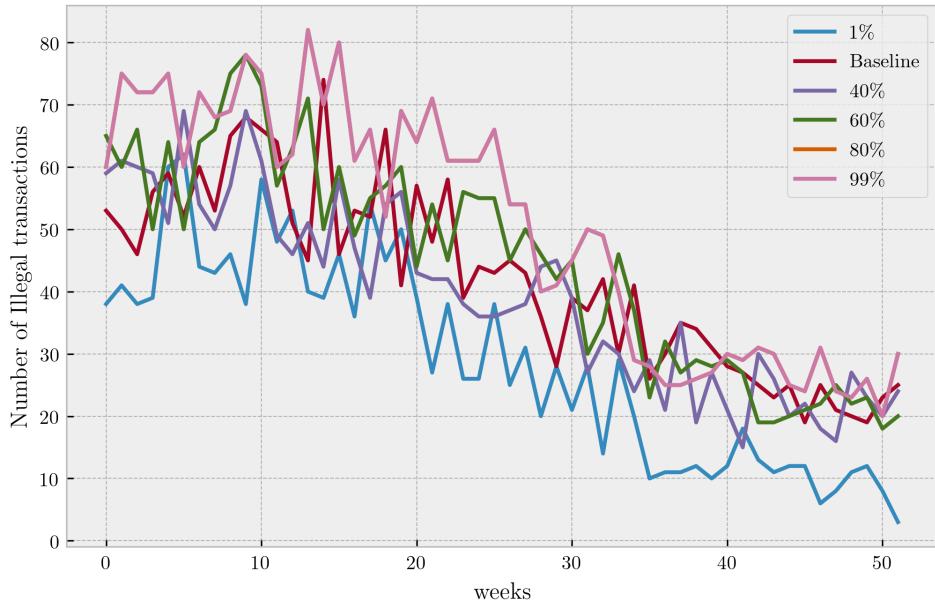


Figure 6: Different setting for risk aversion with different percentages of corrupt agents, homophily network.

## 6 Conclusion

As we mentioned earlier, a closer look at the linkages between the major players in the *Lava Jato* case, allow us to better understand the patterns of corruption in this case, and how individual choices from the bottom-up can have created large scale illegal activities. We learn from the results of our model that small world networks are the ones where the illegal transactions are low and stable. Therefore, what happened in Brazil clearly was not a small world, and it also wasn't random. This helps show that we have evidence to think that there was a homophily process in the formation of the corruption network. Homophily networks as a more specific topic within this corruption scandal would also be an interesting investigation in future works, particularly with new data and networks that are revealed as the investigation develops.

This is a theoretical model based on the details we have of the *Lava Jato* investigation thus far, and it does show similar behavior to the known networks that have already been revealed. Despite what we know of the investigation, there is limited data for us to thoroughly validate this model. Our results are robust and conceptually represent the way in which corruption occurs and has an effect on others within a network, but real data could provide more information about the structures of the networks, and the motivations and decision-making processes of the individuals engaged in corruption. However, even with additional data, we would at most have the illegal transactions that occurred, but not the attempted ones- which may be another benefit of simulating these types of behaviors to understand corruption.

While our results are consistent, further experimentation with empirical data and different model assumptions may produce variations of these results. Gaining a better understanding of the linkages between the major players can also help with targeted corruption detection and prevention. In further works, we can experiment with targeting links and targeting nodes to understand what types of strategies can be effective in cases of large-scale corruption, as well as experimenting with removing nodes from the network, targeting specific types of agents (highest degree, centrality, etc) to understand the influence and relationship. By building an agent-based model to look at the frequencies of agents committing crimes and engaging in corruption, we can explore the different thresholds and understand what types of effects result in higher proportion of agents committing crimes, as well as understanding if some types of social networks are more or less vulnerable to crime detection.

Overall this model serves as a theoretical representation of corruption within the case of the *Lava*

*Jato* scandal and understanding these networks and effects can help enhance our understanding of individual behaviors with large-scale impacts. Modelling can show clearer patterns and underlying networks, which can lead to better policies that can be implemented to combat corruption in the early stages.

## References

- Angelico, Fabiano. 2017. "BRAZIL: Open Data Just Made Investigating Corruption Easier."
- Araújo, Eduardo Carvalho Correa De, Victor Rangel Dos Santos Rodrigues, Danilo Soares Montemor, and Rogério Dias Correia. 2018. "Corrupção E Valor De Mercado: Os Efeitos Da Operação Lava Jato Sobre O Mercado De ações No Brasil." *REVISTA CATARINENSE DA CIÊNCIA CONTÁBIL* 17:nl.
- Besley, Timothy. 2006. *Principled agents? : the political economy of good government*. Oxford New York: Oxford University Press.
- Centola, Damon, Robb Willer, and Michael Macy. 2005. "The Emperor's Dilemma: a Computational Model of Self-enforcing Norms." *American Journal of Sociology* 110:1009–1040.
- Federal, Policia. 2018. "Operacão Lava Jato."
- Garay, Jorge, Guillermo Macías, Nathalia Guerra, Laura Rojas, Ernesto Navarro, Diana Santos Cubides, and Eduardo Salcedo-Albarán. 2018. "The "Lava Jato" Network: Corruption and Money Laundering in Brazil." Technical report. The Global Observatory of Transnational Criminal Networks – Working Paper No. 29.
- Gilbert, E. N. 1959. "Random Graphs." *The Annals of Mathematical Statistics* 30:1141–1144.
- Gilbert, Nigel and Klaus Troitzsch. 2005. *Simulation for the social scientist*. Maidenhead, England New York, NY: Open University Press.
- Goldstein, Donna M. and K. Drybread. 2018. "The Social Life of Corruption in Latin America." *Culture, Theory and Critique* 59:299–311.
- Guerrero, Omar A. and Gonzalo Castañeda. 2019. "Does Better Governance Guarantee Less Corruption? Evidence of Loss in Effectiveness of the Rule of Law." *CoRR* .

Hammond, Ross. 2000. “Endogenous Transition Dynamics in Corruption: An Agent-Based Computer Model.” Technical report. Center on Social and Economic Dynamics Working Paper No. 19.

Lopes, Arthur S. 2016. “The Hermeneutics of Brazil’s Impeachment.” *Harvard International Review* 37:11–14.

Milgram, Stanley. 1963. “Behavioral Study of Obedience.” *The Journal of Abnormal and Social Psychology* 67:371–378.

Moro, Sérgio Fernando. 2018. “Preventing Systemic Corruption in Brazil.” *Daedalus* 147:157–168.

Murase, Yohsuke, Hang-Hyun Jo, János Török, János Kertész, and Kimmo Kaski. 2019. “Structural Transition in Social Networks: the Role of Homophily.” *Scientific Reports* 9:4310.

Parsons, Talcott. 1991. *The Social System*. Routledge Sociology Classics. Routledge, new edition.

Raven, Bertram H. and John R. P. French. 1958. “Group Support, Legitimate Power, and Social Influence1.” *Journal of Personality* 26:400–409.

Watts, Duncan J. and Steven H. Strogatz. 1998. “Collective Dynamics of ‘small-World’ Networks.” *Nature* 393:440–442.

Weyland, Kurt Gerhard. 1998. “The Politics of Corruption in Latin America.” *Journal of Democracy* 9:108–121.