7 Corruption and the shadow of the future

A generalization of an ABM with repeated interactions

Nick van Doormaal, Stijn Ruiter, and Andrew M. Lemieux

Introduction

Corruption is a widespread, complex phenomenon with detrimental impacts on social and economic development around the world (Banuri & Eckel, 2012; Transparency International, 2019; Warf, 2016). High levels of corruption lead to unfair distributions of resources and income, while undermining democracy and the rule of law. Corruption manifests itself in a variety of ways, including individual acts such as accepting bribes, and grand larceny on an organized, institutional scale (Transparency International, 2019; Warf, 2016). Ultimately, a better understanding of the effects and causes of corruption has the potential to inform corruption control and anti-corruption campaigns. Policies that seek to deter corruption can be directed at influencing individual behaviour and perception through penalties, public relations campaigns or organisational structures and procedures. Such policies have the potential to increase the perceived costs for engaging in corrupt activities and deter potential offenders.

Deterrence of corrupt behaviour

Early modern deterrence theorists argued that people weigh costs and benefits in their decision-making process (Becker, 1968). Assuming rational decision-making, individuals can be deterred from crime if the costs outweigh the benefits. This model of deterrence is known as the “economic model of rational deterrence” (Becker, 1968). The theory does not differentiate between criminals and non-criminals, and states that every individual has their own cost–benefit assessment for committing crimes (Paternoster, 2010).

Deterrence theory divides the costs of committing a crime into three components related to punishment: severity, certainty and celerity (Nagin, 2013). The severity of punishment needs to be strong enough to sufficiently reduce the benefits of a particular crime. Certainty applies to the likelihood of receiving punishment, whereas celerity applies to the timing of imposing punishment. Empirical studies on the deterrence of crime show mixed results regarding the severity and celerity of punishment. The certainty of punishment, however, is often found to play an important role in crime deterrence (Nagin, 2013).

Deterrence is general or specific. We refer to specific deterrence when an individual is deterred from committing future crimes through the experience of punishment. General deterrence refers to the idea that individuals respond to the threat of punishment; the punishment of those who commit crimes will also serve as an example to potential offenders among the general population.

Most forms of corruption are interactions between at least two individuals or groups. Choosing to behave corruptly will yield the highest rewards that the individual otherwise would not be able to get. Although the rewards of corruption are generally higher than for following the rules, the rewards could be counterbalanced by high costs. According to deterrence theory, the costs of corruption could be increased by increasing the perceived severity, certainty or celerity of punishment. Empirical studies commonly find the costs of crime can be increased by increasing the likelihood of punishment (Nagin, 2013).

Modelling corruption and deterrence

The level of corruption in a society is the aggregate-level outcome of all individual decisions. The associated costs of corruption are perceived differently among individuals and it is the perception of the severity, certainty and celerity of punishment that matters (Paternoster, 2010). The conditions to deter someone from corruption can vary depending on the individual's perception of punishment. The reciprocal relationships among these three components and its effect on deterrence complicate our understanding of what leads to higher or lower levels of corruption. To tackle the complexity of corruption and deterrence, agent-based models can be used to assess these conditions.

Agent-based modelling is a simulation technique in which the behaviour and decision-making of autonomous individuals are modelled to identify relevant factors for the entire system (Epstein and Axtell, 1996). It allows researchers to study complex systems and problems in an abstract environment, without them being influenced by specific characteristics of certain locations, and without the need for direct observations in the real world. Corruption is a widespread, complex phenomenon and difficult to observe directly in the real world. Therefore, agent-based modelling can be used to study the dynamics of corruption and its deterrence in an abstract environment. For further explanation of agent-based models and their application in criminological research we refer to the introductory chapter of this book.

Studies on corruption have used agent-based models or similar bottom-up approaches to explore the mechanisms, drivers (Farjam et al., 2015; Ye et al., 2011; Zausinová et al., 2019) and emergence of corruption (Kim et al., 2013; Situngkir & Khanafiah, 2006; Voinea, 2013). Hammond (2000) modelled corruption as a game-theoretic interaction between two agent populations and showed the effects of general deterrence on corruption levels in an artificial society. His model showed that the effect of general deterrence or "fear of enforcement" can spread rapidly throughout society and can lead to a transition from a high corrupt society to a low corrupt one. The simulation results showed that the transition from a corrupt society to an honest one can happen endogenously. Hammond's (2000) findings contradict existing political economy and economics literature, which assumes these transitions are the result of an exogenous force like a new election, government policy or economic shock (Di Vita, 2007; Goel & Nelson, 2005).

Improving corruption research

Hammond's model offered an alternative explanation for how general deterrence can cause transitions in a corrupt society. However, his model was based on a specific instance of corruption. The interactions between agents were modelled as one-shot, random encounters. Certain scenarios can indeed be regarded as a one-time random interaction, for example citizens declaring taxes or applying for a driver's licence. However, not all interactions involving corruption can be represented with one-shot interactions. Some types of interaction involving corruption are improperly captured with one-shot interactions or even a disconnected series of one-time interactions among individuals. For example, a supplier and a purchasing agent often interact with each other multiple times and over an extended time, as do corrupt police officers who facilitate the work of local criminal groups. To better understand those forms of corruption, assuming repeated interactions between the same individuals would better reflect reality than one-shot interactions.

A key element that is omitted in one-shot interactions is some form of reciprocity (Dal Bo & Frechette, 2011; Gossner & Tomala, 2009). Corrupt individuals working together operate in a risky environment. Interaction is based on the expectation that cooperation will be beneficial to both sides (Dal Bo & Frechette, 2011; Gossner & Tomala, 2009). If one individual does not hold up his or her end of the bargain, then it is likely that the other side will not cooperate in the future. Repeated interaction favours a solid base for mutual trust, and this will also have an impact on the certainty of apprehension. An individual's cooperation in social situations depends strongly on the degree to which others cooperate, which is why offenders are more likely to co-offend with family, long-time friends and other confidants (Kleemans & de Poot, 2008). In terms of deterrence theory, repeated interactions and mutual trust can reduce the certainty of punishment as long as both parties hold up their end of the bargain. Previous studies of human decision-making in exchange relationships have noted that the possibility of there being future interactions can hold important consequences for understanding cooperation (Axelrod, 1984; Balliet et al., 2011).

Objective

In this chapter we focus on the theory of general deterrence and the certainty of punishment. The work of Hammond (2000) showed the effect of general deterrence on the spread of corruption in an artificial society, but was based on a specific instance of corruption with one-shot interactions. Repeated interactions, on the other hand, likely reduce the certainty of punishment and hence should lead to higher levels of corruption. Our stylized fact is therefore that certainty of punishment is a necessary component for the general deterrence of crime. We extend the original model by Hammond (2000) to examine if and how repeated interactions change the ability of high corrupt societies to transition into low corrupt societies. We explore several scenarios in which we systematically vary the likelihood of punishment to better understand the conditions that lead to higher or lower levels of corruption.

The following sections provide an overview of the modelling framework, and describe how we generalized Hammond's model to include repeated interactions between agents.

Model framework

Payoff structures

We first explain the model from a purely game-theoretic perspective. Every player can choose one of two strategies: "corrupt" or "honest". Their decision is based on the expected payoff of that strategy (Table 7.1). We assume that the corrupt action yields the highest payoff (x), but only if both players choose "corrupt". A game-theoretic analysis shows that choosing "corrupt" results in a strict Nash equilibrium because no one has an incentive to change their decision; the outcome with both players choosing "corrupt" is better than all other outcomes. If one or both players choose "honest", then they both receive the lowest payoff (y). Choosing "honest" leads to a weak Nash equilibrium. Neither player can do better by choosing "corrupt" if the other player chooses "honest" because both options yield the low payoff y.

Table 7.1 The 2 × 2 corruption game payoff structure: x and y are the payoffs for choosing that particular strategy

		Player 2 **Corrupt**	**Honest**
Player 1	**Corrupt**	x, x	y, y
	Honest	y, y	y, y
$x > y$			

Agent-based modelling approach

Although the decisions in the game are straightforward, it is still an over-simplification of decision-making. No two persons in reality are the same, and each individual has a unique set of values that affects their perception. Furthermore, people's decisions are not only influenced by their own set of values, but also by the behaviour of others, especially friends and family. Therefore, individuals will perceive the payoffs for engaging in corruption differently and make different decisions. The game can be played by a heterogeneous population to better reflect reality, but would also be difficult to solve under the traditional game-theoretic framework. The behaviour of heterogeneous players in a dynamic environment is difficult to predict with game theory alone, but can be captured and quantified with agent-based models. Agent-based modelling allows for heterogeneous and autonomous agents capable of exhibiting human-like behaviours, for example, corrupt behaviours.

Agent characteristics and behaviour

Morality and perceived payoff

Based on the offender motivation literature, we assume that agents possess some intrinsic core values on how to behave (McMurran & Ward, 2004; Ryan & Deci, 2000). We refer to this as an agent's "morality" (Hammond (2000) labelled it as "Honesty".) It influences how the agent perceives the payoffs for choosing "corrupt". An agent that scores high on morality gains little from a corrupt interaction, while only an agent with the lowest morality score gains the full benefits. Increasing levels of morality thus decrease the perceived payoff of corruption. Morality takes a random value between 0 and 1 and is assigned to every agent before a simulation run starts. The assigned morality value is fixed throughout the simulation run. The perceived payoff for acting corrupt (x_i) is calculated as:

$$x_i = (1 - morality)\,x$$

Networks

Every agent has its own network of other agents (Hammond (2000) referred to the agents within a network as "friends".) The size of an agent's network is fixed and set by the modeller at the start of the simulation. Every agent creates an undirected link with a number of other random agents until it reaches the specified network size. Agents can be part of multiple networks but will never exceed the specified network size. The size of the network is fixed and set by the modeller at the start of the simulation. We are uncertain if our network setup is the same as the one described by Hammond because little information was provided. Hammond (2000) described it as follows: "These networks are of fixed (standard) size, but the specific contents of each agent's network is randomly assigned during initialization".

The agent has access to certain information of other agents within its network. The agent can observe the most recent actions of the members within its network ("honest" or "corrupt"), and observe which members are suspended.

Agent decision-making

Each agent first calculates the perceived payoffs for acting corrupt (x_i). Next, every agent estimates the probability of encountering a corrupt agent as follows.

The agent keeps track of the actions chosen by the agents it interacted with in previous rounds and will remember those actions for a certain period of time. This is referred to as the agent's "memory". The size of memory (i.e. the number of past interactions the agent can remember) is set by the modeller at the start of the simulation and is fixed throughout the simulation run. The agent examines its memory to count the number of corrupt partners it has encountered in previous interactions. The agent calculates the probability of encountering a corrupt agent as: $A = \frac{n}{N}$, in which n is the number of corrupt partners encountered in N previous interactions.

Every agent also estimates the probability of apprehension for acting corrupt in this round. The agent does this by examining the behaviour and status of other agents in its social network. Every agent can only observe the most recent action chosen by all network members, and observe which network members are suspended. The probability of apprehension is calculated based on the number of suspended network members and the number of corrupt network members in the last round. The probability of apprehension for a corrupt action in a round is calculated as: $B = \frac{m}{M}$, in which m is the number of suspended network members and M the number of corrupt network members in the previous round.

Finally, agents know the length of suspension k. Suspended agents are removed from play for the duration of k. The decision rule for each agent to act corrupt is then:

$$(1-\boldsymbol{B})\left[\boldsymbol{A}\boldsymbol{x}_i + (1-\boldsymbol{A})\boldsymbol{y}\right] + \boldsymbol{B}\left[\boldsymbol{y} - \boldsymbol{k}\boldsymbol{y}\right] > \boldsymbol{y}$$

The cost of being suspended is ky. The partner of the suspended agent will randomly choose another agent as its new partner. Suspended agents and agents that already have a partner cannot be chosen. It is possible that no agents will be available, because some are suspended and others already have a partner. If so, then the agent without a partner will not interact and wait until agents become available again.

Compare actions with partner

Every agent randomly chooses a partner to interact with. Only agents who are not suspended and do not have a partner at that moment can be selected. If no other agent is available as a partner, then the agent will remain without a partner until other agents become available. Depending on the model settings, the interaction between the agent and its partner may only last one round (one-shot interaction) or multiple interactions over a longer period of time (repeated interactions). Each agent decides between the two actions ("corrupt" and "honest") immediately before each interaction by using the decision rule outlined before.

Next, the agents compare their actions, which leads to one of three possible outcomes (Figure 7.1):

1. Honest interaction: both agents act honest and will receive the lowest payoff.
2. Mismatch: a corrupt agent meets an honest agent. Both agents will receive the lowest payoff and the honest agent will "report" the corrupt agent. If the number of reports reaches a certain threshold, then the corrupt agent will be temporarily suspended.
3. Collusive corruption: both agents act corrupt and will receive the highest payoff.

Enforcement

Just as in Hammond's model, a punishment component will be triggered if a corrupt agent meets an honest agent. The model keeps track of how many reports every agent has received for acting corrupt throughout the simulation run. If an agent has been reported a certain number of times, then the reported agent will be suspended temporarily. Suspended agents cannot interact with other players and therefore cannot gain payoffs. The agent is allowed to interact with other agents again and able to gain payoffs after

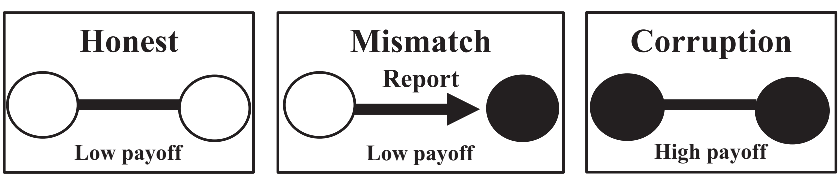

Figure 7.1 Diagram visualizing the three potential outcomes of two agents interacting.

serving the suspension time. An agent's decision-making is unaltered after being suspended.

The agents in the model are aware that it is possible to be suspended for acting corrupt and do know the length of the suspension term. The length of suspension is set during the model initialization and can be specified by the user. Agents take the length of suspension k into account when deciding between acting corrupt or honest. However, the agents themselves do not know how many reports they have received or how many reports are required to get suspended. The next section will describe the decision-making process of the agents in the model.

A model round can be summarized as follows:

1. *Select agent*: every round, an agent will be randomly paired with another agent.
2. *Select strategy*: each agent decides simultaneously to act corruptly or honestly. The decision rule is based on the agent's bounded rationality.
3. *Receive payoff*: acting corrupt yields the highest payoff (x), but only if the other agent also chooses the corrupt strategy. If both agents choose to act honest, they both receive the lowest payoff (y). If only one of two agents acts corrupt, the honest agent reports the corrupt agent and both receive the lowest payoff (y).
4. *Suspend agents*: if an agent is reported a predefined number of times, the agent will be suspended for a period of time. A suspended agent cannot interact with other agents or gain payoffs.
5. *Release agents*: after serving the suspension time, agents are allowed to interact with other agents again.

From random one-shot to repeated interactions

In Hammond's original model, an agent was randomly paired with another available agent. Hammond did not explore if the transition from a high to low corrupt state can be reached when the same pair of agents interact for consecutive rounds. Our model extension is aimed at changing the current one-shot interaction to repeated interactions between the same agents. The decision-making of agents and model processes described earlier still works in the same way. The only difference is that agents will interact with the same agent over multiple rounds. The number of interactions can be specified by the user and applies to all agents. The number of repeated interactions will never be longer than specified, but can end earlier if one of the agents is suspended.

Agent interactions and memory weights

We expect that the length of memory will play an important role in scenarios with repeated interactions. The agent's memory in the original model reflects the different choices by the other players. The agent uses its memory to assess the likelihood of encountering a corrupt agent in the next round. However, when pairs of agents interact with each other over multiple consecutive rounds, their memory will mostly consist of interactions with that particular agent. The agent's challenge is then to assess the likelihood that its current partner will choose the corrupt action in the next round, rather than assessing the likelihood of encountering a corrupt agent in the entire population. We incorporate this by assigning a weight to the most recent interaction of the agent's memory. The higher the weight, the larger the influence of the most recent interaction on the agent's estimation of encountering a corrupt agent. When the weight is 100%, the agent will only include the most recent interaction in the decision-making for estimating the likelihood of encountering a corrupt agent and disregard all other, older memories. Our extension is built in such a way that when no additional weight is assigned to the most recent interaction, the model reflects the original model by Hammond.

Scenario simulations

Scenarios with repeated interactions were compared with scenarios with Hammond's original one-shot model settings. We run scenarios with agents repeatedly interacting with each other for two to eight consecutive rounds. We run these scenarios in combination with situations in which no addition weight, 50% and 100% weight were assigned to the most recent interaction of the agent's memory (Table 7.2). All other parameters in the model were kept constant throughout all simulation runs (Table 7.3). Each simulation run lasts no longer than 2.000 time units (called ticks in the model). The model keeps track of the number of honest and corrupt agents throughout the simulation run. The outcome variable was the number of corrupt agents at the end of each simulation run. We also recorded when a transition from high corruption to low corruption took place. This was recorded as the moment when the number of honest agents exceeded the number of corrupt agents. Each combination of settings was run 150 times.

Table 7.2 Overview of the parameters introduced to the original model and the values in the simulation runs

Parameter introduced to the original model	*Description*	*Values*
Interactions	The number of times an agent consecutively interacts with the same agent before being matched with a different agent	1–8
Weight	Percentage of weight assigned to the most recent interaction in the agent's memory. The higher the weight, the larger its influence on the agent's estimation of encountering a corrupt agent	50, 100

Table 7.3 Overview of the corruption model's defaults setting, based on Hammond (2000)

Parameter	*Description*	*Default value*
Morality	Agent variable to reflect an inherent propensity for "doing good". Increasing levels of morality decrease the perceived payoff of corruption	Randomly distributed [0,1]
Corruption payoff	Benefit an agent receives for a successful collusive corrupt action. Corruption payoff is always the highest payoff	20
Honest payoff	Benefit that an agent receives for choosing honest actions. Both players receive the honest payoff in a mismatch	1
Reports	Number of reports required to get suspended	2 reports
Suspension term	Number of rounds an agent will be suspended from interaction with other players	4 rounds
Memory	Every agent remembers a certain number of actions chosen by the other players it interacted with in previous rounds	5 rounds
Network	Represents the number of "friends" that every agent has. Agents are randomly assigned to a network	10 agents
Population	Total number of agents in every simulation round	300 agents

We used the software NetLogo version 6.1.0 (Wilensky, 1999), building upon an earlier implementation of Hammond's original model (Lonsdale, 2017). Our model and code are published online at: http://modelingcommons.org/browse/one_model/6210.

The Shapiro–Wilk test was used to test for normality. Further analyses were performed with the Kruskal–Wallis test. A post-hoc comparison using Dunn's test with the Bonferroni adjustment was performed if the Kruskal–Wallis showed statistical differences between the groups. These conservative non-parametric methods were applied to reduce the possibility of a type I error.

Results

One-shot vs. repeated interactions

The effect of repeated interactions was compared with the one-shot interactions based on the original model. All other model parameters were kept constant throughout the simulation runs (Table 7.3). The number of corrupt agents at the end of a simulation was statistically significantly different between the different number of repeated interactions (Kruskal–Wallis test; $H = 550.96$; d.f. $= 7$; $p < 0.001$). The number of corrupt agents was the lowest for one-shot interactions, and the two and three repeated interactions (Figure 7.2). The scenario with three repeated interactions

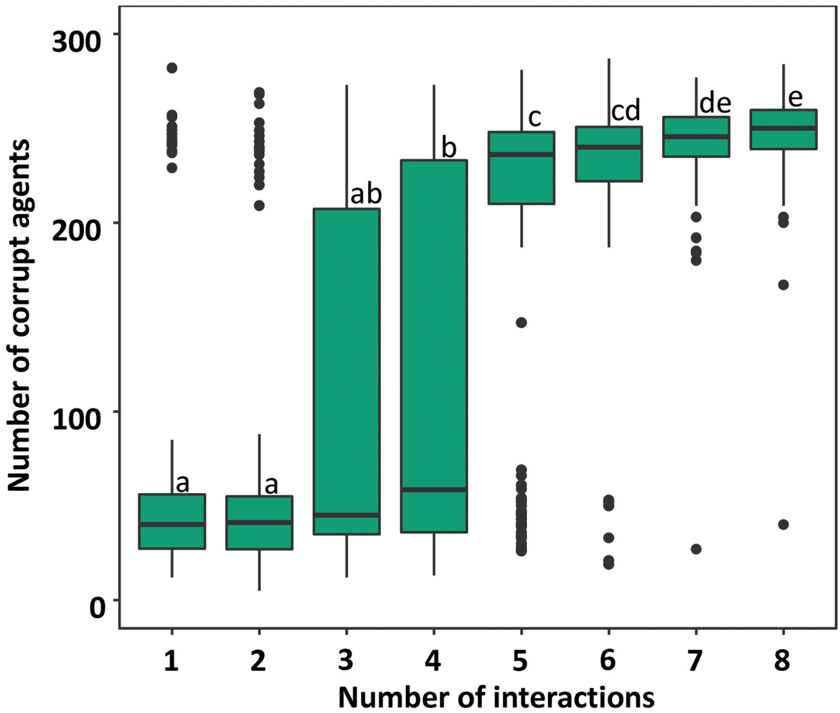

Figure 7.2 Simulation results for the effect of repeated interactions on the number of corrupt agents at the end of each simulation run. No weights to the agent's memory were assigned. For statistical comparisons, a Kruskal–Wallis test was conducted followed by a Dunn's post-hoc test. The plots bearing the same letters are not statistically significantly different at the 5% level.

showed more variation in the number of corrupt agents but it was not statistically significantly different from single- and double-shot interactions. The outliers in the one-shot and two-shot repeated interactions showed that the transition from high to low levels of corruption does not always happen; 11% (n = 16) of the simulation runs with one-shot interactions and 13% (n = 20) of the simulations with two repeated interactions did not result in a transition to low levels of corruption.

The scenarios with three and four repeated interactions showed more variation in the number of corrupt agents than the other scenarios (Figure 7.2). The majority of these runs still resulted in a transition from high to low levels of corruption, as indicated by the low median of corrupt agents (45 and 58 corrupt agents for the three repeated and four repeated scenarios respectively). For the simulation runs with three repeated interactions, 26% (n = 39) did not result in a transition to low levels of corruption. For four repeated interaction runs, this was 47% (n = 71).

The number of corrupt agents at the end of every simulation run was higher when agents repeatedly interacted for five rounds or more compared to the other scenarios (Figure 7.2). In 78% (n = 117) of the runs with five repeated interactions, the levels of corruption remained high and did not show a transition to low levels. Simulation runs with seven and eight repeated interactions showed the highest level of corrupt agents. A transition to low levels of corruption was observed on one occasion only for scenarios with seven and eight repeated interactions.

Weight to memory

Weights were assigned to the agents' most recent interaction to explore their effect on the transition from a high corrupt society to a low one. The higher the weight, the larger the influence of the most recent interaction on the agent's estimation of encountering a corrupt agent. The results for adding weights to the agent's memory resulted in statistically significant, but small, differences in the number of corrupt agents at the end of each simulation run (Kruskal–Wallis test; $H = 1088.5$; d.f. = 15; $p < 0.001$). When the weight parameter was set to 50, the results were similar to scenarios with zero weight; the number of corrupt agents was the lowest for the one-shot interactions (Figure 7.3). The three repeated interactions scenario showed more variation in the number of corrupt agents as well. The majority of those runs (55%, $n = 83$) showed a transition from high to low levels of corruption. The number of corrupt agents at the end of the simulation runs was highest for scenarios with five or more repeated interactions. Only one transition from a high level of corruption to a low level was observed for the six repeated interactions scenario, while no transition was observed for seven or eight repeated interaction scenarios.

When the weight was set to 100%, the agent will only include its most recent interaction in the decision-making for estimating the likelihood of encountering a corrupt agent and disregard all other, older memories. For these scenarios, the results look similar to the scenarios with no addition weights and 50% weights; the number of corrupt agents at the end of the simulation runs was highest for scenarios with six or more repeated interactions (Figure 7.3). Scenarios with the weight value set to 100% showed a high variability in the number of corrupt agents for the two repeated interactions. Just over half of all the simulation runs resulted in a transition from high to low corruption levels (53%, n = 79). Three times a transition from high levels of corruption to low levels occurred for the six repeated interactions scenario. No transition was observed for the seven or eight repeated interaction scenarios.

Time until transition

We also explored the time until transition for the different number of repeated interactions. When no weight was assigned to the agent's memory, transitions

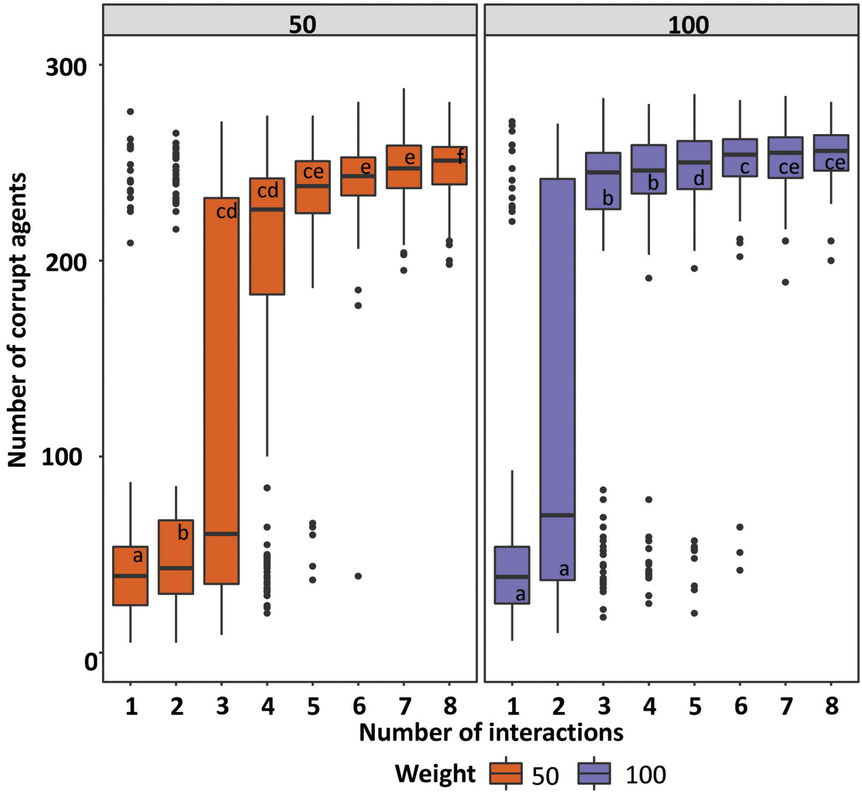

Figure 7.3 Simulation results for the effect of repeated interactions together with assigning weights to agent's memory on the number of corrupt agents at the end of each simulation run. For statistical comparisons, a Kruskal–Wallis test was conducted, followed by a Dunn's post-hoc test. The plots bearing the same letters are not statistically different at the 5% level.

from high to low levels of corruption tend to happen earlier for scenarios with one-shot interactions, followed by scenarios with two, three and four repeated interactions (Figure 7.4). The median "survival" time for a high corrupt state was 32 ticks for scenarios with one-shot interactions. For the repeated interactions, this was 71 ticks, 457 ticks and 1459 ticks for the two, three and four repeated interactions respectively (Table 7.4). For the scenarios with five or more repeated interactions not enough transitions from high to low corruption were observed to estimate a median survival time.

In scenarios with weights assigned to the agent's memory, the scenarios with two and three repeated interactions were affected most (Figure 7.5A). The median "survival" time for a high corrupt state in the two repeated interaction scenarios and a weight of 50% was 120 ticks. When the weight parameter was

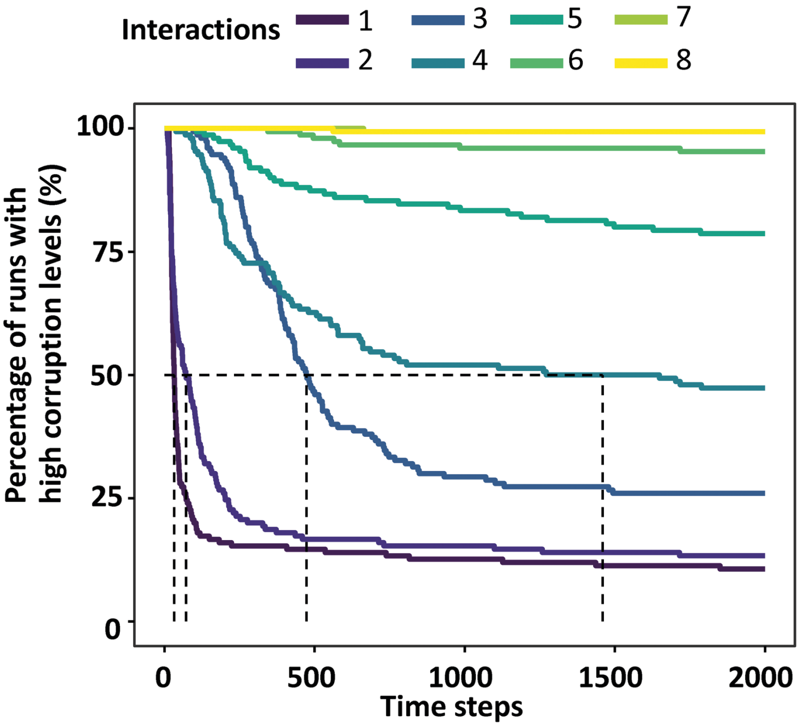

Figure 7.4 Survival curves for duration until transition from high levels to low levels of corruption for simulation runs with one and multiple repeated interactions. Each coloured line represents a different number of interactions. No weights were assigned to agent's memory. The dashed lines show the median survival time for that scenario.

Table 7.4 Summary of the median time until transition for the different scenarios.

Number of interactions	*Median survival time (ticks)*		
	No weight	*50 weight*	*100 weight*
1	32	33	22
2	71	120	762
3	457	1100	–
4	1459	–	–
5–8	–	–	–

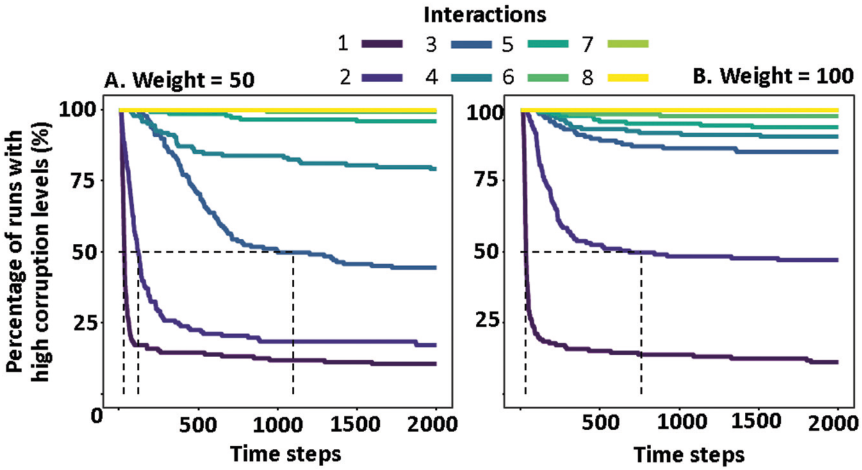

Figure 7.5 Survival curves for duration until transition from high levels to low levels of corruption for simulation runs with one and multiple repeated interactions. Each coloured line represents a different number of interactions. The left graph (A) shows scenarios with the weight parameter set to 50%; the right graph (B) has weight set to 100%. The dashed lines show the median survival time for that scenario.

set to 100%, the number of ticks increased to 762 (Table 7.4). For the three repeated interactions scenario, this was 1100 ticks for the 50% weights. When the weight parameter was set to 100%, not enough transitions occurred to estimate the median "survival" time (Figure 7.5B). The two weight values did not influence the one-shot interactions because both median survival times (respectively 33 and 22 ticks) were similar to one without the weight parameter (32 ticks).

Model sensitivity to parameter settings

To get a better understanding of how our model behaves under different parameter settings, we ran scenarios with varying parameters for corrupt payoff, honest payoff, suspension term and memory size. These results are summarized in Table 7.5. Our results for the corruption and honest payoffs and suspension terms in the one-shot scenarios were in line with the findings from Hammond (2000).

Discussion

This chapter introduces an extension of an agent-based model for corruption originally developed by Hammond (2000). The original model represents a

Table 7.5 Summary of the scenario runs with varying values of corruption payoff, honest payoff and suspension term

Parameter	*One-shot interaction*	*Repeated interactions*
Corruption payoff	Transition to low corruption less likely with increasing payoffs, but still occurs regularly. The transitions that do happen take longer	Higher corruption payoffs decrease the likelihood of a transition to low corruption even further. From six or more repeated interactions, also low payoffs do not show a transition most of the time
Honest payoff	Increased payoffs for honest behaviour always resulted in a transition to low corruption and transitions happen almost immediately	Increased payoffs for honest behaviour always resulted in a transition to low corruption, but transitions do take longer to occur with increasing numbers of repeated interactions
Suspension term	For suspension term = 3, a transition rarely occurred. When this parameter was set to 4 and 5, a transition occurred in most runs	For suspension term = 3, a transition never occurred. When set to 4, transitions only occurred in scenarios with four or fewer repeated interactions. Transitions often occurred when suspension term was set to 5 in all repeated interactions
Memory	Transition to low corruption was observed for the majority of runs for both small and large sizes of memory	Dynamics for both small and large sizes of memory start to shift from around 3–4 repeated interactions. From 5 repeated interactions onwards, a transition to low levels of corruption was not observed for the majority of runs of both small and large sizes of memory

specific instance of corruption, that of one-shot interactions. By also studying repeated interactions between agents, we could evaluate our stylized fact. Our stylized fact was based on the empirical regularity that certainty is an important component for the general deterrence of crime. Our expanded model shows that high levels of corruption were more likely to persist in scenarios with repeated interactions compared to one-shot scenarios. Transitions from high to low levels of corruption rarely occurred in scenarios with increasing number of repeated interactions. The time it took for a transition to occur also increased with increasing number of repeated interactions. We were able to reproduce our stylized fact only for the one-shot interactions and therefore found support that certainty is indeed an important component of general deterrence.

While we were able to reproduce our stylized fact, we were not able to exactly replicate Hammond's original findings. According to Hammond, all agents would eventually decide to behave honestly, until all agents within the population are honest. Our results from the one-shot interactions are similar to those described by Hammond (2000), but none of our simulation runs eventually led to a population with only honest agents. We found that some corrupt agents persisted even while the society was in a low corrupt state. Furthermore, 11% of the simulation runs did not result in a transition. It is unclear what caused these differences in results, because the original code of Hammond's model was unfortunately not published or made available elsewhere. His model was described in great detail, but still information on certain elements were missing from his description. We took NetLogo implementations of Hammond's model published by others as our starting point and were forced to make our own assumptions, for example on the exact network configuration and reporting system.

Deterrence and repeated interactions

In our model, the number of corrupt agents remained high with repeated interactions and often a transition to low levels of corruption did not happen. We interpret this result as the certainty of punishment diminishing with increasing numbers of repeated interactions. When an agent has "learned" that its partner is willing to act corruptly, the optimal decision is to act corruptly as well. Learning that your partner is corrupt removed the certainty of punishment. From that moment onwards, agents have no incentive to report their partner and no one will get suspended. The results of this are then signalled throughout society via the social networks of the agents. Agents use their own network to get a subjective estimate of how likely it is to be caught for a corrupt act in the future. If no one inside the network is suspended, then a corrupt agent has no incentive to change its strategy. The agent feels certain of encountering another corrupt agent, and believes he will not get caught for acting corruptly. Similarly, honest agents are tempted to choose the "corrupt" strategy because none of the corrupt agents inside the network are receiving punishment. Hence, repeated interactions reduce the certainty of punishment and with it the effect of general deterrence dissolves.

This chapter focused on the theory of general deterrence. The theory states that the interplay of severity, certainty and celerity of punishment can deter crime. In our model, we only varied the certainty of punishment, while the severity and celerity were held constant. Deterrence theorists argue that at least certainty of punishment is important for the deterrence of crime (Nagin, 2013; Paternoster, 2010). Empirical studies show that the effects of severity and celerity on crime are still not well understood (Nagin, 2013; Paternoster, 2010; Friesen, 2012). Therefore, we did not fully explore the interactive effects between severity and certainty. Our simulation runs do suggest that some level of severity may be required for certainty of punishment to be effective. A similar suggestion has been proposed by other researchers (Engel & Nagin, 2015; Stafford et al., 1986). The results in this chapter suggest that at least certainty plays an important role in the deterrence of corruption. Our model can be used to explore these components further to detect if a particular severity threshold is indeed needed to effectively deter corruption.

One-shots to increase certainty

Our results are in line with the current literature on behaviour in repeated interactions; cooperation becomes more likely if there is "a shadow of the future" (Axelrod, 1984; Sabater & Sierra, 2005). Cooperation, in our model, refers to two agents successfully colluding in a corrupt act. In the example of corrupt officers working together with local criminal groups, the officers routinely work together with offenders for longer periods of time. By working together, they can get a higher payoff and reduce the certainty of getting caught. To increase the perceptions of certainty of punishment, one should aim to create a setting that mimics a one-shot scenario or fewer repeated interactions. These scenarios should represent situations in which individuals do not learn about the behaviour of their partner. An example to approach this is through the introduction of rotation schemes. Abbink (2004) showed that the number of bribery attempts and their volume are cut by approximately half when an agent is paired with another random agent in every round. Staff rotation schemes have the potential to reduce levels of collusive corruption in organizational or institutional settings.

Model assumptions and improvements

The general approach in extending the model leads to two important assumptions related to the enforcement and reporting system. The strength of this general model is that these scenarios can be easily incorporated in the current model.

In our model, the model has perfect information on all agents. Although the individual agents do not keep track of the number of reports they received, the model will automatically remove a corrupt agent from the game when that agent received a certain number of reports. This assumption is related to the certainty of apprehension in deterrence theory. Our results show that, even when one has perfect information, high levels of corruption can still persist if agents repeatedly interact with one another. Only through mismatches in which an honest agent reports a corrupt agent will the system know who is corrupt. Our model can be extended further to compare with different enforcement systems – for example, a new enforcement system in which agents are randomly inspected for corrupt behaviour or by focusing only on agents who were apprehended in the past.

For simplicity, we assumed that an honest agent will always report a corrupt agent. An agent may decide not to report a corrupt agent under certain circumstances. For example, an honest agent might not blow the whistle if the reporting agent was involved in corrupt practices in the past. A possible extension would be to relate the likelihood of reporting a corrupt agent to an agent's morality, or to the number of corrupt acts in the past. These potential model extensions can provide more insights into how whistleblowing could affect levels of corruption over time.

Conclusion

This chapter improved upon the work of Hammond (2000) to test the theory of general deterrence and the role of certainty of punishment on controlling corruption in an artificial environment. Our stylized fact, that certainty of punishment is a necessary component for the general deterrence of crime, was only replicated for one-shot interactions. Repeat interactions between agents reduce the certainty of punishment and corruption is therefore more likely to persist inside the artificial society. This led us to suggest that the certainty of punishment is indeed an important component of general deterrence theory. The general framework of the model can be easily expanded to explore different elements and conditions for deterring corruption. The model presented here should be regarded as a theoretical exploration to better understand the complexity of corruption and deterrence.

References

Abbink, K. (2004). Staff rotation as an anti-corruption policy: an experimental study. *European Journal of Political Economy, 20*(4), 887–906. doi:10.1016/j.ejpoleco.2003.10.008

Axelrod, R. (1984). On the evolution of cooperation. New York: Basic Books.

Balliet, D., Mulder, L. B., & Van Lange, P. A. (2011). Reward, punishment, and cooperation: a meta-analysis. *Psychological Bulletin*, *137*(4), 594. doi:10.1037/a0023489.

Banuri, S., & Eckel, C. (2012). Experiments in culture and corruption: a review. In D. Serra & L. Wantchekon (eds.), New advances in experimental research on corruption (Vol. 15, pp. 51–76). Bingley: Emerald.

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

Dal Bo, P., & Frechette, G. R. (2011). The evolution of cooperation in infinitely repeated games: experimental evidence. *American Economic Review, 101*(1), 411–429. doi:10.1257/aer.101.1.411

Di Vita, G. (2007). A note on exogenous changes in incentives for and deterrence of corruption. *European Journal of Law and Economics, 24*(1), 15–27.

Engel, C., & Nagin, D. (2015). Who is afraid of the stick? Experimentally testing the deterrent effect of sanction certainty. *Review of Behavioral Economics, 2*(4), 405–434. doi:10.1561/105.00000037

Epstein, J., & Axtell, R. (1996). Growing artificial societies: social science from the bottom up. Cambridge: Brookings Institution Press & MIT Press.

Farjam, M., Faillo, M., Sprinkhuizen-Kuyper, I., & Haselager, P. (2015). Punishment mechanisms and their effect on cooperation: a simulation study. *Journal of Artificial Societies and Social Simulation*, *18*(1). doi:10.18564/jasss.2647

Friesen, L. (2012). Certainty of punishment versus severity of punishment: an experimental investigation. *Southern Economic Journal*, *79*(2), 399–421. doi:10.4284/0038-4038-2011.152

Goel, R. K., & Nelson, M. A. (2005). Economic freedom versus political freedom: cross-country influences on corruption. *Australian Economic Papers, 44*(2), 121–133.

Gossner, O., & Tomala, T. (2009). Repeated games with complete information. In R. Meyers (eds.),Encyclopedia of complexity and systems science, pp. 7616–7630. Berlin: Springer.

Hammond, R. (2000). Endogenous transition dynamics in corruption: an agent-based computer model. Washington, DC: Center on Social and Economic Dynamics.

Kim, Y., Zhong, W., & Chun, Y. (2013). Modeling sanction choices on fraudulent benefit exchanges in public service delivery. *Journal of Artificial Societies and Social Simulation, 16*(2), doi:10.18564/jasss.2175

Kleemans, E. R., & de Poot, C. J. (2008). Criminal careers in organized crime and social opportunity structure. *European Journal of Criminology, 5*(1), 69–98. doi:10.1177/1477370807084225

Lonsdale, C. (2017). *Creating an agent-based model of Hammond's model of social inequality using netlogo*. http://charleslonsdale.co.uk/portfolio/advanced-2.php. Accessed 9 January 2019.

McMurran, M., & Ward, T. (2004). Motivating offenders to change in therapy: an organizing framework. *Legal and Criminological Psychology, 9*, 295–311. doi:10.1348/1355325041719365

Nagin, D. S. (2013). Deterrence in the twenty-first century. *Crime and Justice in America, 1975–2025*, *42*, 199–263. doi:10.1086/670398

Paternoster, R. (2010). How much do we really know about criminal deterrence? *Journal of Criminal Law & Criminology, 100*(3), 765–823.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68–78. doi:10.1037//0003-066x.55.1.68

Sabater, J., & Sierra, C. (2005). Review on computational trust and reputation models. *Artificial Intelligence Review, 24*, 33–60. doi:0.1007/s10462-004-0041-5

Situngkir, H., & Khanafiah, D. (2006). Theorizing corruption through agent-based modeling. Paper presented at the 9th Joint International Conference on Information Sciences (JCIS-06).

Stafford, M. C., Gray, L. N., Menke, B. A., & Ward, D. A. (1986). Modeling the deterrent effects of punishment. *Social Psychology Quarterly, 49*(4), 338–347. doi:10.2307/2786773

Transparency International. (2019). *Corruption perception index 2018*. www.transparency.org/cpi2018. Accessed 30 January 2019.

Voinea, C. (2013). Bribery-scape: an artificial society-based simulation model of corruption's emergence and growth. *European Quarterly of Political Attitudes and Mentalities-EQPAM, 2*(1), 27–54.

Warf, B. (2016). Global geographies of corruption. *Geojournal, 81*(5), 657–669. doi:10.1007/s10708-015-9656-0

Wilensky, U. (1999). Netlogo . Evanston, IL: Center for Connected Learning and Computer-Based Modeling, Northwestern University. Retrieved from http://ccl.northwestern.edu/netlogo/

Ye, H., Tan, F., Ding, M., Jia, Y., & Chen, Y. (2011). Sympathy and punishment: evolution of cooperation in public goods game. *Journal of Artificial Societies and Social Simulation*, *14*(4). Article no. 20. doi:10.18564/jasss.1805

Zausinová, J., Zoričak, M., Vološin, M., & Gazda, V. (2019). Aspects of complexity in citizen–bureaucrat corruption: an agent-based simulation model. *Journal of Economic Interaction and Coordination*, *15*, 527–552. doi:10.1007/s11403-019-00240-x