Games and Economic Behavior 124 (2020) 319–353

Contents lists available at ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

Partners in crime? Corruption as a criminal network

Romain Ferrali [1]

Social Science Division, New York University Abu Dhabi, United Arab Emirates

ARTICLE INFO

Article history:
Received 17 January 2019
Available online 7 September 2020

JEL classification:
C92
C93
D73
D85

Keywords:
Corruption
Networks
Lab experiment
Diffusion

ABSTRACT

How does the structure of an organization affect corruption? This paper analyzes a model that views organizations as networks on which coalitions of corrupt accomplices may form. This network approach to corruption provides new insights into the problem: (i) corruption will arise in enclaves, i.e. coalitions that minimize joint exposure to witnesses, (ii) making the organization more connected may increase corruption, and (iii) corruption will involve larger coalitions under better monitoring. Simulation results also suggest that more hierarchical organizations are more corrupt than flatter organizations. I test these predictions in the lab. Results confirm the predictions and reveal a systematic deviation that has implications for why better monitoring reduces corruption: participants disproportionately fail to realize larger coalitions, which are more necessary under good monitoring. Results suggest it would be sensible to redesign public agencies to puncture the isolation of enclaves.

© 2020 Elsevier Inc. All rights reserved.

A corrupt bureaucracy is a challenge to any government. With a cost of at least two percent of global GDP, corruption is rampant in developing countries and persists in developed countries (International Monetary Fund 2016). Existing theories of corruption usually rely on the principal-agent framework (for a review, see Olken and Pande, 2012). In their simplest form, these models often consider two dyads, where a welfare-maximizing principal optimizes some aspect of the environment of a potentially corrupt agent who interacts with a client. The metaphor best describes isolated acts of less profitable, *petty* corruption, such as a policeman pocketing a traffic bribe.[2] While the principal-agent approach is very well-suited to study how institutions affect individual acts of petty corruption, it is less equipped to study cases of more profitable, *grand* corruption – such as the 2015 FIFA scandal, where 25 top-ranking members of FIFA have been indicted for collusion with sports marketing executives (US Department of Justice, 2015a,b). Grand corruption usually involves complex networks of bureaucrats that often span across several divisions of their organization and whose members cooperate to subvert the institutions that were designed to deter them. In response, a range of theoretical and experimental work has moved beyond the standard principal-agent framework to consider how organizational structures affect corruption, or how several

E-mail address: rferrali@nyu.edu.

[1] I thank Stefano Caria, Johannes Haushoffer, Robert Gibbons, Guy Grossman, Rebecca Morton, Betsy Paluck, Ernesto Reuben, Bassel Tarbush, Leonard Wantchekon, as well as seminar participants at the CASBS, CSSO, and WESSI workshops, the Imai Research Group, Harvard University, Imai Research Group, Oxford University, and Princeton University for their helpful comments. I thank the staff of MEDA Solutions for their excellent assistance. I extend my special gratitude to Olivier Bochet, Matias Iaryczower, Kosuke Imai, Amaney Jamal, Jennifer Larson, and John Londregan for their comments and support. This work was generously supported by the Mamdouha S. Bobst Center for Peace and Justice (2015-06-Bobst).

[2] See, e.g., Besley and McLaren (1993) and Banerjee et al. (2013) for principal-agent approaches to the impact of wage incentives and monitoring technologies on corruption, respectively. See Abbink et al. (2002); Barr and Serra (2009); Ryvkin et al. (2017) for lab experiments on bribery.

https://doi.org/10.1016/j.geb.2020.08.013
0899-8256/© 2020 Elsevier Inc. All rights reserved.

corrupt agents may cooperate and form criminal networks.[3] Yet, both approaches suffer from a few shortcomings. Models of corruption in organizations usually consider highly stylized organizational structures (perfect hierarchies or perfectly flat organizations). This typology, however, hardly allows discriminating between large organizations such as FIFA for they are all largely hierarchical but show very different organizational charts. Models and experiments of collective corruption and criminal network formation usually occur outside of any pre-existing organizational structure, which misses a specificity of corruption: unlike other forms of organized criminal activities, corruption occurs within organizations that endow their members with colleagues who may be potential accomplices or witnesses.

This paper introduces a network approach to corruption that tells us how corruption networks form, and how this depends on the structure of the organization where corruption occurs. I examine a model and a lab experiment that look at corruption as the result of a diffusion process on a network (e.g. Polanski, 2007; Manea, 2017). In this setting, an agent may take an illegal rent. Organizations matter because they structure opportunities for corruption: the social ties they establish expose corrupt agents to witnesses, but may also be exploited to turn those agents into accomplices. Accomplices "cover up" for each other but cost resources, and may have other witnesses monitoring *them*. The agent faces a tradeoff between resources and secrecy: she needs to form the coalition that best protects her against detection, but costs a minimal fraction of the rent.

The model yields three core findings with substantive implications. First, characterizing how corruption is organized, I show that members of equilibrium coalitions jointly minimize exposure to witnesses; in other words, these *minimal* coalitions are the most *enclaved* parts of the organization. This result shifts the unit of analysis from the individual to the coalition, which may reconcile previous puzzling empirical findings about whether more or less connected individuals are more likely to be corrupt. The concept of minimal coalitions emphasizes the benefits of having few ties to the out-group, and is largely agnostic about the structure of ties within the in-group.

Second, changing the organizational structure has subtle effects on corruption. Changing organizational structures at the margin by adding ties to the network is no cure-all. Additional ties may decrease corruption by exposing minimal coalitions to additional witnesses. However, these ties have no effect if they do not target those minimal coalitions. Worse, they may increase corruption by facilitating access to existing minimal coalitions. I compare across organizational structures using simulations and reproduce the finding more hierarchical organizations are more corrupt, because hierarchies give birth to minimal coalitions. While the finding holds true in the aggregate, some noise remains, due to small structural details: relatively flat organizations may feature enclaves, while relatively hierarchical organizations can feature few such enclaves. This suggests that while it would be sensible to redesign government agencies to puncture enclaves, but that such redesigns may be counterproductive if not done carefully.

Third, better monitoring technologies reduce corruption but do not eliminate it, because accomplices adapt. Better monitoring increases the risk of detection, which makes buying off accomplices more attractive, and drives up the size of the coalition. Less profitable, petty corruption cannot cover the extra cost entailed by more accomplices and disappears; only grand corruption survives, making overall corruption less frequent. This provides a testable rationale for Kaufmann's (2004) observation that corruption persists in developed countries. I support this with preliminary evidence from a comparison of corruption cases in India and the US.

In a series of extensions, I show that predictions are largely robust to alternative assumptions about the cost of corruption, the informational environment, and the ability of corrupt individuals to strike informal contracts that police their interactions. Those extensions also add nuance to the results. Incomplete information about accomplices' willingness to cooperate with law enforcement penalizes larger coalitions, for they are more likely to contain agents that would report to law enforcement. Additionally, the lawlessness (Dixit, 2004) inherent to criminal activities introduces inefficiencies[4] that benefit *brokers* – accomplices that recruits other accomplices, – who exploit their control over the diffusion process to extract higher shares of the surplus. Agents may exploit two features to limit such inefficiencies in this environment. First, accomplices may leverage ties among one another. Additional ties among accomplices allow circumventing brokers, which reduces inefficiencies. Additionally, agents may leverage repeated interactions to devise self-enforcing contracts that solve brokers' commitment problems.

I test whether the main substantive implications of the theory are robust to behavioral traits that are assumed away in the model using a lab-in-the-field experiment. The lab has the dual advantage of allowing for easy inducement of network structures and easy measurement of corruption, two constructs that are notoriously hard to measure and manipulate in the field.

I introduce a minimal design in which subjects play a diffusion game analogous to the model in a face-to-face setting. The design is minimal in that it contains only the necessary ingredients to speak of corruption in organizations: (1) an organizational structure represented by a social network that conditions offers and monitoring, and (2) a lawless environ-

[3] See Tirole (1986); Laffont (1990); Laffont and Tirole (1991); McAfee and McMillan (1995); Melumad et al. (1995); Ting (2008) for principal-agent work on the impact of organizational structures on corruption. For theoretical work on corruption, see Andvig and Moene (1990); Shleifer and Vishny (1993); Burgess et al. (2012) for market approaches, and Baccara and Bar-Isaac (2008, 2009) for games of criminal network formation. Less related, see Calvó-Armengol and Zenou (2004); Ballester et al. (2010) for models of peer-effects in crime. Finally, see Gonzalez et al. (2002); Schickora (2011); Barr et al. (2009); Azfar and Nelson (2007); Berninghaus et al. (2013); Morton and Tyran (2015) for experiments on collective corruption.

[4] Throughout the paper, efficiency is understood from the point of view of corrupt agents. As such, while corrupt agents would like to maximize efficiency, a benevolent social planner would like to minimize it. See section 2.3 for a formal definition and further discussion.

ment where corruption is risky and where corrupt individuals cannot commit to recruit the agents that their accomplices find desirable. Being minimal, the design controls for potential sources of confounding, such as moral considerations about corruption, save for one. I allow participants to leverage cheap talk in order to devise informal contracts to alleviate the commitment problem. To increase ecological validity, I hold the experiment in Morocco, a mid-income country with median levels of corruption,[5] and compare a subject pool of service sector employees to a subject pool of undergraduate students.

The experimental data confirm the model's predictions. I consider a few small networks, and manipulate the monitoring technology and the profitability of corruption. I show that minimal coalitions are more corrupt, and only some ties reduce corruption. Consistently with model predictions, the overall incidence of corruption falls under better monitoring, but the corruption that does occur takes place on a larger scale, involving more accomplices. Behavior does not differ across subject pools and thus appears to be largely robust to factors outside the model. There is, however, one striking deviation from theoretical predictions: agents disproportionately fail to realize larger coalitions, presumably because they pose more challenging backward induction problems. This suggests another, behavioral reason as to why corruption decreases under better monitoring: better monitoring prompts for larger coalitions, which are more difficult to form.

Most closely related to the theoretical part of the paper are other models of strategic diffusion on networks, which study how information goods diffuse an on a network, starting from an initial pool of sellers (Polanski, 2007; Manea, 2017). This approach departs from the principal-agent framework by considering simpler forms of interactions between agents – notably, hierarchy is implicitly embedded in network structure instead of being an explicit feature,[6] – in order to model more realistic organizational structures. Closest to the experimental part of the paper is Berninghaus et al. (2013), whose design I supplement with an exogenous network, hence introducing a minimal design to assess the impact of organizational structure on corrupt behavior.

In the remainder of this paper, I first expose the model and derive the main theoretical propositions in a simple environment (section 1). I then consider a series of extensions (section 2). I finally take the model to the lab to test these propositions (section 3). I conclude by situating results in the literature, and discussing the main model assumptions and design choices made in the experiment.

1. A simple model of corruption as a criminal network

This section describes a simple model of corruption as criminal network formation. I describe the setting and the main results, that I complement with a series of conjectures derived from simulations. This simple environment delivers a rich set of results that are robust to a variety of extensions. I examine these extensions in the next section.

1.1. Setting

I model corruption as a dynamic game of complete information. The setting is very simple. Agents are the nodes of a network. Nature picks an agent at random, and is offered a rent of value 1. This agent, the *seed*, may reject the rent and end the game, or take it and initiate a diffusion process that results in the formation of a *coalition* of corrupt nodes. In this process, she formulates a vector of transfer offers to her neighbors. Her neighbors then respond to those offers sequentially and, if they accept, may similarly offer to transfer to their neighbors fractions of their holdings. Offers between any two neighbors can only be made once and the process carries on until either all nodes join the coalition, or no further offer can be extended. Once the process is over, an enforcer detects the coalition with some probability. In what follows, I describe this game formally, and introduce an important simplifying assumption; namely, that members of the coalition agree upon a *division rule* named *equal-sharing* whereby they divide the rent equally among one another. I relax this assumption in an extension (Section 2.3).

Agents are the nodes of the finite exogenous multiplex graph[7] $g \equiv (\mathcal{N}, \mathcal{G}_c, \mathcal{G}_m)$ where $\mathcal{N}$ is a set of nodes indexed from 1 to $|\mathcal{N}|$, and $\mathcal{G}_c$ and $\mathcal{G}_m$ are sets of ties. $\mathcal{G}_c$ is an undirected *communication* network, with $ij \in \mathcal{G}_c$ denoting a channel of communication between i and j. Communication ties allow existing members of the coalition to recruit new ones. Because organizations form a coherent unit, I assume that $\mathcal{G}_c$ is connected.[8] $\mathcal{G}_m$ is a directed *monitoring* network, with $i \to j \in \mathcal{G}_m$ meaning that i monitors j; in other words, i would hold incriminating evidence on j (and turn into a witness should j be corrupt). $\mathcal{G}_m$ is directed to allow for asymmetries of information: a manager may monitor her employees, but the converse may not be true. Furthermore, if two people do not interact, I assume that they do not know about each others' activities: $i \to j \in \mathcal{G}_m \Rightarrow ij \in \mathcal{G}_c$.

[5] Morocco is ranked 81 out of 180 countries in the Transparency International Corruption Perception Index 2017.

[6] Hierarchy is implicitly embedded in network structure because it has little impact on organizing corruption. Coercing lower-level employees to join the coalition is difficult: they are often critical to some task within the coalition, and know of the wrong-doing of their managers, which gives them leverage (Jávor and Jancsics, 2013). Regarding reporting corruption, a meta-analysis (Mesmer-Magnus and Viswesvaran, 2005) shows that hierarchy matters little compared to holding evidence, which stems from close interaction with the wrongdoer. The balance of accomplices and witnesses matters: larger coalitions face less risk of being reported because whistleblowers face a higher risk of retaliation.

[7] I use the terms graph and network interchangeably. A multiplex graph is a graph that contains several types of ties.

[8] That is, that there is a path on $\mathcal{G}_c$ between any $i, j \in N$.

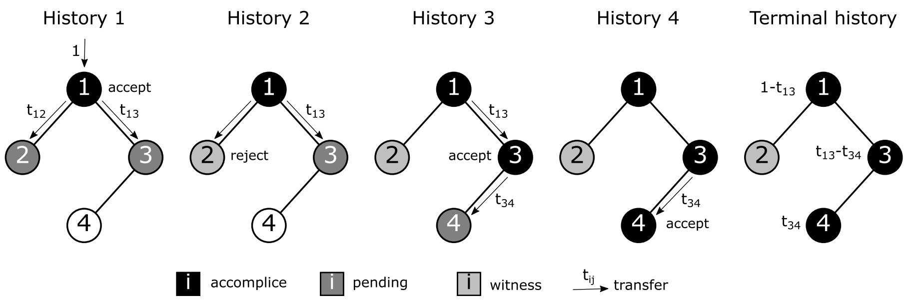


Fig. 1. Example diffusion process. Node 1 is the seed. Ties denote communication and mutual monitoring. At the terminal history, nodes 1, 3, 4 hold $1-t_{13}$, $t_{13}-t_{34}$, and t_{34} respectively.

A randomly drawn node $s \in \mathcal{N}$, the *seed*, discovers an illegal stream of rents whose value is normalized to 1 and which may represent a bribe or an opportunity for embezzlement. The seed can reject the rent or accept it.

- If she *rejects* the rent, the game is over, and all players gain 0.
- If she *accepts* the rent, she becomes an accomplice, and (1) she pays an exogenous cost $\epsilon \geq 0$; (2) all the agents that monitor the seed turn into witnesses; and (3) she makes the vector of offers t_s to the agents she communicates with.

The cost ϵ represents effort expended by the seed when engaging in corruption; for instance, learning about corrupt practices, or engaging in the criminal activity itself. Conversely, $1-\epsilon$ represents the benefit of corruption *relative to its cost*; that is, the scale of corruption. When ϵ is low, corruption is very profitable compared to its cost, indicating high-scale, grand corruption. Conversely, high values of ϵ indicate low-scale, petty corruption. While it seems natural to assume that the seed and her accomplices expend criminal effort irrespective of whether they get caught, it might be that the seed and accomplices incur a loss only when they get caught. I explore this possibility in an extension (see Online Appendix A).

Once the seed has made her offers, the nodes that have been made a strictly positive offer are *pending*. They move sequentially, with lower indices acting first, and face a similar action space.[9] They can reject the offer they have been made, or accept it. If node i accepts, she becomes an accomplice and holds the transfer t_{si}. Like the seed, she pays the cost ϵ, her non-pending, non-accomplice in-neighbors on $\mathcal{G}_m$ turn into witnesses, and she makes the vector of offers t_i to her *susceptible* neighbors; that is, her non-pending, non-accomplice neighbors on $\mathcal{G}_c$. Once all pending nodes have moved, the players to whom they have made offers (if any) can act. They face the same action space (i.e. reject their incoming transfer, or accept it and make transfer offers to their non-pending, non-accomplice neighbors), and their moving order is determined the same way. This process is repeated until no accomplice makes a positive offer, or until all nodes in g have become accomplices (Fig. 1).

There are four types of players at any history h: pending nodes, accomplices, witnesses and neutral nodes. *Pending* nodes are all the nodes that have been made an offer prior to history h and will play at, or after h. *Accomplices* are all the nodes that have accepted an offer to share the rent. Together, they form a criminal conspiracy, the *coalition*. *Witnesses* are the non-accomplice, non-pending in-neighbors of accomplices on $\mathcal{G}_m$. Finally, *neutral* nodes are all the remaining nodes, and do not play any role.

Coalition c on graph g has $a_c \equiv |c|$ accomplices. The set of witnesses of coalition c on g at a terminal history is $\mathcal{W}_{cg} \equiv \{i \in \mathcal{N} : i \to j \in \mathcal{G}_m, i \notin c, j \in c\}$, and $w_{cg} \equiv |\mathcal{W}_{cg}|$ the number of witnesses. Let $\mathcal{C}$ be the set of coalitions that can be formed on any graph with $|\mathcal{N}|$ nodes. A coalition c is feasible on graph g if it is consistent with some diffusion process originating from the seed; formally:

Definition 1. Let $\mathcal{C}(g,s) \subseteq \mathcal{C}$ be the set of feasible coalitions on graph g for seed s. A coalition $c \in \mathcal{C}(g,s)$ is *feasible* on g for seed s if for any node $i \in c$, there is a path between s and i on $\mathcal{G}_c$ such that all nodes on that path are in c.

Once the coalition is formed, an exogenous enforcer detects the coalition with probability $1-p$, where $p \equiv p(a, w, q)$: $\{1, ..., |\mathcal{N}|\} \times \{0, ..., |\mathcal{N}|\} \times (0,1) \to (0,1)$ is the coalition's probability of success. The probability of success p is a function of a, the number of accomplices in the coalition, w, its associated number of witnesses, and $q \in (0,1)$, a parameter for the *monitoring technology*, that captures the ability of an organization to detect and punish corruption. I make several additional

[9] Under this setup, offers are answered sequentially, which eliminates possibilities for multiple equilibria. Future research could usefully relax this assumption.

assumptions on p. Of course, better monitoring makes detection more likely; as such, $\frac{\partial p}{\partial q} < 0$. I also assume that p is increasing in a, with $p(a+1, w, q) - p(a, w, q) > 0$, and decreasing in w, with $p(a, w+1, q) - p(a, w, q) < 0$.

This specification for p incorporates, in a reduced form, a variety of existing results derived from the literature. First, witnesses increase the risk of detection (Baker and Faulkner, 1993), capturing the fact that accomplices can only imperfectly hide corruption from their non-corrupt colleagues who, in turn, have an incentive to report them to law enforcement – perhaps due to a moral inclination to act as a whistleblower, or because law enforcement compensates them to collect evidence. Second, additional accomplices pose a tradeoff. On the one hand, they help the coalition by decreasing the probability of detection, which captures deliberate activities to “cover up” illegal activities (Wade, 1982; Ledeneva, 1998).[10] On the other hand, they cost fractions of the rent and, *depending on the structure of the network*, they create additional witnesses among their colleagues, which increases risk. Finally, the specification makes detection all-or-nothing. This models in a reduced form the variety of self-enforcing contracts that criminals often use to prevent denouncing each other to law enforcement (Gambetta, 1996; Vannucci and Della Porta, 2013). These contracts prevent accomplices from denouncing each other prior to detection, but typically make the whole coalition collapse if one member gets caught.

If player i is not a member of the coalition at a terminal history her payoff is 0. Otherwise, she incurs cost ϵ and holds some share of the rent $\pi_i \geq 0$. For simplicity, I assume that accomplices implement a division rule where they divide the rent equally. In other words, they schedule transfers such that each member of the coalition c ends up with $\pi_i = \frac{1}{a_c}$.[11] With probability p, agent i gets her share π_i, and gets 0 otherwise. With risk-neutral agents, the expected utility of agent i writes

$$u_i(c, q) = \begin{cases} \frac{p(a_c, w_{cg}, q)}{a_c} - \epsilon, & \text{if } i \in c \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

To distinguish graphical and algebraic considerations, I define the *valuation* of a coalition $v : \{1, ..., |\mathcal{N}|\} \times \{0, ..., |\mathcal{N}|\} \times (0, 1) \to \mathbb{R}$. The valuation of a coalition is defined for an arbitrary number of accomplices and witnesses, with $v(a, w, q) = \frac{p(a,w,q)}{a}$, while the expected utility of a coalition $u_i : \mathcal{C}(g, s) \times (0, 1) \to \mathbb{R}$ is the valuation of existing coalitions on specific graphs, and incorporates the cost of corruption ϵ: $u_i(c, q) = 1\{i \in c\}[v(a_c, w_{cg}, q) - \epsilon]$.

The diffusion process leads to the formation of a *coalition* of accomplices: a corrupt subgraph of g. I analyze the Subgame Perfect Nash Equilibria (SPNE) of this game, and describe them focusing on three dimensions: *frequency*, the likelihood that some corruption occurs–whether the seed takes the rent; *scope*, the number of accomplices; and *scale*, capturing the profitability of corruption on a continuum from less profitable, petty corruption to more profitable, grand corruption, and represented by the quantity $1 - \epsilon$.

1.2. Results

This section derives equilibrium and comparative statics. All proofs are in the Appendix.

1.2.1. Characterization of equilibrium coalitions

When should the seed take the rent? I show that the seed has a threshold strategy where she rejects the rent below some threshold in scale. Consider her favorite coalition, $c^* \in \arg\max_{c \in \mathcal{C}(g,s)} u(c, g, q)$. If $u(c^*, q) < \epsilon$, then s does not take the rent, since her favorite coalition does not cover her cost. If $u(c^*, q) \geq \epsilon$, then the problem is more complicated: in principle, because incentives are dynamic, there is no guarantee that accomplices will cooperate to realize c^*. However, because accomplices divide the rent equally, incentives within the coalition are sequentially aligned: all accomplices value the same coalitions equally, and as such, members of c^* have no incentive to deviate to some other coalition. Formally:

Lemma 1.1 *(Threshold strategy). Let $\mathcal{C}^*(g, s, q) \equiv \arg\max_{c \in \mathcal{C}(g,s)} u_s(c, q)$ and $c^* \in \mathcal{C}^*(g, s, q)$. There is a threshold $\hat{\epsilon}_s(g, q) = v(a_{c^*}, w_{c^*g}, q) \in (0, 1)$ such that all equilibria have the same outcome where s rejects the rent if $\epsilon > \hat{\epsilon}(g, q)$. Otherwise, she accepts it, and some coalition $c \in \mathcal{C}^*(g, s, q)$ is realized.*

Saying more about which coalitions are realized in equilibrium requires characterizing the coalitions that the seed prefers. Yet, very different coalitions could be realized in equilibrium. I make a technical assumption that ensures that equilibrium coalitions are *essentially unique*; that is, that, even though there may be multiple equilibria, all equilibrium coalitions have the same counts of accomplices and witnesses. The assumption reads

[10] In reality, accomplices may also help by extracting more ressources (Jávor and Jancsics, 2013). To simplify interpretation, accomplices only help through their impact on the probability of detection, with the rent normalized to 1.

[11] Technically, this assumption is akin to either (1) restricting the strategy profiles to those that match the division rule, or (2) analyzing a different game in which agents only offer their neighbors to join the coalition, without making them a transfer, and each get transfered a π_i that matches the division rule once the coalition has formed.

Assumption 1.1. If $v(a_1, w_1, q) = v(a_2, w_2, q)$ for some $a_1 \leq a_2$, $w_1, w_2 \in \{0, ..., |\mathcal{N}|\}$, $q \in (0, 1)$, then $\frac{\partial p(a_2, w_2, q)}{\partial q} \Big/ \frac{\partial p(a_1, w_1, q)}{\partial q} \neq \frac{a_2}{a_1}$ for any $q \in (0, 1)$.

Assumption 1.1 implies that two essentially different coalitions may give the same payoff on at most a curve in (ϵ, q). As such, those two coalitions generically give different payoffs. Therefore, although there are many equilibria, generically, equilibrium coalitions are essentially unique. Formally:

Proposition 1.1 *(Essential uniqueness). Suppose Assumption 1.1 holds. Then equilibrium coalitions are essentially unique for any $(\epsilon, q) \in (0, \infty) \times (0, 1) \setminus U$, where U has measure zero.*

Having pinned down equilibrium allows characterizing how corruption varies within, and across organizations. Two questions seem particularly relevant. Keeping organizational structure constant, how does corruption change as organizations adopt better monitoring technologies? Keeping the monitoring technology constant, which parts of the organization are more likely to be corrupt, and how does corruption vary as the organizational structure changes?

1.3. How does corruption vary as monitoring improves?

Examining how corruption varies as organizations adopt better monitoring technologies requires additional assumptions about how various coalitions fare as the monitoring technology increase. In what follows, I present this assumption, detail the result that follows from it, and then discuss the assumption in light of this result. I assume that, compared to smaller coalitions, the additional protection afforded by larger coalitions does not vanish too quickly as the monitoring technology improves. The assumption strengthens Assumption 1.1 and reads as follows:

Assumption 1.2 *(Larger coalitions are sufficiently resistant against better monitoring).* If $v(a_1, w_1, q) = v(a_2, w_2, q)$ for some $a_1 \leq a_2$, $w_1, w_2 \in \{0, ..., |\mathcal{N}|\}$, $q \in (0, 1)$, then $\frac{\partial p(a_2, w_2, q)}{\partial q} \Big/ \frac{\partial p(a_1, w_1, q)}{\partial q} < \frac{a_2}{a_1}$ for any $q \in (0, 1)$.

Assumption 1.2 implies that corruption is less frequent but has a higher scale and a broader scope under better monitoring (Fig. 2). Because detection is more likely, the seed prefers giving away resources to benefit from the protection of additional accomplices, hence increasing the scope of corruption. Because larger coalitions are more costly, accepting the rent requires the project's scale to be high enough to offset this increase in costs: only projects with a high enough scale can now be sustained. As such, the seed now only takes the rent for large-scale projects. Corruption therefore becomes less frequent by *selecting* on grand corruption: under poor monitoring, the seed would both take the rent for small- and large-scale projects, while under good monitoring, she only takes the rent for large scale project. Formally:

Proposition 1.2 *(Corruption is less frequent under better monitoring technologies but grows in scope and selects on large-scale). Let $c_1^* \in \mathcal{C}^*(g, s, q_1)$, $c_2^* \in \mathcal{C}^*(g, s, q_2)$. If $q_1 < q_2$, then $\hat{\epsilon}_s(g, q_1) \geq \hat{\epsilon}_s(g, q_2)$. If Assumption 1.2 holds, then $q_1 < q_2 \Rightarrow a_{c_1^*} \leq a_{c_2^*}$.*

Secondary literature and additional empirical results provide empirical support for Proposition 1.2. Kaufmann (2004) shows that, although it is less frequent than in developing countries, corruption persists in developed countries. He also shows that while developing countries feature both petty and grand corruption, developed countries only show grand corruption. Yet, Proposition 1.2 also provides a mechanism for this stylized fact: better monitoring prompts for more protection, which can only be afforded by additional accomplices. Online Appendix D provides empirical support for this mechanism, using a comparison of 110 cases of corruption in India and the US. Controlling for the scale of corruption, instances of corruption in the US – a country that presumably has better monitoring than India – involve more accomplices than in India.

Assumption 1.2 seems fairly reasonable in light of Proposition 1.2. First, it is a relatively weak assumption, because $\frac{a_2}{a_1} > 1$. As such, the assumption does not require that the marginal effect of monitoring be smaller for larger coalitions (i.e. $\frac{\partial^2 p}{\partial a \partial q} > 0$) but only that, should larger coalitions suffer more from improved monitoring than smaller ones, this disadvantage be small enough. Second, reversing Assumption 1.2 (i.e. assuming that $\frac{\partial p(a_2, w_2, q)}{\partial q} / \frac{\partial p(a_1, w_1, q)}{\partial q} > \frac{a_2}{a_1}$) would make Proposition 1.2 state that as monitoring increases, the size of the coalition decreases, which contradicts existing empirical evidence (Online Appendix D).

1.4. How does corruption vary within and across organizations?

Studying how corruption varies within and across organizations requires answering two related questions. First, within given organizations, which nodes and/or sets of nodes are more likely to be corrupt? Second, which kinds of network structures best mitigate corruption? Analytical tools have little traction for these questions, because (1) nodes are interdependent;

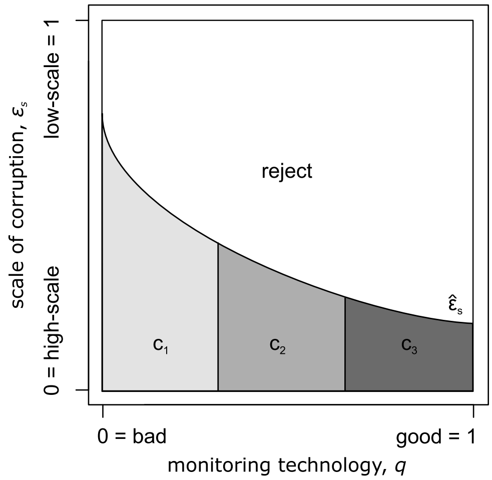


Fig. 2. Equilibrium outcomes for any (q, ϵ)**.** As q increases, the rejection area grows, weeding out low-scale corruption, and coalitions of increasing size are realized (Proposition 1.2).

therefore, changing a node's structural attribute changes the structure of the entire network, hence making comparisons across nodes difficult, and (2) two networks can differ in a variety of ways, making comparisons across networks difficult. As such, I complement analytical results with computational results derived from simulations. The remainder of this section presents a series of analytical results, moves on to computational results, and then discusses all results together in light of the literature.

Analytical results

The setup allows an easy characterization of the sets of nodes that are more likely to be corrupt. Since the probability of success p is decreasing in the number of witnesses w, comparing same-sized coalitions, one prefers the one with fewer witnesses. I define these coalitions as *minimal.*

Definition 2. $\mathcal{M}(g,s) \equiv \{c \in \mathcal{C}(g,s) : w_{cg} \leq w_{c'g}$ for any $c' \in \mathcal{C}(g,s)$ such that $a_{c'} = a_c\}$ is the set of *minimal coalitions* on graph g for seed s. Let $\mathcal{M}_a(g,s) \equiv \{c \in \mathcal{M}(g,s) : a_c = a\}$ is the set of minimal coalitions on graph g for seed s and size a.

As such, equilibrium coalitions must be minimal.

Proposition 1.3 *(Equilibrium coalitions are minimal). If $c \in \mathcal{C}^*(g,s,q)$ for some $q \in (0,1)$, then $c \in \mathcal{M}(g,s)$.*

Proposition 1.3 has an important implication: within an organization, minimal coalitions should be more corrupt than non-minimal coalitions. Minimality captures the idea that a coalition is *jointly* isolated from the out-group; in other words, this coalition has few monitoring ties pointing to it from the out-group. I say that a set of nodes that has few monitoring ties pointing to it from the out-group is relatively *enclaved*. Minimal coalitions are the most enclaved coalitions of a graph.

Proposition 1.3 also has an immediate implication for the kinds of nodes that are more likely to be corrupt. Since only minimal coalitions are realized in equilibrium, then for node i to be corrupt, she must belong to a minimal coalition.

Corollary 1.3.1 *(Corrupt nodes belong to minimal coalitions). If node $i \in \mathcal{N}$ is corrupt in equilibrium, then $i \in c$ such that $c \in \mathcal{M}(g,s)$ for some $s \in \mathcal{N}$.*

Since saying more about which nodes are more likely to be corrupt using analytical methods is intractable, I supplement this result with a series of simulations.

The last set of analytical results investigates how corruption varies as the organizational structure changes. I characterize analytically the impact of a marginal change to an existing organization: adding a tie to the network. The exercise reveals that making organizations more connected is no cure-all. It helps when adding monitoring ties but hurts when adding communication ties. Additional monitoring ties expose existing coalitions to more witnesses, which makes taking the rent more risky, and decrease the frequency of corruption. Additional communication ties, however, do not make existing coalitions more exposed. Worse, they may allow forming new, more enclaved coalitions, hence increasing corruption. Most additional ties, however, have no effect: because corruption only occurs in minimal coalitions (Proposition 1.3), ties may change predictions only if they affect those, which is increasingly unlikely as the graph gets larger.

Consider graphs g and g', constructed either by adding a monitoring tie to g ($g' = g + i \to j$), or a communication tie ($g' = g + ij$). We have

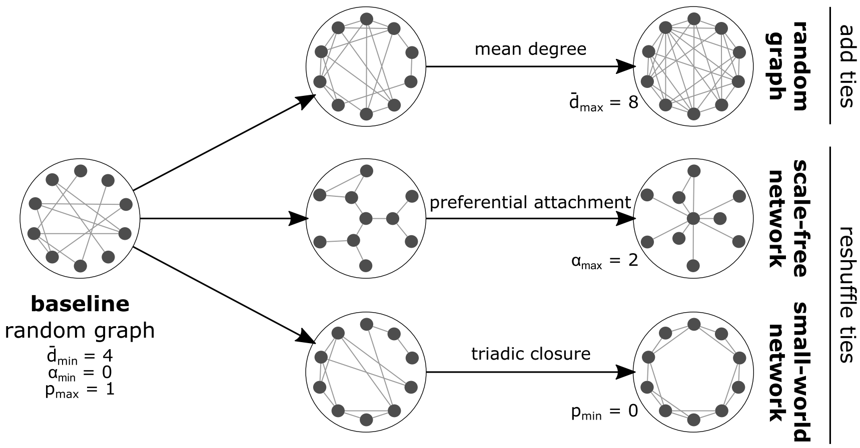


Fig. 3. Simulation parameters. The parameters $\bar{d}$, α, p are mean degree, preferential attachment (Barabási and Albert, 1998), and rewiring probability (Watts and Strogatz, 1998) respectively. Each series of simulations samples 699 values uniformly from the parameter ranges in the figure. The graphs represented in this figure are stylized representations of the output of each of the generative models used in the simulations.

Proposition 1.4 *(Adding monitoring ties weakly decreases the frequency of corruption). If $g' = g + i \to j$, then $\hat{\epsilon}_s(g', q) \leq \hat{\epsilon}_s(g, q)$ for all $s \in \mathcal{N}$. Furthermore, if there is $s \in \mathcal{N}$ and $q \in (0, 1)$ such that $\hat{\epsilon}_s(g', q) < \hat{\epsilon}_s(g, q)$, then there is $a \in \{1, ..., |\mathcal{N}|\}$ such that for all minimal coalition $c \in \mathcal{M}_a(g, s)$, $j \in c$ and $i \notin c \cup \mathcal{W}_{cg}$.*

Proposition 1.5 *(Adding communication ties weakly increases the frequency of corruption). If $g' = g + ij$, then $\hat{\epsilon}_s(g', q) \geq \hat{\epsilon}_s(g, q)$ for all $s \in \mathcal{N}$.*[12]

Computational results

Simulations examine the extent to which a randomly chosen seed takes the rent for graphs of varying modularity. I consider the expected area under the $\hat{\epsilon}_s$ curve (AUC) for a random seed (Fig. 2). For seed s on graph g, the AUC writes $AUC_{sg} \equiv \int_0^1 \hat{\epsilon}_s(g, q)dq$. Its expected value is $AUC_g \equiv \mathbb{E}(AUC_{ig})$, where higher values of AUC_g denote more corruption. Simulations are computationally intensive, for finding the minimal coalitions of a graph of size N requires enumerating its connected subgraphs which, using the Depth-First-Search Enumeration algorithm (Skibski et al., 2019), is $\mathcal{O}(|\mathcal{C}(g)||\mathcal{G}|)$. As such, simulations consider moderately-sized graphs (36 nodes) in which communication and monitoring ties are collapsed. Simulations use the following probability of success[13]:

$$p(a, w, q) = 1 - \left[q + \frac{w}{N-1}(1-q) - \frac{a-1}{N-1}q\right] \tag{2}$$

This function has several properties that make it appealing. In the absence of social structure, detection depends only on the monitoring technology: $p(1, 0, q) = 1 - q$. It is linear in a and w. Success is certain when the whole organization is corrupt, with $p(|\mathcal{N}|, 0, q) = 1$. Symmetrically, detection is certain when the all remaining members of the organization are witnesses, with $p(1, |\mathcal{N}| - 1, q) = 0$.

I conduct three series of simulations that each depart from the same baseline: a random graph with a mean degree of 4. Each series of simulations samples a specific kind of network in order to explore the impact of a structural feature on corruption. The first simulation adds ties. It samples random networks with mean degree varying from 4 to 8. The other two simulations fix mean degree to 4, but reshuffle ties. I look into the effect of flatter vs. more hierarchical organizations by sampling scale-free networks (Barabási and Albert, 1998) and varying the preferential attachment parameter α from 0 to 2. As α increases, nodes depart from a random graph and become increasingly hierarchical, as increasingly large hubs appear. Finally, I investigate the effect of triadic closure (i.e. the extent to which i's friends are friends with one another) by sampling small-world networks (Watts and Strogatz, 1998), and varying the rewiring parameter from 0 (a lattice) to 1 (a random graph). For each of these series of simulations, I draw 699 samples using a uniform distribution on the parameter space. Fig. 3 summarizes the parameters.

I then analyze the results in two ways. First, I examine each series of simulations separately to see how each graph-level structural feature impacts overall corruption AUC_g. Second, I pool all three series together and examine how node-

[12] Note that as in Proposition 1.2, Propositions 1.4 and 1.5 imply that when the frequency of corruption decreases, it selects on higher-scale projects.
[13] See SI, Appendix C for proof that p satisfies Assumption 1.2.

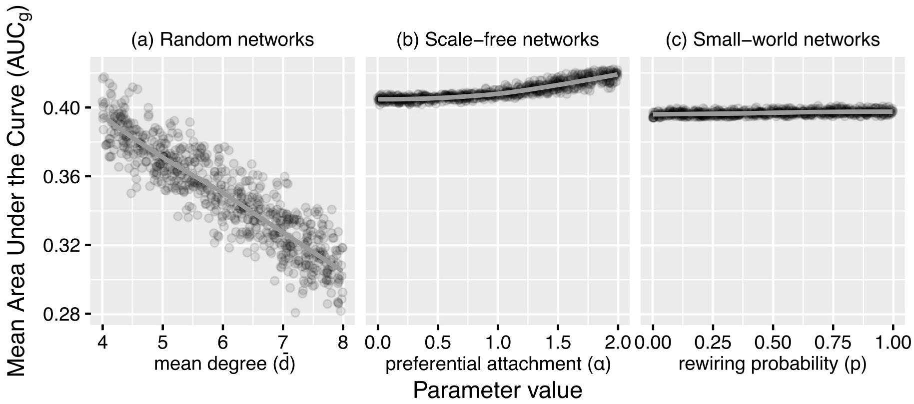


Fig. 4. Simulation results, aggregate level. The grey line is a loess fit. Denser organizations are less corrupt (panel a). More hierarchical organizations are more corrupt (panel b). Triadic closure has no discernible effect on corruption (panel c).

level structural features impact individual-level corruption AUC_{ig}. I examine two such features: degree and the clustering coefficient; that is, the percentage of closed triads in node i's neighborhood.[14]

Of course, making *ceteris paribus* claims when it comes to network structure is challenging. At the network level, changes to one network statistic usually correlate with changes to other network statistics. At the node level, adding or removing a tie to node i also affects other nodes. While the randomization procedure employed in this set of simulations should alleviate these concerns, one should keep in mind that those results are inherently correlational.

Fig. 4 reports results at the graph level. Denser organizations are less corrupt (panel a), suggesting that the corruption-reducing effect of additional monitoring ties (Proposition 1.4) more than offsets the corruption-increasing effect of additional communication ties (Proposition 1.5). More hierarchical organizations are more corrupt (panel b). Indeed, as the organization increasingly resembles a star, i.e. as it becomes increasingly centered around one large hub, such hubs become more exposed, which decreases their corruption, but spokes become more enclaved, which increases their corruption. Since there are more spokes than hubs, corruption increases overall. Finally, clustering has little effect on corruption (panel c). This is because ties that close triads are *redundant*, in that they neither increase monitoring nor allow forming new coalitions. Consider seed s with ties to j and k. Because s is connected to both j and k, she can form coalitions that include either nodes. Furthermore, should s and j be corrupt, k would be a witness because of its tie to s. As such, the tie jk is redundant, as highlighted in the following proposition:

Proposition 1.6 *(Increasing clustering has no effect). Consider connected graph g such that $ij \in \mathcal{G}_c \iff i \to j \in \mathcal{G}_m$ and $j \to i \in \mathcal{G}_m$. Within g, consider a node $s \in \mathcal{N}$ that has two neighbors i, j such that $ij \notin \mathcal{G}_c$. Construct graph g' by adding $ij, i \to j, j \to k$ to g. We have that $AUC_{sg} = AUC_{sg'}$.*

Finally, results suggest that the effect of adding ties (panel a) is more important than that of reshuffling them (panels b and c).

Results at the individual level, reported in Table 1 confirm the intuition derived from aggregate-level results. Nodes with a higher degree are less corrupt, because they are more observable. Consistently with Proposition 1.6, clustering has little effect: the effect size is very small – one order of magnitude smaller than that of degree – and flips signs when introducing graph-level fixed effects. That the correlation is significant at the 1% level is unsurprising given sample size (see Online Appendix B for power calculations).

Discussion

Analytical results revealed that more enclaved coalitions – that is, *minimal* coalitions – are more corrupt, because they minimize the number of witnesses (Proposition 1.3). As a consequence, members of those coalitions are more likely to be corrupt (Corollary 1.3.1). However, making organizations more connected may have a dual effect. On the one hand, additional ties enable further monitoring, which weakly decreases corruption (Proposition 1.4). On the other hand, additional ties enable further communication, which weakly increases corruption by potentially allowing the formation of new minimal coalitions (Proposition 1.5). Computational results showed that sparser, or more hierarchical organizations correlave with more corruption (Fig. 4, panels a and c), while there is little correlation between clustering and corruption (panel b). Indeed, sparser and/or more hierarchical organizations leave more room for enclaves, which furthers corruption. Conversely,

[14] Other measures of centrality, such as betweenness, closeness, or eigenvector centrality highly correlate with degree centrality, and hence yield similar results. Results using these measures are available upon request.

Table 1
Simulation results, individual level. The Table reports OLS estimates with standardized effect sizes. Standard errors are clustered at the graph-level. Model 2 includes graph-level fixed effects. Higher degree nodes are less corrupt. Nodes with a higher clustering coefficient are more corrupt, but the effect size is small (one order of magnitude smaller than that of degree).

	Dependent variable:	
	AUC_{ig}	
	(1)	(2)
degree	− 0.060***	− 0.058***
	(0.0005)	(0.0004)
clustering	0.003***	− 0.001***
	(0.0001)	(0.0001)
Graph-level FE	_	✓
Observations	74,053	74,053
R^2	0.946	0.965
Note:	* p < 0.1; ** p < 0.05; *** p < 0.01.	

clustering introduces redundant ties that have no effect (Proposition 1.6). Results carry over to the individual level (Table 1): higher degree nodes correlate with fewer corruption, while clustering coefficient has little correlation with corruption.

The concept of minimal coalitions has an important implication: it shifts the unit of analysis from the individual to the coalition, and is, to some extent, agnostic about tie structure within the in-group. In other words, what matters is only whether a coalition is relatively enclaved. The structure of ties within the in-group is irrelevant from this model's point of view. This is, of course, a direct consequence of the payoff function, which ignores the structure of ties within the in-group. This insight needs, however, to be nuanced for two reasons. First, we will see in an extension (Section 2.3) that within-coalition ties may facilitate solving commitment problems. Second, coalitions are interdependent. As such, what constitutes a within-coalition tie for one coalition may well constitute a tie to the out-group for another coalition. Dense coalitions have therefore little chance of containing minimal sub-coalitions.

Overall, the concept of minimal coalitions may reconcile a series of mixed findings about the structure of criminal networks. Aven (2015) and Morselli et al. (2007) show that criminal networks are sparser than comparable non-criminal networks, but there is also evidence that better connected individuals are more corrupt (Nyblade and Reed, 2012; Khanna et al., 2015). Collectively, minimal coalitions are sparsely connected to the rest of the organization. However, accomplices may have many ties with each other, and some of them may be exposed to many witnesses. Overall, this prompts for distinguishing ties between accomplices and ties between the coalition and its witnesses.

Minimal coalitions also nuance an old insight from the principal-agent literature – that flatter organizations limit corruption by making the actions of agents more observable to the principal (McAfee and McMillan, 1995; Melumad et al., 1995). Allowing to move beyond contrasting perfect hierarchies to perfectly flat organizations, computational results nuance this insight on two counts. First, results show that while the insight holds true at the aggregate level, there still is variation across organizations that are equally hierarchical. Structural details matter: enclaves may appear in relatively flat organizations and reciprocally, relatively hierarchical organizations may comport few enclaves, depending on the exact layout of communication and monitoring ties. Second, computational results shift the definition of a flat organization. In the proposed approach, a flat organization corresponds to a random graph, a moderately hierarchical organization corresponds to a tree, and a very hierarchical organization to a star. In the principal-agent approach, stars are held to be flat organizations, while lines are held to be hierarchical. The difference in results stems from the weight given to different players. In the proposed approach, players are homogeneous; as such, being observed by the center of the star only creates one additional witness. Conversely, in the principal-agent approach, the center of the star is the principal, who holds a special role.

Finally, analytical and computational results suggest how endogenous network formation may affect the results. Propositions 1.4 and 1.5 imply that corrupt agents would like to sever monitoring ties and add communication ties. Conversely, a benevolent social planner would rather add monitoring ties and sever communication ties. Simulations suggest that corrupt agents should prefer sparser, more hierarchical organizations, while a benevolent social planner should prefer denser and flatter organizations.

2. Extensions

The simple model analyzed in section 1 made a series of stark assumptions that I relax in this section. First, the model assumed that petty and grand corruption are equally likely to be detected. Yet, petty and grand corruption may react differently to law enforcement. For instance, being more profitable, grand corruption might afford better protection against detection. Conversely, grand corruption might be more salient, hence easier to detect. I examine both possibilities. Second, the model assumed complete information, which is unrealistic for corruption, a phenomenon that is characterized by secrecy. I examine an extension with incomplete information in which agents have different (privately known) propensities to cooperate with law enforcement. Third, the model assumed that accomplices divided the rent equally among themselves.

The assumption essentially assumed a *contractual environment*, where accomplices could commit ex-ante to a contract specifying a specific division rule. I relax this assumption by examining other division rules, as well as a *lawless environment* where agents cannot commit to a division rule and instead divide the rent endogenously through take it or leave it offers. I finally examine a repeated game in the lawless environment, which leaves room for contracts to emerge endogenously through repeated interactions.

While each of these extensions add nuances to the simple model examined above, the core findings from the simple model are all robust to these extensions, under qualitatively similar assumptions. Specifically, under all these extensions, (1) better monitoring reduces the frequency of corruption and selects on grand corruption (Proposition 1.2), (2) equilibrium coalitions are minimal (Proposition 1.3), and (3) adding monitoring ties weakly decreases corruption (Proposition 1.4), while adding communication ties weakly increases corruption (Proposition 1.5).

2.1. Detection as a function of the scale of corruption

The simple model analyzed in section 1 assumed that the probability of detection is independent of the scale of corruption. This assumption simplifies the analysis, but may be unrealistic. However, as mentioned above, it is unclear how the scale of corruption should affect the probability of detection. On the one hand, more profitable, grand corruption could be more salient, hence more likely to be detected. On the other hand, grand corruption might allow agents to spend some of the additional profit to thwart efforts by law enforcement. I consider both cases. Assuming that grand corruption is less likely to be detected does not change the results. Assuming that grand corruption is more likely to be detected, results do not change if the effect is sufficiently small. Conversely, if the effect is sufficiently large, then one result changes: as monitoring improves, corruption now decreases by weeding out grand corruption instead of petty corruption (but still involves more accomplices). Because empirical evidence is more supportive of the original result (see section 1.3), I favor the assumption that the probability of detection either decreases for more profitable schemes, or does not increase by much. The rest of this subsection details changes in the setting, and the main changes in the results, leaving to Appendix B.1 results that only change in their formulation.

Setting

In this extension, I assume that $\epsilon \in (0, 1)$, and make the probability of success dependent on ϵ by amending the original probability of success as follows:

$$\tilde{p}(a, w, q, \epsilon) = \rho(\epsilon) p(a, w, q), \tag{3}$$

where $\rho : (0, 1) \to (0, 1)$ is twice-differentiable and rescales the probability of success according to ϵ. Recall that $1 - \epsilon$ measures the scale of corruption, with large values of ϵ indicating petty corruption. Assuming that $\rho'(\epsilon) > 0$ makes grand corruption more likely to be detected, while assuming $\rho'(\epsilon) \leq 0$ makes petty corruption more likely to be detected. Note that this formulation implicitly assumes that the effect of scale on detection is independent of the composition of the coalition. This is in the spirit of the motivating question: how would results change if grand corruption was more or less likely to be detected than petty corruption, *independently* of the composition of the supporting coalition? Payoffs now write:

$$u_i(c, q, \epsilon) \equiv \begin{cases} \frac{\tilde{p}(a_c, w_{cg}, q, \epsilon)}{a} - \epsilon, & \text{if } i \in c \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

Assumptions 1.1 and 1.2 are maintained.

Results

Most of the intuition does not change. For a fixed level of monitoring q, when considering whether to accept the rent, the seed looks at $\mathcal{C}(g, s)$ and considers the utility of her favorite coalition. Because the effect of scale on detection is independent of the composition of the coalition, the seed's favorite coalition stays the same for any ϵ. When grand corruption is less likely to be detected than petty corruption, then that coalition becomes increasingly profitable as ϵ decreases. As such, the seed accepts all projects above some threshold in scale. When grand corruption is more likely to be detected than petty corruption, but the effect is not too strong, then the fact that grand corruption is more profitable offsets the fact that it is more risky. The seed still accepts all projects above some threshold in scale. Conversely, when that effect is very strong, although grand corruption is more profitable, it is too risky. As such, the seed accepts all projects below some threshold in scale (i.e. the pettiest projects). Lemma 1.1 changes to accommodate this new result. The lemma now reads:

Lemma 2.1. *Let $\mathcal{C}^*(g, s, q) \equiv \arg\max_{c \in \mathcal{C}(g,s)} u_s(c, q, \epsilon)$. We have that $c \in \mathcal{C}^*(g, s, q)$ for some $\epsilon > 0$ if and only if $c \in \mathcal{C}^*(g, s, q)$ for any $\epsilon \in (0, 1)$. Furthermore, there is a threshold $\hat{\epsilon}_s(g, q) \in [0, 1]$ such that:*

- *If $\rho'(\epsilon) \leq \frac{a_c}{p(a_c, w_{cg}, q)}$ for any $c \in \mathcal{C}(g, s)$, $q \in (0, 1)$, $\epsilon \in (0, 1)$, then all equilibria have the same outcome where s rejects the rent if $\epsilon > \hat{\epsilon}_s(g, q)$. Otherwise, she accepts it, and some coalition $c^* \in \mathcal{C}^*(g, s, q)$ is realized.*

- *If $\rho'(\epsilon) > \frac{a_c}{p(a_c, w_{cg}, q)}$ for any $c \in \mathcal{C}(g, s), q \in (0, 1), \epsilon \in (0, 1)$, then all equilibria have the same outcome where s rejects the rent if $\epsilon < \hat{\epsilon}_s(g, q)$. Otherwise, she accepts it, and some coalition $c^* \in \mathcal{C}^*(g, s, q)$ is realized.*

As a consequence, the result that corruption decreases by selecting on grand corruption as monitoring improves (Proposition 1.2) changes to accommodate Lemma 2.1. Corruption still decreases as monitoring improves. When the seed selects projects above some threshold in scale, then the result remains unchanged: corruption decreases by selecting on grand corruption. However, when the seed selects projects below some threshold in scale, then corruption decreases by selecting on petty corruption. The new proposition reads:

Proposition 2.1. *Let $c_1^* \in \mathcal{C}^*(g, s, q_1), c_2^* \in \mathcal{C}^*(g, s, q_2)$. If $\rho'(\epsilon) \leq \frac{a_c}{p(c, g, q, \epsilon_s)}$ for any $c \in \mathcal{C}(g, s), q \in (0, 1), \epsilon \in (0, 1)$, then $q_1 < q_2 \Rightarrow \hat{\epsilon}(g, q_1) \geq \hat{\epsilon}(g, q_2)$. If $\rho'(\epsilon) > \frac{a_c}{p(c, g, q, \epsilon)}$ for any $c \in \mathcal{C}(g, s), q \in (0, 1), \epsilon \in (0, 1)$, then $q_1 < q_2 \Rightarrow \hat{\epsilon}(g, q_1) \leq \hat{\epsilon}(g, q_2)$. If Assumption 1.2 holds, then $q_1 < q_2 \Rightarrow a_{c_1^*} \leq a_{c_2^*}$.*

Other propositions are left virtually unchanged. They are available in Appendix B.1.

2.2. Incomplete information

The simple model analyzed in section 1 assumed complete information. This assumption may be unrealistic in an environment that is characterized by secrecy. In this extension, I assume that some agents may not be suitable partners, and that agents lack information on who is a suitable partner. Specifically, I assume that some agents may cooperate with law enforcement by reporting their fellow accomplices, and that there are two types of agents, whereby high types have a higher incentive to cooperate with law enforcement than low types. I rule out by assumption the uninteresting cases where no agent ever cooperates with law enforcement or all agents always cooperates with law enforcement, and focuses on the case where high types always cooperate with law enforcement, while low types cooperate if the coalition contains high types.

The extension introduces a few changes. First, when high-type agents have an incentive to cooperate with law enforcement while low types do not, larger coalitions are at a disadvantage, because they have a higher probability of containing a high type. Second, equilibria largely depend on assumptions about how high types resolve indifference conditions, which complicates the analysis. Most results from the simple model travel under assumptions that are comparable to the original Assumptions 1.1 and 1.2, with one exception: minimal coalitions are no longer realized in equilibrium (Proposition 1.3), unless high types always pool with low types.

I amend the original setting by adding types and a second stage to the simple model. There are two private types H (igh) and L (ow). Agents of type H have stronger incentive to cooperate with law enforcement than agents of type L. Before the game begins, Nature assigns to each player $i \in \mathcal{N}$ type $\tau_i \in \{H, L\}$ with probability $r \equiv \Pr(\tau_i = L) \in (0, 1)$. Players then enter the *coalition formation stage*, during which they proceed to form a coalition as in the simple model. Once the coalition has formed, the *enforcement stage* begins, in which a law-enforcer offers agents to provide her with hard evidence against the coalition. If the enforcer gathers a piece of hard evidence, then the coalition gets detected with probability 1. Otherwise, the enforcer detects the coalition with probability p as in the simple model. Only coalition members have access to such hard evidence, and the enforcer rewards such evidence with benefit $b \geq 0$. If they provide hard evidence, agents incur a cost κ for betraying fellow coalition members. This cost is lower for high-type agents than for low-type agents: $\kappa_H < \kappa_L$.

Let $d_i = 1$ if agent i cooperates with law enforcement, and $d_i = 0$ otherwise. If i is a member of coalition c, she pays cost ϵ. If no accomplice cooperates with law enforcement – that is, if $d_j = 0$ for all $j \in c$, – then she gets her share of the rent $p(a_c, w_{cg}, q)/a_c$. If she cooperates with law enforcement, then she gets the benefit b but pays the cost κ_{τ_i} associated with her type τ_i. Utilities therefore write:

$$u_i(c, q) = \begin{cases} d_i(b - \kappa_{\tau_i}) + \frac{p(a_c, w_{cg}, q)}{a_c} \prod_{j \in c}(1 - d_j) - \epsilon, & \text{if } i \in c \\ 0 \text{ otherwise.} \end{cases}$$

I analyze this extension using Markov Perfect Equilibrium (MPE) as a solution concept. To simplify the analysis, I focus on pure-strategy equilibria that are *symmetric* in the enforcement stage; that is, equilibria in which all players of a given type have the same strategy.

I assume that all agents $i \in \mathcal{N}$ have a symmetric prior $\Pr_i(\tau_j = L) = r$ for any $j \neq i \in \mathcal{N}$. Note that since we focus on pure-strategy equilibria and players only move once, it must be that at any information set h, the posterior $\Pr_i(\tau_j = L|h)$ satisfies $\Pr_i(\tau_j = L|h) \in \{0, r, 1\}$. In other words, for any strategy profile σ and at any information set h, player j has either revealed her type (i.e. $\Pr_i(\tau_j = L|h) \in \{0, 1\}$) or not (i.e. $\Pr_i(\tau_j = L|h) = r$).

I characterize equilibria using backward induction, and first focus on the enforcement stage. At this stage, members of the coalition c may have either revealed their type or not in the coalition-formation stage. I make the following assumption on payoffs:

Assumption 2.1. Parameters b, κ_L, and κ_H are such that for any $s \in \mathcal{N}$, and any $c \in \mathcal{C}(g,s)$,

$$b - \kappa_H > \frac{p(a_c, w_{cg}, q)}{a_c} > (r^{a_c - 1})\frac{p(a_c, w_{cg}, q)}{a_c} > b - \kappa_L$$

This assumption rules out the uninteresting cases in both types cooperate or none do. Instead, with the assumption, high types always cooperate with law enforcement, because their payoff from cooperating $b - \kappa_H$ is greater than the payoff of a coalition that she knows contains no other high types ($\frac{p(a_c, w_{cg}, q)}{a_c}$). Conversely, low types cooperate only if a high type has revealed her type in the coalition formation stage: their payoff from cooperating $b - \kappa_L$ is lower than the payoff of a coalition in which she ignores the types of all members. Formally:

Proposition 2.2. *Suppose that coalition c with $a_c > 1$ was formed in the coalition formation stage. Let $Pr_i(\tau_j)$ the belief hed by i over the type of j. If Assumption 2.1 holds, then in equilibrium*

1. *If $\tau_i = H$, then i cooperates with law enforcement.*
2. *If $\tau_i = L$ then i cooperates with law enforcement if and only if $\Pr_i(\tau_j = L) = 0$ for some $j \neq i \in c$.*

I maintain Assumption 2.1 for all results of this extension. Note that this assumption has low types cooperate only if a high type has revealed her type. An alternative could be that low types cooperate only if they are certain that all other members of c are low types, assuming instead that $\frac{p(a_c, w_{cg}, q)}{a_c} > b - \kappa_L > r\frac{p(a_c, w_{cg}, q)}{a_c}$. The alternative assumption preserves all the results deriving from Assumption 2.2, and adds one nuance. I present results using Assumption 2.2 and discuss the alternative in the end of this section.

Having characterized equilibrium in the enforcement stage, consider now the enforcement stage. An important question is whether equilibria are pooling or separating. From Proposition 2.2, high types always cooperate with law enforcement in the enforcement stage. As such, they are willing to join any coalition in the coalition-formation stage, to then pocket the reward from cooperation. Conversely, Proposition 2.2 states that low types only cooperate if they are certain that the outcome coalition contains at least one high type. As such, low types are only willing to join coalitions in which the risk of containing a high type is sufficiently low. Furthermore, as soon as low types learn that the coalition they are forming contains a high type, they also become indifferent as to the specific outcome of the coalition formation stage, since they will pocket the benefit from cooperation with law-enforcement. In this context, a separating equilibrium would have high types behave differently from low types *once they have joined the coalition*. For instance, high types could extend no offer after having received the rent.

Note however that fully separating equilibria – i.e. equilibria in which all members of the outcome coalition reveal their type – may well not exist. If coalition c is an equilibrium, then low-type must have an incentive to join c. Additionally, we have seen that high types have an incentive to join any coalition and that as such, a separating equilibrium would separate between high and low types by having them make different offers once they have joined the coalition. Yet, this may well not be feasible. Indeed, the path leading to the formation of coalition c may have histories with action sets that only contain the actions of accepting or rejecting offers to join the coalition – for instance, when a player joins the coalition at a history in which all her neighbors are accomplices. It is impossible to construct separating equilibria for the players that move at those histories.

Note finally that separating equilibria may appear unrealistic. Indeed, if high types are whistleblowers that infiltrate the coalition in order to report it to law-enforcement, it seems somewhat inconsistent to assume that such whistleblowers would successfully mimick low types in order to join said coalition but then reveal their type once they have joined it. In other words, pooling equilibria assume that high types are most competent; they perfectly blend in with low types until the enforcement stage, and therefore post the most challenging learning problem.

With that in mind, we can now describe equilibrium behavior: if strategy profile σ is a MPE, then the seed considers all the coalitions she may form and, under σ, how many members reveal their type in each of these coalitions. She then picks the coalition that maximizes her expected payoff. Suppose that $l_{-ic\sigma}$ members of c different from i do not reveal their type under profile σ – if the profile σ satisfies full pooling, then $l_{-ic\sigma} = a_c - 1$ for any i, c. Then with probability $r^{a_c - l_{-ic\sigma} - 1}$, all of the members of c that reveal their type are of type L. With probability $r^{l_{ic\sigma}}$, the remaining $l_{ic\sigma}$ members are also of type L. As such, with probability $r^{a_c - l_{ic\sigma} - 1} r^{l_{ic\sigma}} = r^{a_c - 1}$, an agent of type L pockets the benefit $u(c, g)$. With probability $1 - r^{a_c - l_{ic\sigma} - 1}$, at least one member of c that reveals her type under σ turns out to be of type H. In this event, an agent of type L cooperates, and earns the payoff from cooperation $\overline{u} = b - \kappa_L$. As such, the expected payoff to the seed for any coalition at the initial information set h_0 is $\mathbb{E}_s[u_s(c, g, q)|h_0, \sigma] = r^{a_c - 1}u(c, g, q) + (1 - r^{a_c - l_{sc\sigma} - 1})\overline{u}$. Assuming that c^* is the coalition that maximizes the seed's payoff, it is realized in equilibrium, because as the game unfolds and no agent is revealed to be of type H, coalition c^* becomes more attractive to subsequent agents. In other words, if coalition c^* was more attractive than coalition c at h_0, it will be even more attractive at a later history, since subsequent information revealed that some members of c^* are low types. As such, equilibrium behavior is similar to the main model, and Lemma 1.1 becomes:

Lemma 2.2. *Suppose $\tau_s = L$. If σ is a MPE then there is a threshold $\hat{\epsilon}_s(g, q, \sigma) \in (0, 1)$ such that s rejects the rent if $\epsilon < \hat{\epsilon}_s$. Otherwise, she accepts it and a coalition $c \in C^*(g, s, \sigma) \equiv \arg\max_{c \in \mathcal{C}(g,s)} \mathbb{E}_s[u_s(c, g, q)|h_0, \sigma]$, with*

$$\mathbb{E}_s[u_s(c,g,q)|h_0,\sigma] = r^{a_c-1}u(c,g,q) + (1 - r^{a_c-l_{sc\sigma}-1})\overline{u} - \epsilon,$$

is realized in any state where $\tau_i = L$ for any $i \in c$ that reveals her type on the path of play from h_0 under σ.

Immediately, incomplete information puts larger coalitions at a disadvantage, because they have a higher probability to contain at least one high type. Among essentially equal coalitions, pooling proves to be a disadvantage, since it reduces the probability of pocketing the reward from cooperation $\overline{u}$. In the limit equilibrium where all high-type agents pool, $l_{sc\sigma} = a_c - 1$ for any coalition $c \in \mathcal{C}(g,s)$; that is, no agent expects to pocket the reward from cooperation, since high types never reveal their type at the coalition formation stage.

The rest of the results remain largely unchanged. Assumptions 1.1 and 1.2 must also change slightly, in order to accommodate this change in the payoff function. With $v(a_c, w_{cg}, q, \sigma) \equiv \mathbb{E}_s[u_s(c,g,q)|h_0,\sigma]$, Assumptions 1.1 and 1.2 become:

Assumption 2.2. If $v(a_1, w_1, q, \sigma) = v(a_2, w_2, q, \sigma)$ for some $a_1 \le a_2$, $w_1, w_2 \in \{0, ..., |\mathcal{N}|\}$, $q \in (0,1)$, and some MPE profile σ, then

$$\frac{\partial p(a_2, w_2, q, \sigma)}{\partial q} \Big/ \frac{\partial p(a_1, w_1, q, \sigma)}{\partial q} \neq \frac{a_2}{a_1 r^{a_2-a_1}}$$

for any $q \in (0,1)$.

Assumption 2.3. If $v(a_1, w_1, q, \sigma) = v(a_2, w_2, q, \sigma)$ for some $a_1 \le a_2$, $w_1, w_2 \in \{0, ..., |\mathcal{N}|\}$, $q \in (0,1)$, and some MPE profile σ, then

$$\frac{\partial p(a_2, w_2, q, \sigma)}{\partial q} \Big/ \frac{\partial p(a_1, w_1, q, \sigma)}{\partial q} < \frac{a_2}{a_1 r^{a_2-a_1}}$$

for any $q \in (0,1)$.

Note that Assumptions 2.2 and 2.3 are weaker than their counterparts from the simple model, since $\frac{a_2}{a_1} \le \frac{a_2}{a_1 r^{a_2-a_1}}$. Recall that Assumption 1.2 is instrumental to the result that as monitoring improves, larger coalitions are realized (Proposition 1.2), because it guarantees that larger coalitions do not disproportionately suffer from better monitoring. Because the extension puts larger coalitions at a disadvantage, those coalitions may also suffer additional losses from better monitoring until they become too unattractive.

Furthermore, results must slightly change in their formulation to accommodate the fact that the coalition that solves $\max_{c \in \mathcal{C}(g,s)} \mathbb{E}_s[u_s(c,g,q)|h_0,\sigma]$ is only realized in states in which $\tau_i = L$ for any $i \in c$ that reveals her type on equilibrium path under profile σ, and that the threshold $\hat{\epsilon}_s(g,q,\sigma)$ now varies with σ. In particular, with $\mathcal{S}(s,g,q)$ the set of MPE profiles for seed s on graph g with monitoring q, we define $\hat{\epsilon}_s(g,q) \equiv \max_{\sigma \in \mathcal{S}(s,g,q)} \hat{\epsilon}_s(g,q,\sigma)$.

Additionally, under this extension, the result that minimal coalitions are realized in equilibrium (Proposition 1.3) does not hold anymore, because the payoff from a coalition also depends on whether agents reveal their type under the equilibrium profile. To see why, consider two coalitions c_1 and c_2 such that $a_{c_1} = a_{c_2}$ and $w_{c_1 g} < w_{c_2 g}$. Under the simple model, c_1 yields a higher payoff than c_2 and is therefore realized in equilibrium. However, under this extension, it may be that there is a profile σ such that $l_{sc_1\sigma} \gg l_{sc_2\sigma}$, to the point that c_1 becomes less attractive than c_2. In fully pooling equilibria however, Proposition 1.3 holds again, since all coalitions with the same counts of accomplices and witnesses yield the same payoff. I detail the changes in Appendix B.2.

Finally, recall that Assumption 2.1 assumed that low-type agents preferred the payoff from the coalition than the reward from cooperation, as long as they are not certain that a member of the outcome coalition is a high type; i.e. $\frac{p(a_c, w_{cg}, q)}{a_c} > r^{a_c-1}\frac{p(a_c, w_{cg}, q)}{a_c} > \overline{u}$. This assumption admits an alternative; namely that low-type agents are more risk-averse, and only prefer the payoff from the coalition to the reward from cooperation if they are certain that the outcome coalition contains no high type: $\frac{p(a_c, w_{cg}, q)}{a_c} > \overline{u} > r\frac{p(a_c, w_{cg}, q)}{a_c}$. With more risk-averse agents, low types cooperate in a broader subset of equilibria that under Assumption 2.1, since it suffices that one high-type agent pools for low types to cooperate in the enforcement stage. Results do not change significantly otherwise.[15]

2.3. Alternative ways of dividing the surplus

The simple model assumed *equal-sharing*; that is, it assumed that agents divide the rent equally among accomplices. This simplifying assumption is akin to a contract whereby accomplices commit ex-ante to a contract defining a division rule. Additionally, the assumption gives much traction, because it aligns incentives within the coalition which, in turn, was key to proving Lemma 1.1. Relaxing this assumption raises new interrogations. How would accomplices distribute the rent

[15] Proofs available upon request to the author.

endogenously? Do different division rules prompt for different coalitions? Could such division rules emerge endogenously? In what follows, I first assume a lawless environment; that is, I consider the sequential take it or leave it bargaining that equal-sharing assumed away. I refer to this environment as "bargaining" or "lawless environement" interchangeably. I show that lawlessness introduces a commitment problem that creates inefficiencies. I then return to the contractual environment and contrast equal-sharing with another division rule, *monopoly*. Under monopoly, the seed pockets all the surplus. While equal-sharing assumed maximally egalitarian accomplices, monopoly assumes maximally unegalitarian accomplices. I show that monopoly achieves more efficiency than equal-sharing, and that all results from the simple model travel to this division rule. I finally consider repeated interactions within the lawless environment to show that efficient contracts (i.e., contracts that implement the same coalitions as monopoly) may emerge endogenously through repeated interactions.

Before moving further, let's define efficiency. The surplus is the expected benefit from the rent, net of the cost ϵ paid by each accomplice. A coalition is efficient if it maximizes the surplus. An equilibrium profile is efficient if its equilibrium outcome is an efficient coalition. Since the diffusion process requires that feasible coalitions $\mathcal{C}(g,s)$ include the seed, I use a local notion of efficiency, and only look for efficient coalitions among the coalitions that are feasible for a given seed. Furthermore, the notion is defined from the players' viewpoint. While players would like to maximize efficiency, in order to realize the coalition that maximizes the surplus, a benevolent social planner would like, conversely, to minimize efficiency and thereby generate as little corruption as possible. Formally, let $\Pi(c,g,q) \equiv p(a_c, w_{cg}, q) - a_c\epsilon$ be the surplus of coalition c on graph g for monitoring level q, and define $\Pi(\emptyset, g, q) \equiv 0$, the payoff from the empty coalition that is realized when the seed rejects the rent. Let $\mathcal{E}(g,s) \equiv \{c \in \mathcal{C}(g,s) : c \in \arg\max_{c\in\mathcal{C}(g,s)} \Pi(c,g,q)\}$. Then,

Definition 3. An equilibrium profile is *efficient* if for any seed s, its equilibrium coalition c satisfies $c \in \mathcal{E}(g,s)$.

Additionally, note that bargaining and monopoly create multiple equilibria, some of which arise from uninteresting resolution of indifference conditions. I rule them out by considering equilibria that satisfy *deference*. Consider a strategy profile. I say that a node is a *broker* in this strategy profile if she recruits other nodes. I call *operatives* the nodes that do not recruit other nodes in this strategy profile (e.g. because they only have one neighbor, the one who recruited them). In equilibrium, an accomplice may be indifferent between her broker's favorite outcome and that broker's outside option. There may be an equilibrium in which she picks the outside option. To rule out this case, I consider equilibria that satisfy *deference*; that is, equilibria where if node i_0 is indifferent, she defers to the preference of her broker i_1. If i_1 is indifferent, then she defers to i_1's broker i_2, etc. Formally:

Definition 4. Consider strategy profile σ. Suppose that node $i_0 \neq s$ moves at history h, and considers the path of brokers $s = i_k \to i_{k-1} \to \dots \to i_0$, $k \geq 1$ leading to i_0's offer. Let $u_i(a,h,\sigma)$ be the payoff to agent i from taking action $a \in \mathcal{A}_h$ at history h under profile σ. The strategy profile σ satisfies *deference* if, whenever $u_{i_0}(a_1,h,\sigma) = u_{i_0}(a_2,h,\sigma)$ for any $a_1, a_2 \in \mathcal{A}_h$, and there is $i_j \in \{i_1, \dots, i_k\}, a_j^* \in \mathcal{A}_h$ such $u_{i_j}(a_j^*, h, \sigma) \geq u_{i_j}(a', h, \sigma)$ for any $a' \in \mathcal{A}_h$, then i takes the action a_j^* of the node i_j that has the lowest index j.

Lawlessness introduces a commitment problem that benefits brokers. Operatives do not recruit other nodes on the equilibrium path. As such, they only need to be made indifferent between accepting and rejecting their incoming transfer, and receive the smallest possible transfer. Brokers recruit other nodes on equilibrium path. They cannot commit to the equilibrium schedule of transfers, and may have a profitable deviation in making no transfers, or different ones. Consequently, their incoming transfer must be larger, to make them indifferent between the equilibrium schedule and these deviations: brokers extract more surplus than operatives because they have more outside options.

Fig. 5 summarizes all division rules considered in this section, and exemplifies the commitment problem (bottom panel). Consider a profile leading to the formation of coalition $c \in \mathcal{C}(g,s)$. Node i is indifferent between joining c or not if her holdings $\pi_i = \frac{\epsilon}{p(a_c, w_{cg}, q)}$. In our example, the outcome is coalition $c = \{1,3,4\}$, which has node 2 as a witness. Node 3 receives transfer $t_{13} = \pi_3 + \pi_4$ from node 1. Under monopoly, nodes 3 and 4 pocket none of the surplus. Their need is indifferent between joining the coalition or not, so they both gain $\pi_i = \frac{\epsilon}{p(3,1,q)}$. Under bargaining, since node 4 is an operative, she is also only indifferent between joining the coalition or not. Thus $\pi_4 = \frac{\epsilon}{p(3,1,q)}$. Node 3, however, is a broker, and keeping π_4 instead of transferring it to node 4 may prove a profitable deviation. As such it must be that $\pi_3 \geq \pi_4$. The following proposition generalizes the intuition:

Proposition 2.3 *(Brokers extract more surplus). In the lawless environment, if c is an equilibrium coalition, then $u_i(c,q) \geq 0$ for any $i \in c$. If i is an operative, then $u_i(c,q) = 0$.*

An important implication is that under lawlessness, additional communication ties may facilitate corruption for an additional reason. We saw that in the simple model, additional ties facilitate corruption because they may allow forming more enclaved coalitions (Proposition 1.5). Under lawlessness, additional ties may ease commitment problems for the same coalitions, because, by giving more direct access to nodes, they turn would-be brokers into operatives. This is especially true of ties that introduce redundancies. As such, while in the contractual environment, ties among accomplices did not matter, they do matter under lawlessness because they facilitate the resolution of commitment problems.

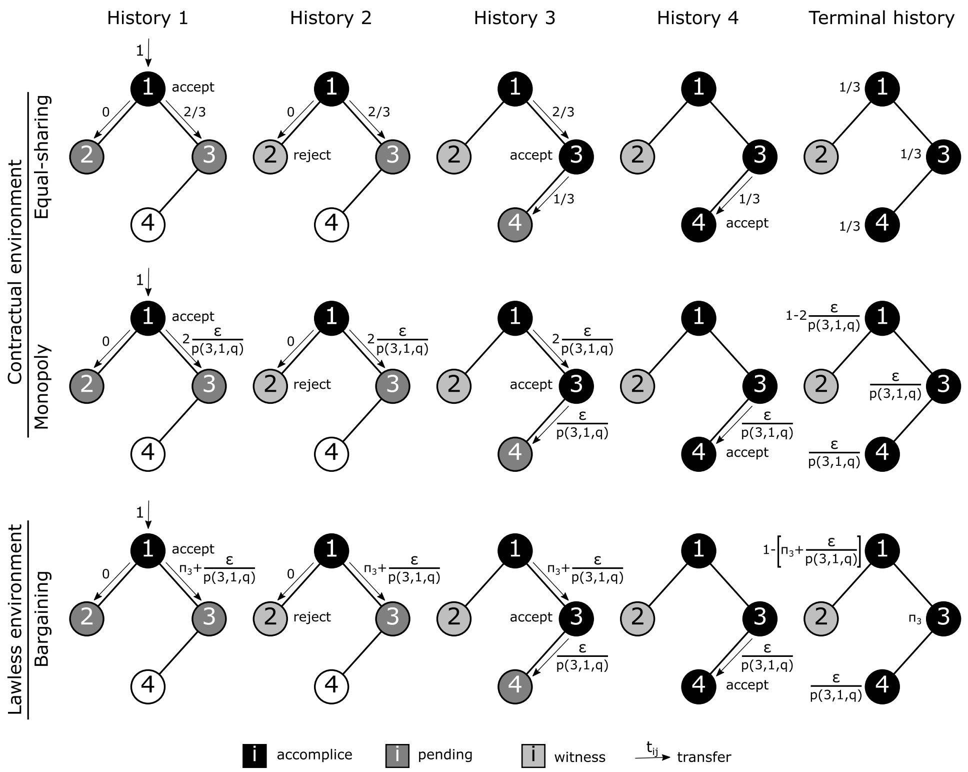


Fig. 5. Example diffusion process under different division rules. This figure reproduces the diffusion process examined in Fig. 1 with the transfers implied by different division rules.

Note finally that in the lawless environment, all nodes have threshold strategies such that they join the coalition if the transfer they receive is above some threshold, and make a vector of transfers that implements the coalition that gives them the highest payoff among all coalitions in $\bar{\mathcal{C}}(g,h) \subseteq \mathcal{C}(g,s)$, the set of all coalitions that can be formed at this history. Formally:

Proposition 2.4 *(Threshold strategies in the lawless environment). Let $\bar{\mathcal{C}}^*(g,h,q) \equiv \arg\max_{c \in \bar{\mathcal{C}}(g,h)} u_i(c,q)$. Suppose node i answers offer t_{ji} at history h. There is a threshold $t^*_{ji} \geq 0$ such that all equilibria of the subgame that begins at history h have the same outcome where i rejects t_{ji} if $t_{ji} < t^*_{ji}$. Otherwise, she accepts it, some coalition $c \in \bar{\mathcal{C}}^*(g,h,q)$ is realized, and she makes a (possibly null) vector of transfers $t^*_i(h,c)$.*

I now move on to comparing the efficiency of all three division rules (bargaining, equal-sharing, and monopoly). Doing so requires pinning down the seed's equilibrium behavior. In all division rules, the seed has a threshold strategy:

Proposition 2.5 *(Threshold strategy, extension). Under the monopoly rule and in the lawless environment, the seed has a threshold $\hat{\epsilon}_s > 0$ such that she rejects the rent if $\epsilon > \hat{\epsilon}$. Otherwise, some coalition in $\mathcal{C}(g,s)$ is realized.*

Knowing when the seed takes the rent allows to discuss efficiency. The monopoly rule is efficient because it aligns the seed's incentives towards efficient coalitions, while making other accomplices equally satisfied with any coalition. Indeed, under monopoly, the seed's payoff writes $u_s(c,g,q) = p(a_c, w_{cg}, q) - a_c\epsilon = \Pi(c,g,q)$. Other arrangements are less efficient. Under lawlessness, brokers pose two problems. First, although accomplices all agree on which coalitions provide most protection, brokers' incentives may still not align with the seed's: depending on the distribution of the rent, they may disagree on the optimal coalition. Second, even when some distribution of the rent would make brokers and the seed better off, brokers may not be able to commit to implementing it. These problems may be so acute as to prevent the seed from taking the rent, making corruption less frequent than under monopoly. Equal-sharing is qualitatively more efficient than lawlessness, for it aligns the incentives of accomplices. Accordingly, corruption is as frequent as under monopoly. Equal-sharing is, however, less efficient than monopoly, because it veers incentives towards smaller coalitions. Indeed, compared to monopoly, the seed's share is smaller in larger coalitions, leading her to favor smaller ones. The next result encapsulates the discussion:

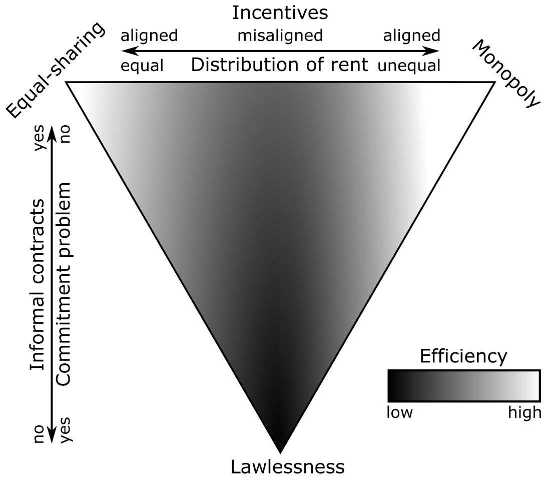


Fig. 6. Effect of informal contracts on cooperation among accomplices. Stronger informal contracts and contracts that align incentives increase efficiency. Monopoly is more efficient than equal-sharing.

Proposition 2.6 *(Efficiency). The monopoly rule admits an efficient equilibrium profile. Let $\hat{\epsilon}^m$, $\hat{\epsilon}^e$, $\hat{\epsilon}^l$ be the seed's thresholds under monopoly, equal-sharing, and lawlessness respectively. We have $\hat{\epsilon}^m = \hat{\epsilon}^e \geq \hat{\epsilon}^l$. Let $\mathcal{C}^m$ and $\mathcal{C}^e$ be sets of equilibrium coalitions under equal-sharing and monopoly for some (q, ϵ). Then $\min_{c \in \mathcal{C}^e} a_c \leq \min_{c \in \mathcal{C}^m} a_c$, and $\max_{c \in \mathcal{C}^e} a_c \leq \max_{c \in \mathcal{C}^m} a_c$.*

I then examine the robustness of the findings to these new division rules. I show in Appendix B.3 that all the propositions derived under equal-sharing hold in the monopoly rule under similar assumptions. None of the findings travels to the lawless environment, because the cost of a coalition depends on the outside options of all of its brokers. Without strong assumptions, there is considerable heterogeneity in how the price of such outside options varies with parameter values, which upsets the regularities observed in the contractual environment.

Fig. 6 considers these three division rules jointly to suggest how informal contracts affect cooperation among accomplices (Fig. 6). Coordination problems among accomplices create inefficiencies and have two sources: commitment problems and misaligned incentives among accomplices due to distributional considerations. Lawlessness compounds both problems. Increasingly binding contracts address the commitment problem by setting payoffs in each coalition ex-ante; as such, efficiency increases as one moves from lawlessness to the top of the simplex in Fig. 6. The contracts considered here also align incentives within the coalition, and display the regularities discussed in the previous section. Consequently, in Fig. 6, moving from the middle of the simplex to the left or to the right – that is, to perfectly egalitarian or inegalitarian contracts – increases efficiency. Although monopoly is more efficient than equal-sharing, both contracts show that informal contracts can be welfare-enhancing.

While the previous analysis has shown that contracts may improve efficiency, it is unclear whether agents can enforce such contracts endogenously. By definition, corruption is illegal, which implies that corrupt agents cannot enforce contracts in a court of law. Yet, a large body of literature shows that repeated interactions among criminals may support a variety of informal institutions that allow enforcing informal contracts (e.g. Gambetta, 1996).

To answer the question, I analyze a repeated game Γ, where agents repeat infinitely the stage game Γ_l in the lawless environment. At each period $t \in \mathbb{N}$, a seed $s_t \in \mathcal{N}$ is drawn at random from a distribution $f : \mathcal{N} \to [0, 1]$ that needs not be full-support.[16] Agents discount the future with common rate $\delta \in (0, 1)$. Outcome coalitions $c \in \mathcal{C}$ are now indexed by t. The present value of the flow of future payoffs $U_{i\tau}$ for agent i from period $\tau \in \mathbb{N}$ is given according to the function

$$U_{i\tau} = (1 - \delta) \sum_{t=\tau}^{\infty} \delta^{t-\tau} u_i(c_t, q),$$

where the function u_i remains unchanged from the lawless environment. If agent i is a member of coalition c_t and holds a share of the rent π_{it}, her payoff at that period is $u_i(c_t, q) = \pi_{it} p(a_{c_t}, w_{c_t g}, q) - \epsilon$. Her payoff is 0 if $i \notin c_t$.

Analyzing this game is non-trivial because the stage game is itself dynamic. In this setting, the folk theorem does not readily apply. Indeed, cooperation when the stage game is static is typically ensured by resorting to minimax threats. However, while those threats are credible in a static stage game, they may not be credible in all the subgames of a dynamic stage game.

[16] That f needs not be full-support allows, in particular, for different agents to have different opportunities for corruption. Specifically, some agents may never be the seed and have $f(i) = 0$.

It turns out that the repeated game Γ admits an efficient equilibrium profile if agents are sufficiently patient. This profile amends the SPNE of the stage game at the margin. Let σ_s be a SPNE of the stage game $\Gamma_l(s)$ that begins with seed $s \in \mathcal{N}$, and suppose that it has coalition c_s as an outcome. Agents play according to σ_s if c_s is efficient. Otherwise, they implement an efficient coalition c_s^* and make sure that each $i \in c_s^*$ pockets a strictly higher share of the surplus than she would under σ_s. This is possible because c_s^* generates more surplus than c_s. Deviations from c_s^* by agent $i \in c_s^*$ are prevented using the threat of reverting back to σ_s. Since c_s^* grants i a higher payoff than c_s, if i is sufficiently patient, she has no incentive to deviate. I now state the results formally, starting by defining our candidate strategy profile σ.

Definition 5 *(Candidate strategy profile σ).* For any seed $s \in \mathcal{N}$, pick a strategy profile σ_s that is a SPNE of the stage game $\Gamma_l(s)$. If the outcome coalition c_s of σ_s satisfies $c_s \in \mathcal{E}(g,s)$, then agents follow σ_s. Otherwise, they implement a schedule of transfers that has coalition $c_s^* \in \mathcal{E}(g,s)$ as an outcome, with payoffs that satisfy $u_i(c_s^*, q) > u_i(c, q)$ for any $i \in c_s^*$ and such that no offer is rejected. Agents follow σ_s at any off-path history of the stage game. Should any agent deviate on the path of the stage game, agents revert back to σ_s at any subsequent time period where s is picked as the seed.

I show that σ can be sustained in equilibrium if players are sufficiently patient.

Proposition 2.7 *(σ can be sustained in equilibrium). There is $\bar{\delta} \in (0,1)$ such that for any $\delta \geq \bar{\delta}$, the strategy profile σ is a SPNE of Γ.*

By construction, the strategy profile σ is efficient. Since it has the same equilibrium outcomes as the monopoly rule, it also displays the same regularities as in the simple model, for the same assumptions (see Appendix B.3 for details).

Overall, we have seen that contracts may improve efficiency over bargaining, because they solve a commitment problem, and may align incentives among accomplices. In particular, one division rule (monopoly) implements efficient coalitions. Proposition 2.7 shows that such division rule may be attained endogenously using repeated interactions. Additionally, these profiles need not be as unequal as monopoly, since the candidate profile σ allows for a variety of ways to allocate the surplus.

3. Experiment

We learnt from the theory that corruption in organizations exhibits a series of regularities. First, as organizations adopt better monitoring technologies, corruption becomes less frequent but increases in scope and selects on high-scale projects (Proposition 1.2). Second, more enclaved subgraphs are more corrupt, because they generate fewer witnesses (Proposition 1.3). Third, some ties make corruption less frequent, some increase corruption, and others have no effect, depending on whether they enable additional monitoring, or reaching better accomplices (Propositions 1.4 and 1.5). These results hold under a variety of assumptions about the way accomplices divide the surplus (section 2.3). Specifically, they hold whenever accomplices manage to devise informal contracts that commit them ex-ante to some division of the surplus; in particular, either equal division of the rent, or assignment of all the surplus to the seed. While lawlessness introduces significant noise and inefficiencies to the way accomplices cooperate (Proposition 2.6) and upsets the abovementioned regularities, agents may be able to devise such contracts endogenously (Proposition 2.7).

I now take the model to the lab in order to test whether the regularities highlighed by theory can be given any empirical support, irrespective of the assumptions that underlie them. Indeed, if those regularities are not borne out empirically, then no set of assumptions is relevant. A second order goal is then to examine which set of assumptions best matches the behavior of agents.

I examine an experimental design that primarily tests whether the regularities hold, and leaves some room to examine which assumptions over the division of the surplus best describe behavior. To do so, I examine a baseline condition characterized by a graph g, a monitoring technology q, and a cost of corruption ϵ. I derive a series of treatments that manipulate g, q, ϵ, and consider a payoff function and graphs such that the comparative in Propositions 1.2 to 1.5 hold for any q, ϵ, for all division rules examined in the theory, *including* the one-shot, lawless environment. I then pick parameter values q, ϵ such that each treatment has the same predicted coalition for all division rules. I simply compare across treatments and show that behavior largely matches the predictions from Propositions 1.2 to 1.5. I then look at how participants divided the surplus within each treatment to try to match it to the division rules examined in the paper.

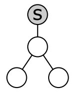
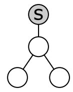
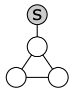
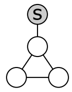
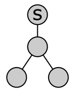
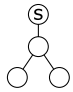
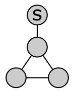
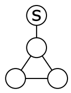
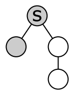
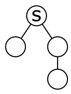
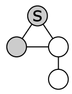
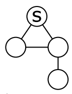
3.1. Design

In the experiment, subjects are put in groups of four players to play 12 repetitions of the diffusion game under lawlessness Γ_l in various treatment conditions. I first describe the basic setup of the game, common to all treatment conditions, then the treatment conditions and their associated predicted outcomes, followed by the setup of a lab session.

3.1.1. Basic setup

The baseline game occurs in an environment with face-to-face environment where interactions are mediated by an enumerator. Four subjects sit at a table. They are endowed with ϵ discrete experimental units (EU) and assigned an index number and network positions through a network diagram drawn on a paper handout.

Table 2
Experimental parameters and equilibria. s is the seed. Grey nodes are coalition members. Each group of participants plays either under grand or petty corruption for an entire session. They play each treatment condition for 2 repetitions. The number of games indicates the total sample size in each condition.

Condition	Monitoring (q)	Irrelevant tie	Predicted equilibrium coalition: Grand corruption ($\epsilon = 2$) (48 groups)	Predicted equilibrium coalition: Petty corruption ($\epsilon = 4$) (20 groups)
Baseline	.1	– (2 repetitions)	 (96 games)	 (40 games)
		✓ (2 repetitions)	 (96 games)	 (40 games)
Hard	.75	– (2 repetitions)	 (96 games)	 (40 games)
		✓ (2 repetitions)	 (96 games)	 (40 games)
Exposing tie	.1	– (2 repetitions)	 (92 games)	 (40 games)
		✓ (2 repetitions)	 (92 games)	 (40 games)

As in the game Γ_l in the lawless environment, one player is designated as the seed and offered a rent of 12 EU. Rejecting the rent terminates the game and has each player win their endowment. Accepting it initiates a diffusion process mediated by the enumerator. Accomplices – i.e., players that hold some of the rent – sequentially make take-it-or-leave-it transfer offers to their non-accomplice neighbors, who then sequentially accept/reject these offers and, if they accept, give up their initial endowment and make subsequent offers to their neighbors.[17] The process ends when either all players are accomplices, or no offer can be made anymore. Then, non-accomplices each earn their endowment ϵ. The enumerator randomly determines whether accomplices earn their share of the rent, using a paper handout listing the probability of success p from equation (2) for all feasible coalitions,[18] and a hundred-sided dice as a randomization device.

For simplicity, all networks used in the experiment collapse the monitoring and communication networks into one undirected network: if two players are able to communicate, they also monitor each other. All interactions are public and communication for transfers are mediated by the enumerator to implement the sequential take-it-or-leave-it offers of the diffusion process. Cheap talk is otherwise allowed, to allow participants to devise informal contracts as in the contractual environment.

3.1.2. Treatments and predictions

I compare a baseline condition to a series of treatments to test the main predictions of the model. Those treatments are summed up in Table 2. Within each condition, I vary the profitability of corruption by manipulating the endowment ϵ held by subjects, using either $\epsilon = 4$ EU (petty corruption), or $\epsilon = 2$ EU (grand corruption). The treatments manipulate the graph g, the cost of corruption ϵ, and the monitoring technology q. The graphs g and the probability of detection p used in the experiment (equation (3)) are such that the comparative statics in Propositions 1.2 to 1.5 hold true for any q, ϵ irrespective

[17] See section 1.1 for details about how the sequence of offers is determined.
[18] The experiments rescale p by a factor of .83: $\tilde{p}(a, w, q) = .83p(a, w, q)$, to ensure that all outcomes are uncertain. Indeed, since $p(N, 0, q) = 1$, the rescaling implies $\tilde{p}(N, 0, q) = .83$.

of whether subjects use equal-sharing, monopoly, or bargaining when dividing the surplus.[19] As such, I picked q, ϵ such that the predicted coalitions are the same under all division rule for each treatment condition. This setup puts most weight on the first-order goal of the experiment – that is, testing for the main predictions of the model irrespective of assumptions about the decision rule, – because those main predictions are tested through experimental manipulation. The setup also allows, to some extent, exploring the division rule used by agents. Indeed, although agents evolve in a lawless environment, they may leverage cheap talk in order to devise informal contracts. As such, comparing observed patterns of transfers to the predictions of each division rule allows evaluating which division rule examined in the theory section best matches agents' behavior. This approach is by no means a definitive test of the division rule used by agents, but it sill gives some indication as to which set of assumptions is most reasonable.

The first row of Table 2 describes the baseline condition. The baseline uses a star network has the seed be one of the spokes. It uses a monitoring technology $q = .1$. In equilibrium, the seed should accept the rent under both grand and petty corruption, and keep it to herself.

One theoretical prediction, derived from Proposition 1.2, described the consequences of better monitoring. It is encapsuled in the following hypothesis:

Hypothesis 1. Better monitoring technologies decrease the frequency of corruption, increase its scope and select on grand corruption.

I introduce the *hard* treatment to test this hypothesis. This treatment holds the network structure constant but increases the monitoring technology from $q = .1$ in the baseline to $q = .75$. The second row of Table 2 shows that in this treatment, the seed is expected to recruit more participants than in the baseline under grand corruption, and to reject the rent under petty corruption.

Another series of predictions considered the impact of adding additional ties to the organization. Propositions 1.4 and 1.5 showed that some ties make corruption less frequent, some increase corruption, and others have no effect, depending on whether they enable additional monitoring, or reaching better accomplices. I call the ties that reduce corruption *exposing* ties, and the ones that have no effect *irrelevant* ties, and test the following two hypotheses[20]:

Hypothesis 2. Exposing ties reduce the frequency of corruption.

Hypothesis 3. Irrelevant ties have no effect on the frequency and the scope of corruption.

I introduce the *exposing tie* treatment to test Hypothesis 2. This treatment holds the monitoring technology constant but adds a tie that should decrease corruption. The third row of Table 2 shows that this treatment amends the star network used in the baseline by adding a tie between the seed and another spoke of the star, hence making the seed exposed to one more participant. The treatment should reduce the frequency of corruption compared to the baseline, since the seed should now reject petty bribes.

I test Hypothesis 3 by adding and removing, within each treatment conditions, ties that should have no effect on outcomes, represented by the sub-rows in Table 2. These ties have no effect because they do not make minimal coalitions more exposed. Indeed, they either connect players that were neither accomplices nor witnesses (e.g. in the baseline), players that were both accomplices (hard treatment under grand corruption), or a player that was already a witness to another accomplice (exposing tie condition under grand corruption). Since the goal of this comparison is to fail to reject the null of no significant differences, it is important that this test be sufficiently well-powered. As such, within each treatment condition, half the games include the irrelevant tie, and half do not.

The last theoretical prediction was that within an organizational structure, minimal coalitions are more corrupt, because they generate fewer witnesses (Proposition 1.3):

Hypothesis 4. Minimal coalitions are more corrupt.

I test this hypothesis by looking, within the exposing tie treatment under grand corruption, at whether the seed prefers recruiting the more enclaved left-hand side node over the less enclaved right-hand side node.

3.1.3. Procedures

The experiments were all conducted in Mohammedia, Morocco, with a sample of 272 subjects comprising of one quarter undergraduate students, and three quarters employees of the service industry. Table 3 shows descriptive statistics about the two samples, revealing that they are very different. As such, comparing between subject pools tests whether behavior is driven by some characteristic held only by students or employees.

[19] Section 2.3 discussed how Propositions 1.2 to 1.5 need not hold true under one-shot bargaining. It turns out they do hold true under the particular case used in the experiment. See proofs in SI, section C.1.

[20] Due to power considerations, I did not test whether some ties may increase corruption (Proposition 1.5).

Table 3
Sample descriptive statistics. Income is measured from asset ownership and ranges from 0 to 3. Risk-taking ranges from 1 (risk-averse) to 4 (risk-lover). Altruism is measured from the donation in a dictator's game. Extroversion ranges from 0 (introvert) to 5 (extrovert). Tests for differences in means use a t-test; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Variable	Sample	Students	Employees	Δ
age	24.85	20.44	26.17	5.73***
% females	0.21	0.16	0.23	0.07
% secondary education	0.75	0.95	0.69	-0.26***
income	1.65	1.83	1.59	-0.23**
% urban	0.94	1.00	0.93	-0.07***
% Arabs	0.94	0.95	0.94	-0.01
risk-taking	2.78	2.59	2.83	0.25
altruism	1.54	1.83	1.45	-0.37**
extroversion	2.96	2.69	3.04	0.36**
N	272	63	209	

During an experimental session, subjects were randomly assigned to groups of four, and played 12 repetitions of the diffusion game. At the beginning of each session, I randomly decided whether the entire session would be played under petty or grand corruption. The session was divided into three parts corresponding to each treatment, and each treatment would be repeated four times. The first and second part were the baseline and hard treatment in a random order. The exposing tie treatment was always played in the third part, because it was cognitively more taxing. Assignment to treatment was such that, within each treatment condition, each subject would get to be the seed once, and to occupy the remaining three network positions once.

In order to approximate play in a one-shot game, subjects were not informed of how many games they would play, and were not allowed to keep track of their gains. I show in Online Appendix C that there is little evidence for potential learning and pooling effects – that is, whether subjects converge to or diverge from equilibrium predictions over time (learning), and whether they tie their behavior in a repetition to behavior in another repetition (pooling). I also show that participants displayed satisfactory levels of comprehension.

I used standard experimental procedures, including monetary incentives and neutrally worded instructions. In other words, the protocol used a neutral framing that made no mention of corruption to participants, in order to control for moral considerations about corruption and social desirability bias (I discuss the implications of this design choice in the conclusion). The experiment was about 90-minutes long. Each subject took part in only one session. Total compensation, including a $5 show-up fee, averaged $7.6, which amounts to about daily minimum wage. Online Appendix C gives additional details about experimental procedures, including recruitment of subjects, training of enumerators, prompts and materials used in the experiment. This Appendix also conducts post-hoc power analysis and reports tests for potential learning and pooling effects.

3.2. Results

3.2.1. Main theoretical predictions

I first examine support for the main theoretical predictions, that are encapsuled in Hypotheses 1 to 4.

To test these hypotheses, I compare three outcomes across all treatment conditions: whether the seed accepts the rent (Fig. 7), the mean size of realized coalitions (Fig. 8), and the distribution of these coalitions (Fig. 9). Average treatment effects on acceptance behavior and on coalition size are derived using difference in means estimated using OLS: I regress the corresponding outcomes – whether the seed takes the rent, and the size of the resulting coalition – on indicator variables for each treatment, where a *treatment* is the interaction of a main *condition* (baseline, hard, exposing tie), and the scale of corruption (petty, grand). The effect of irrelevant ties on acceptance behavior is estimated within-treatment using OLS: I add to the previous specification an indicator variable for whether the network included the irrelevant tie or omitted it. I cluster errors at the group level, to account for within-group correlations.[21] When examining the size of realized coalitions and their distribution, I restrict the analysis to grand corruption, because predictions on the structure of corrupt coalitions are conditional on the seed taking the rent in equilibrium, which only happens with grand corruption.

In what follows, I discuss how the results in Figs. 7 to 9 show support for Hypotheses 1 to 3 and discuss how behavior does not differ across subject pools. I also point out an important deviation from model predictions, that will motivate the discussion in the next section.

Support for Hypothesis 1. Comparing the hard treatment to the baseline, we find that:

- *Corruption becomes less frequent under the hard treatment by selecting on grand corruption.* With the baseline under grand corruption as a reference, acceptance rates remain comparable under baseline – petty corruption (−7 p.p., p-value

[21] Models reported in Online Appendix C.3.

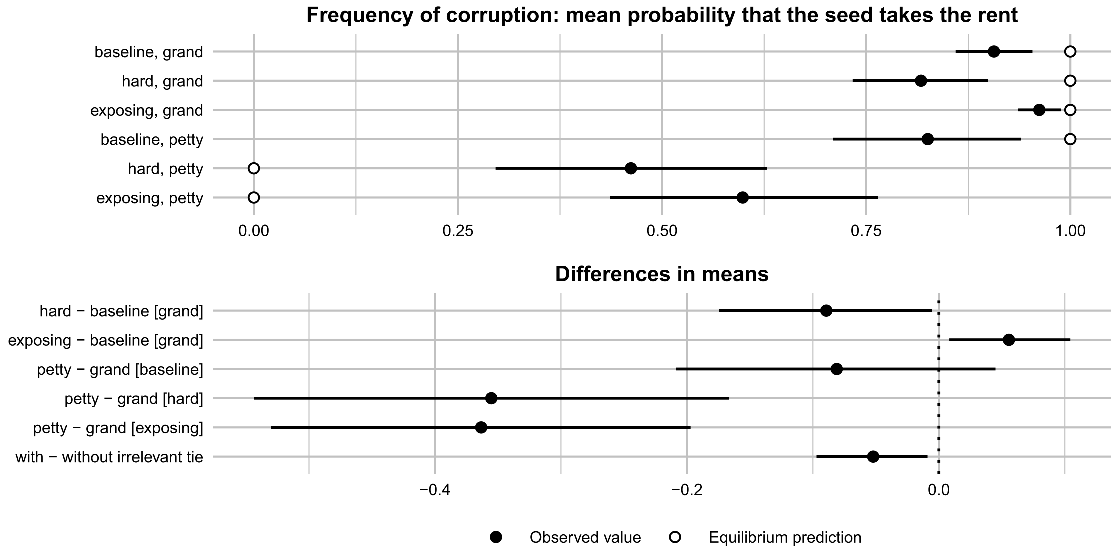


Fig. 7. Frequency of corruption. Bars are semi-parametric bootstrapped 95 percent confidence interval clustered at the group level using 10,000 replicates. Frequency is comparably high in all treatments under grand corruption. Switching to petty corruption has little effect in the baseline, but largely decreases frequency in the other treatments. Irrelevant ties have little effect.

= .21) and hard treatment – grand corruption (−5 p.p., p-value = .04). They drop significantly in hard treatment – petty corruption (−35 p.p., p-value < .001, Fig. 7).

- *The scale of corruption decreases in the hard treatment.* The average size of the coalition increases by 1.1 accomplice (p-value < .001, Fig. 8), and realized coalitions somewhat match predictions (Fig. 9).

Support for Hypothesis 2. Comparing the exposing tie treatment to the baseline, we find that *corruption becomes less frequent under the exposing tie treatment.* With the baseline under grand corruption as a reference, acceptance rates remain comparable under exposing tie treatment – grand corruption (3 p.p., p-value = .03). They drop significantly in exposing tie treatment – petty corruption (−35 p.p., p-value < .001) (Fig. 7).

Support for Hypothesis 3. Comparing the conditions that include the irrelevant tie to those that do not, we find that *irrelevant ties have no effect on the frequency and the scope of corruption*. There is no significant difference in acceptance rates (Fig. 7). Comparing the distribution of realized coalitions with and without the irrelevant tie within-treatment using Fisher exact tests shows no significant difference (Online Appendix C.3).

Support for Hypothesis 4. Looking within the exposing tie condition, we find that *minimal coalitions are more corrupt*. The seed overwhelmingly favors the equilibrium coalition (45 percent of realized coalitions) over the other, more exposed 2-people coalition (Fig. 9).

Few differences across subjects (Online Appendix C.6). Multi-level specifications with random effects at the individual and group level show that there is little group- and individual-level heterogeneity, and that group-level heterogeneity is larger than individual-level heterogeneity. Despite holding different characteristics, students and employees behave similarly. Their behavior in the baseline is similar, and they show comparable effect sizes. This makes the results more credible, suggesting that behavior is not driven by some characteristic held only by students or employees.

Comparing experimental results to model predictions. Predictions largely align with the model except for two points. First, while results relative to acceptance/rejection of the rent align with comparative statics, levels of acceptance in treatment conditions where the seed should reject the rent is relatively high (around 50%), which may reflect experimenter demand effect. Second, realized coalitions under the hard treatment appear smaller than the predicted complete coalition (Figs. 8 and 9).

3.2.2. Division rule

Having shown support for the main theoretical predictions, I now investigate the division rule used by participants. I consider the division rules for the one-shot game examined in section 2.3: *bargaining* in a lawless environment (Γ_l), and, within a contractual environment, equal-sharing and monopoly. The design makes characterizing the division rule difficult. First, because treatments were designed to yield similar outcomes irrespective of the rule, predictions are often identical across rules. Second, in the contractual environment, it is unclear what is the best response to transfers that are inconsistent with the division rule, since this environment assumes that agents make transfers consistently with the division rule. To compare division rules, I restrict the analysis to on-path behavior, and examine the seed's payoffs in treatments where the equilibrium outcome was realized. I then examine deviations from predictions under the lawless environment at all histories, since best-responses are always well-defined. Recall that under bargaining, agents have threshold strategies

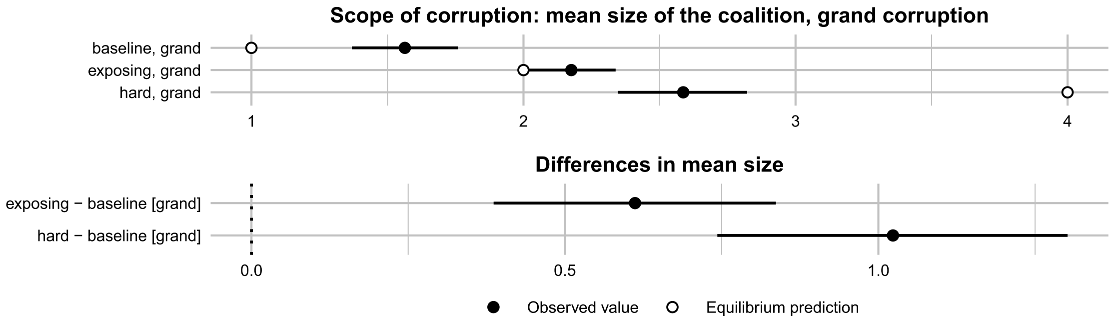


Fig. 8. Scope of corruption under grand corruption. Bars are semi-parametric bootstrapped 95 percent confidence interval clustered at the group level using 10,000 replicates. Conditional on corruption occurring, the realized coalition involves more accomplices in the hard treatment. Results are furthest away from predictions in the hard treatment.

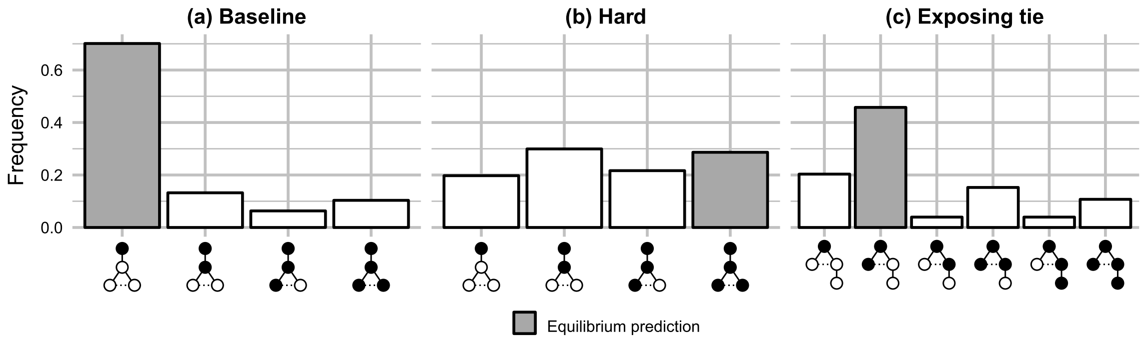


Fig. 9. Distribution of realized coalitions under grand corruption. Labels are coalitions of black nodes. Minimal coalitions are more corrupt (panel c, second vs. third bar). Realized coalitions overwhelmingly match the prediction, except in the hard treatment.

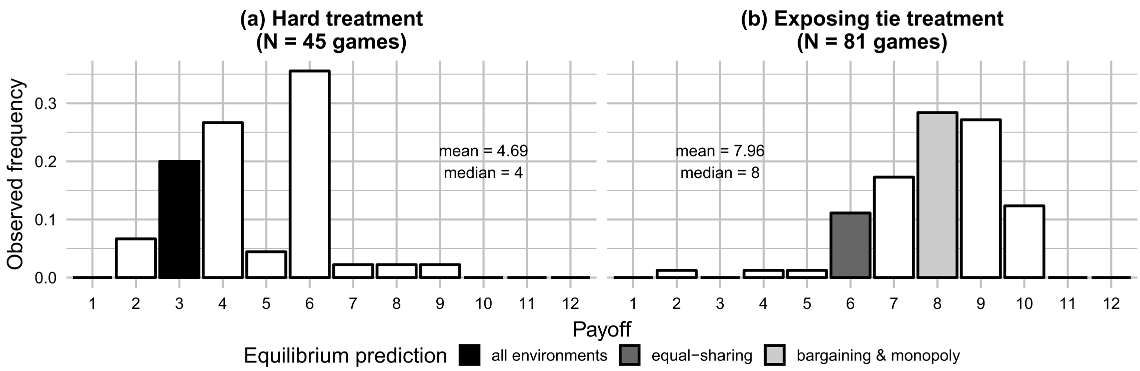


Fig. 10. Seed's share of the rent under grand corruption in equilibrium coalitions. All: much of the mass lies above the monopoly prediction. Right: The seed's payoff is closer to the bargaining/monopoly predictions.

(Proposition 2.4): recipient j should accept transfer t_{ji} from j if it is above some threshold t^*_{ji}. Having accepted the transfer, recipient i then turns into an offerer and has an optimal vector of offers t^*_i. Comparing observed transfers to those thresholds highlights offers that are too greedy ($t_{ji} - t^*_{ji} < 0$), i.e. that leave the recipient i with negative surplus and should be rejected, or offers that are too generous ($t_{ji} - t^*_{ji} \geq 0$), i.e. leave the recipient with positive surplus and should not be extended. In what follows, I introduce the results, then discuss.

Division rule. Weak evidence in favor of bargaining and monopoly against equal-sharing. Fig. 10 shows the distribution of the seed's payoff in instances where a multi-player coalition was the equilibrium outcome (hard and exposing tie treatments under grand corruption) and was realized. In the treatment that has different predictions for different division rules (panel b), the observed distributions align more closely with the prediction under bargaining and equal-sharing than with the prediction under equal-sharing.

Deviation from model prediction (1). Recipients accept offers that are too greedy. When the equilibrium coalition is realized, the seed's payoff is strictly above the monopoly allocation (implying that other players have negative surplus)

Table 4
Distribution of deviations from bargaining. Numbers denote observed frequencies. Italicized cells denote deviations from predictions under bargaining.

		did	
		share	not share
should	share	.07	*.00*
	not share	*.50*	.42

(a) Sender's decision

		did	
		accept	reject
should	accept	.24	*.01*
	reject	*.51*	.23

(b) Recipient's decision

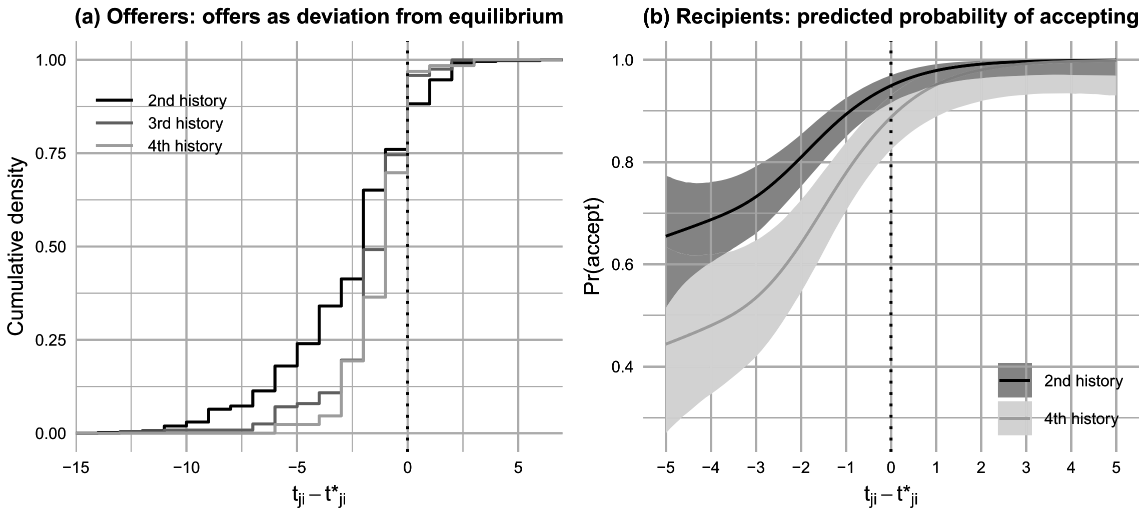


Fig. 11. Deviation of observed transfers t_{ji} from acceptance threshold t^*_{ji}. History 1 is the seed's decision, and is omitted. Right: shaded areas denote 95% confidence intervals. For ease of interpretation, predicted probabilities at the third history are omitted.

in 72 and 40 percent cases in the hard and exposing tie treatments respectively (Fig. 10). Considering deviations from bargaining, I examine binary decisions; i.e. whether a sender makes any offer, and whether a recipient accepts said offer.[22] I find that for both senders and recipients, about 50 percent decisions are errors, and that virtually all errors are false positives (Table 4): when behaving inconsistently with model predictions, senders tend to make offers when they should not (over-sharing) and consistently, recipients accept offers they should reject (over-acceptance).

Deviation from model prediction (2). Deviations attenuate in later histories. The left panel of Fig. 11 examines the distribution of $t_{ij} - t^*_{ij}$, the deviation of observed offers from the acceptance threshold. About 75 percent offers are greedy, but deviations attenuate at later histories: compared to history 2, histories 3 and 4 have much fewer offers with $t_{ij} - t^*_{ij} < 2.5$. The right panel examines the probability that these offers are accepted using generalized additive logistic regression.[23] Greedy offers have a high chance of being accepted: an offer that is greedy by 1 EU has more than 80 percent changes of being accepted. However, just like offering behavior, deviations attenuate at later histories: the probability of accepting an offer greedy by 1 EU drops by about 10 percentage points between the second and the fourth history.

Discussion. We find that recipients accept offers that are too greedy, but that these deviations from model prediction attenuate at later histories. The finding is consistent with the commonly observed fact that backward induction problems are cognitively taxing, especially at early histories (Johnson et al., 2002; Spenkuch et al., 2018). At early histories, offerers and recipients make and accept offers that only seem generous because they both fail to internalize these future transfers. That they fail to internalize these future transfers may, in turn, owe to experimenter demand effect, in the sense that subjects would want to please the experimenter and diffuse the rent in a game that is about diffusion. At later histories however, there are fewer steps to backward-induce, making the problem easier. Recipients are better able to identify greedy offers and potential tensions between maximizing one's payoff and pleasing the experimenter. Accordingly, offers adjust to get closer to the equilibrium prediction. This may explain why results are furthest away from predictions in the hard treatment: this treatment prompts for the largest equilibrium coalition, which poses a more complex induction problem to participants. This also has substantive implications for corruption under better monitoring. Better monitoring not only reduces corruption by making it more risky, but also by prompting for larger coalitions, which requires agents to solve harder backward induction problems.

[22] For acceptance, I only examine non-seed nodes. Indeed, the seed faces an exogenous offer, which is very different from subsequent endogenous offers, and her decision is examined in detail in Fig. 7 in the previous subsection.

[23] The model includes a parameter for the history, and uses thin plate regression splines. It is estimated on offers with deviation ranging from −5 to 5, to exclude outliers, and is reported in Online Appendix C.3, Table C.1.

4. Conclusion

We began by highlighting that extant approaches to corruption face difficulties answering two questions: when do large networks of corrupt agents emerge? how does organizational structure affect corruption?

This paper proposed a model and an experimental design that treat corruption as the outcome of a process of strategic diffusion on a network. This simple, easily expandable idea provides a framework to think about the relationship between how corruption is organized, and how this is affected by pre-existing organizational structures. The model provides insights that echo, reconcile, and sharpen extant findings. The lab experiment confirms most of the model's predictions in a field environment with relative ecological validity, and shows divergences that have substantive implications. I now discuss the implications of these findings in light of the literature, and highlight how further research may address limitations of the approach.

The first finding is that corruption occurs in minimal coalitions (Proposition 1.3). The concept shifts the unit of analysis from the individual to the coalition, and distinguishes between two types of ties: ties among accomplices, and ties between accomplices and witnesses. The concept itself is largely agnostic about ties among accomplices, but states that coalitions should minimize exposure to witnesses. This, in turn, may reconcile mixed findings about the structure of criminal networks: Aven (2015) and Morselli et al. (2007) show that criminal networks are sparser than comparable non-criminal networks, but there is also evidence that better connected individuals are more corrupt (Nyblade and Reed, 2012; Khanna et al., 2015). Collectively, minimal coalitions are sparsely connected to the rest of the organization, but both sparse or dense networks may emerge among accomplices.

Minimal coalitions also nuance an old insight from the principal-agent literature – that flatter organizations limit corruption by making the actions of agents more observable to the principal (McAfee and McMillan, 1995; Melumad et al., 1995). Allowing to move beyond contrasting perfect hierarchies to perfectly flat organizations, the proposed approach shows that while the insight holds true in the aggregate, there still is variation among equally hierarchical organizations (see computational results in section 1.3). Structural details matter: minimal coalitions may appear in relatively flat organizations and reciprocally, relatively hierarchical organizations may comport few minimal coalitions, depending on the exact layout of communication and monitoring ties (Propositions 1.4 and 1.5).

Findings on monitoring technologies (Proposition 1.2) may explain why corruption persists and selects on grand corruption in developed countries, although it is less frequent than in developing countries (Kaufmann, 2004). Because detection is more likely under good monitoring, recruiting accomplices are more desirable. Yet, only grand corruption is profitable enough to afford their additional protection, leading to the disappearance of petty corruption. A comparison of corruption cases between the US than in India provides tentative evidence that corruption indeed involves more accomplices in the US, a country that presumably features better monitoring than India (Online Appendix D).

Finally, the results suggest a range of policy avenues that may tackle corruption. Most novel is the fact that changing the organizational structure may substitute for better enforcement, but may also backfire. This is concerning because organizational responses to corruption, ranging from fairly standardized practices such as staff rotation (Abbink, 2004) and competition between agencies (Amir and Burr, 2015) to highly specific organizational redesigns (e.g. Bennet, 2012; Friedman, 2012) are very common, but have not been subjected to careful evaluation. Results suggest that one should consider whether the proposed policy will undermine existing minimal coalitions without creating new ones. Another policy that may reduce corruption is improving monitoring. Immediately, better monitoring increases the risk of sanction. However, results also highlight an indirect channel through which monitoring reduces corruption. Better monitoring prompts for larger coalitions, which are problematic for two reasons. First, under incomplete information (Section 2.2), larger coalitions have a higher probability of containing at least one "traitor" who would cooperate with law enforcement and denounce her accomplices. Second, the lab shows that forming large coalitions poses a challenging backward induction problem. This behavioral trait makes corruption all the more unlikely under good monitoring. Finally, results on alternative division rules (Section 2.3) show that undermining agents' ability to enforce informal contracts – e.g. by shortening their time horizon (Proposition 2.7) through, say, staff rotation – will introduce inefficiencies that will reduce corruption.

Yet, coming to definitive policy recommendations requires further research. Although this paper opens up new avenues for thinking of corruption in organizations, the model and experiment incorporated some stark design choices that could usefully be relaxed.

The model makes several simplifying assumptions: it analyzes a one-shot game of complete information, where agents are largely homogeneous and report corruption mechanistically. While some of these assumptions have been relaxed in extensions – specifically, the assumptions of complete information and the one-shot game, – future research could usefully explore other avenues. In particular, considering the impact of strong ties (e.g., coethnicity) and of strategic reporting of corruption by witnesses may yield interesting insights on corruption in developing countries. In this environment, strong ties could hold better information about each other, be better able to cooperate, or less likely to report corruption. This would, in turn, introduce a new tradeoff in coalition formation. Accomplices might favor including more exposed strong ties because they are more efficient, or are known to be corruptible. Conversely, they might prefer including weak ties, because strong ties would be less likely to report corruption.

Finally, the minimal experimental design proposed in this paper could be extended to evaluate the robustness of the findings to several well-known behavioral traits, and test the model extensions outlined above. The present design made three strong decisions: corruption did not have negative externalities, the design used a neutral framing, and reporting corruption was mechanistic. Future experiments could usefully use this design as a baseline and evaluate these features through additional treatments.

Available evidence suggests that a loaded framing and negative externalities should have little impact. Regarding negative externalities, section 2.3 shows that subjects engaged in greedy bargaining. Showing little altruism to fellow accomplices, it seems unlikely that they would be more altruistic towards society broadly defined. Considering the impact of a loaded framing, Abbink and Hennig-Schmidt (2006) find no framing effects. In a bribery game, Barr and Serra (2009) and Lambsdorff and Frank (2010) do find framing effects, but only on the citizen side. Since this paper focuses on bureaucrats, introducing a loaded framing should have little effect.

In the spirit of the proposed theoretical extension with strong ties, another interesting empirical avenue would be to introduce strategic reporting. The present design could prove particularly useful, because bringing real friendship ties into the lab would pair well with face-to-face interactions.

Appendix A. Proofs of section 1

Proof of Lemma 1.1. Suppose $u_s(c^*, g, q) < 0$. Then, no coalition gives s a positive payoff. She rejects the rent.

Suppose $u_s(c^*, g, q) \geq 0$. Since $c^* \in \mathcal{C}(g, s)$, there is at least one strategy profile that has c^* as an outcome. Since all accomplices have the same payoff function, c^* maximizes the utility of any accomplice i in c^*. No accomplice has an incentive to deviate from the profile, since it yields their highest possible payoff. Because $u_s(c^*, g, q) \geq 0$, s accepts the rent.

Showing that profiles that have as outcomes coalitions that do not belong to $\mathcal{C}^*(g, s, q)$ cannot be sustained in equilibrium is straightforward. Consider a strategy profile that such a coalition as an outcome to one that has c^* as an outcome. At the first history where the two profiles diverge, the player that moves at this history has an incentive to deviate to the profile that has c^* as an outcome. As such, $\hat{\epsilon}_s(g, q) = v(a_{c^*}, w_{c^*g}, q) \in (0, 1)$. □

Proof of Proposition 1.1. Consider two essentially different coalitions c_1, c_2 on some graphs g_1 and g_2 respectively, with probability of success p_1 and p_2 respectively, and suppose without loss of generality that $a_{c_2} \geq a_{c_1}$. Let $u_1 = u_s(c_1, q)$ and $u_2 = u_s(c_2, q)$ be the seed's utility from these coalitions, and let $U = \{(\epsilon_s, q) : u_1 = u_2\} \subset (0, 1)^2$. I show that U has measure 0.

We have $u_2 - u_1 = \frac{p_2}{a_{c_2}} - \frac{p_1}{a_{c_1}}$. Suppose U is non-empty and consider some point $(\epsilon_s, q) \in U$. The directional derivative of $u_2 - u_1$ at this point writes:

$$\nabla_x u_2 - u_1 \equiv \frac{\partial u_2 - u_1}{\partial \epsilon_s} x_{\epsilon_s} + \frac{\partial u_2 - u_1}{\partial q} x_q = \left(\frac{\partial p_2}{\partial q} \Big/ a_{c_2} - \frac{\partial p_1}{\partial q} \Big/ a_{c_1} \right) x_q \tag{A1}$$

where $x \equiv (x_{\epsilon_s}, x_q)$ is a unit-length vector. If the equation $\nabla_x u_2 - u_1 = 0$ has a finite number of solutions in x, then U has measure 0. Assumption 1.1 implies that $\frac{\partial p_2}{\partial q}/a_{c_2} - \frac{\partial p_1}{\partial q}/a_{c_1} \neq 0$, so the only solutions are $(1, 0)$ and $(-1, 0)$.

Since the space for which one is indifferent between any two essentially different coalitions has measure 0, the space for which one is indifferent between any two essentially different equilibrium coalitions also has measure 0. □

Proof of Proposition 1.2. Let's first show that $q_1 < q_2 \Rightarrow \hat{\epsilon}_s(g, q_1) \geq \hat{\epsilon}_s(g, q_2)$. From Lemma 1.1, if $c^* \in \mathcal{C}^*(g, s, q)$, then $\hat{\epsilon}_s(g, q) = u_s(c^*, g, q) + \epsilon$. We have $u_s(c_1^*, g, q_1) \geq u_s(c_2^*, g, q_1)$. Since for a given coalition, u is decreasing in q, we have $u_s(c_2^*, g, q_1) \geq u_s(c_2^*, g, q_2)$. This implies $u_s(c_1^*, g, q_1) \geq u_s(c_2^*, g, q_2)$ which, using Lemma 1.1, implies $\hat{\epsilon}_s(g, q_1) \geq \hat{\epsilon}_s(g, q_2)$.

I now show that $q_1 < q_2 \Rightarrow a_{c_1^*} \leq a_{c_2^*}$. Let c_1^* be the largest coalition in $\mathcal{C}^*(g, s, q_1)$, and c_2^* the smallest in $\mathcal{C}^*(g, s, q_2)$ with sizes $a_{c_1^*}$ and $a_{c_2^*}$. To prove the claim, itsuffices to show that $a_{c_1^*} \leq a_{c_2^*}$. Suppose not. Because $c_1^* \in \mathcal{C}^*(g, s, q_1)$ and $c_2^* \in \mathcal{C}^*(g, s, q_2)$, we have $u_s(c_2^*, g, q_1) - u_s(c_1^*, g, q_1) \leq 0$ and $u_s(c_2^*, q_2) - u_s(c_1^*, q_2) \geq 0$. Since $u_s(c_2^*, q) - u_s(c_1^*, q)$ is continuous in q, there must be some $q \in [q_1, q_2]$ such that $u_s(c_2^*, q) = u_s(c_1^*, q)$. Because $a_{c_1^*} > a_{c_2^*}$, Assumption 1.2 implies that $\frac{\partial}{\partial q}\left[u_s(c_2^*, q) - u_s(c_1^*, q)\right] < 0$. Since $u_s(c_2^*, g, q_1) - u_s(c_1^*, g, q_1) \leq 0$, then $u_s(c_2^*, q_2) - u_s(c_1^*, q_2) < 0$, a contradiction. □

Proof of Proposition 1.3. Suppose not. That is, suppose $c \in \mathcal{C}(g, s, q)$ and $c \notin \mathcal{M}(g, s)$. If $c \notin \mathcal{M}(g, s)$, then $c \notin \mathcal{M}_{a_c}(g, s)$. Since $\mathcal{M}_{a_c}(g, s)$ is non-empty, there is $c' \in \mathcal{M}_{a_c}(g, s)$ such that $w_{c'} < w_c$. As such, $u(c', q) > u(c, q)$. So $c \notin \mathcal{C}(g, s, q)$, a contradiction. □

Proof of Corollary 1.3.1. Suppose not. That is, suppose that node $i \in \mathcal{N}$ is a member of a coalition c^* that is an equilibrium outcome for some $s \in \mathcal{N}$, but $i \notin c$ for any $s' \in \mathcal{N}$, any $c \in \mathcal{M}(g, s)$. Proposition 1.3 implies that $c^* \in \mathcal{M}(g, s)$, a contradiction. □

Lemma A1 *(Old coalitions are weakly dominated). For any $s \in \mathcal{N}$, we have $\mathcal{C}(g, s) = \mathcal{C}(g', s)$ if $g' = g + i \to j$ and $\mathcal{C}(g, s) \subseteq \mathcal{C}(g', s)$ if $g' = g + ij$. Any $c \in \mathcal{C}(g, s)$ satisfies:*

$$w_{cg'} = \begin{cases} w_{cg} + 1, & \text{if } g' = g + i \to j \text{ and } j \in c \text{ and } i \notin c \cup \mathcal{W}_{cg} \\ w_{cg} & \text{otherwise.} \end{cases} \tag{A2}$$

Proof. The proof is immediate. □

Proof of Proposition 1.4. By Lemma A1, we have $\mathcal{C}(g', s) = \mathcal{C}(g, s)$ and for any $c \in \mathcal{C}(g, s)$, $w_{cg} \le w_{cg'} \le w_{cg} + 1$. This implies $v(a_c, w_{cg'}, q) \le v(a_c, w_{cg}, q)$. So $\hat{\epsilon}_s(g', q) \le \hat{\epsilon}_s(g, q)$. Let's show the second part of the proposition. Proposition 1.3 implies that for any $q \in (0, 1)$, there is $c^* \in \mathcal{M}(g, s)$ that is realized in equilibrium on g. I show the contrapositive. Suppose that for some $s \in \mathcal{N}$, and for all $a \in \{1, ..., \mathcal{N}\}$, there is $c \in M_a(g, s)$ such that $j \notin c$ or $i \in c \cup \mathcal{W}_{cg}$. Then by Lemma A1, c on g' is essentially equal to c^* on g. As such, $v(a_c, w_{cg'}, q) = v(a_c, w_{c^*g}, q)$, which implies $\hat{\epsilon}_s(g', q) = \hat{\epsilon}_s(g, q)$ for any $q \in (0, 1)$. □

Proof of Proposition 1.5. By Lemma A1, we have that for any $c \in \mathcal{C}(g, s)$, $v(a_c, w_{cg}, q) = v(a_c, w_{cg'}, q)$. As such, for any $s \in \mathcal{N}$, it cannot be that $\hat{\epsilon}_s(g', q) < \hat{\epsilon}_s(g, q)$. This implies that for all $s \in \mathcal{N}$, we have $\hat{\epsilon}_s(g', q) \ge \hat{\epsilon}_s(g, q)$. □

Proof of Proposition 1.6. We first show that $\mathcal{C}(g, s) = \mathcal{C}(g', s)$. Suppose not; that is, suppose that there is $c' \in \mathcal{C}(g', s)$ such that there is $k \in c'$ such that all paths between k and s such that all nodes on this path are in c' go through the tie ij. Suppose that this path is $k, \dots, i, j, \dots, s$. Since i is a neighbor of s, an alternative path is $k, \dots, i, s$. As such, $c' \in \mathcal{C}(g, s)$, a contradiction.

We then show that for any $c \in \mathcal{C}(g, s)$, we have $w_{cg} = w_{cg'}$. Suppose that there is c such that $w_{cg'} \ne w_{cg}$. Note that by Lemma A1, it must be that $w_{cg'} = w_{cg} + 1$. Additionally, Lemma A1 implies that $i \in c$, $j \notin \mathcal{W}_{cg}$, $j \notin c$. Suppose without loss of generality that $i \in c$, $j \notin c$. Then, since j is a neighbor of s, it must be that $j \in \mathcal{W}_{cg}$, a contradiction.

Since $\mathcal{C}(g, s) = \mathcal{C}(g', s)$ and for any $c \in \mathcal{C}(g, s)$, we have $w_{cg} = w_{cg'}$, it must be that $\hat{\epsilon}(s, g) = \hat{\epsilon}(s, g')$. As such, $AUC_{sg} = AUC_{sg'}$. □

Appendix B. Proofs and additional results of section 2

B.1. Proofs and additional results of section 2.1

Proof of Lemma 2.1. Let's first show that $\arg\max_{c \in \mathcal{C}(g,s)} u_s(c, q, \epsilon) = \arg\max_{c \in \mathcal{C}(g,s)} u_s(c, q, \epsilon')$. Consider $c, c' \in \mathcal{C}(g, s)$ such that $u_s(c, q, \epsilon) \le u_s(c', q, \epsilon)$. This implies $\frac{\rho(\epsilon) p(a_c, w_{cg}, q)}{a_c} \le \frac{\rho(\epsilon) p(a_{c'}, w_{c'g}, q)}{a_{c'}}$. As such, $u_s(c, q, \epsilon') \le u_s(c', q, \epsilon')$, proving the point.

Let's show the rest of the lemma. Note that $\frac{\partial u_s}{\partial \epsilon} = \frac{\rho'(\epsilon) p(a_c, w_{cg}, q)}{a_c} - 1$. So $\frac{\partial u_s}{\partial \epsilon} \ge 0 \iff \rho'(\epsilon) \ge \frac{a_c}{p(a_c, w_{cg}, q)}$.

Let's show the rest of the proposition. Suppose that for a given q, there is $\hat{\epsilon}_s > 0$ such that $u_s(c^*, q, \hat{\epsilon}_s) = 0$. Note that $\frac{\partial u_s}{\partial \epsilon} = \frac{\rho'(\epsilon) p(a_c, w_{cg}, q)}{a_c} - 1$. So if $\rho'(\epsilon) \le \frac{a_c}{p(a_c, w_{cg}, q)}$ for any $c \in \mathcal{C}(g, s), q \in (0, 1), \epsilon \in (0, 1)$, s rejects the rent for $\epsilon > \hat{\epsilon}_s$. Conversely, if $\rho'(\epsilon) > \frac{a_c}{p(a_c, w_{cg}, q)}$ for any $c \in \mathcal{C}(g, s), q \in (0, 1), \epsilon \in (0, 1)$, s rejects the rent for $\epsilon < \hat{\epsilon}_s$. If $u_s(c^*, q, \hat{\epsilon}_s) > (<) 0$ for any $\epsilon \in (0, 1)$ then the seed always accepts (rejects) the rent, so define some $\hat{\epsilon}_s \in \{0, 1\}$.

Whenever $u_s(c^*, q, \hat{\epsilon}) \ge 0$, we show as in Lemma 1.1 that s accepts the rent and c^* is an equilibrium outcome. □

Proof of Proposition 1.1. The proposition proves as in Appendix A, with the exception that $u_2 - u_1 = \rho(\epsilon) \frac{p_2}{a_{c_2}} - \frac{p_1}{a_{c_1}}$. However, the directional derivative at this point writes

$$\nabla_x u_2 - u_1 = \rho'(\epsilon) \left(\frac{p_2}{a_{c_2}} - \frac{p_1}{a_{c_1}} \right) x_\epsilon + \left(\frac{\partial p_2}{\partial q} \Big/ a_{c_2} - \frac{\partial p_1}{\partial q} \Big/ a_{c_1} \right) x_q$$

Since the directional derivative is evaluated at U where $\frac{p_2}{a_{c_2}} - \frac{p_1}{a_{c_1}} = 0$, the directional derivative reduces to equation (A1). The rest of the proof proceeds as in the proof in Appendix A. □

Proof of Proposition 2.1. From Lemma 2.1, $\hat{\epsilon}(g, q)$ is one of the bounds of the interval in ϵ such that s accepts the rent; that is, such that $u_s(c^*, q, \epsilon) \ge 0$ for some $c^* \in \arg\max_{c \in \mathcal{C}(g,s)} u_s(c, q, \epsilon)$. Pick $q_1 < q_2$, and their associated equilibrium coalitions, c_1, c_2. Coalition c_1 satisfies $u_s(c_1, q_1, \epsilon) \ge u_s(c_2, q_1, \epsilon)$. Since for a given coalition, u_s is decreasing in q, we have $u_s(c_2, q_1, \epsilon) \ge u_s(c_2, q_2, \epsilon)$. This implies $u_s(c_1, q_1, \epsilon) \ge u_s(c_2, q_2, \epsilon)$. As such, the interval such that $u_s(c_1, q_1, \epsilon) \ge 0$ has a weakly greater range than the interval such that $u_s(c_2, q_2, \epsilon) \ge 0$. That is, $\hat{\epsilon}(g, q_1) \ge (\le) \hat{\epsilon}(g, q_2)$ if $\rho'(\epsilon) \le (>) \frac{a_c}{p(c_a, w_{cg}, q)}$. The rest of the proposition proves as in Proposition 1.2. □

Proposition 1.3 and Corollary 1.3.1 remain unchanged and prove as in the simple model.

Proposition 1.4 becomes:

Proposition B1. *Suppose $g' = g + i \to j$. If $\rho'(\epsilon) \le (>) \frac{a_c}{p(c, g, q, \epsilon)}$ for any $s \in \mathcal{N}, c \in \mathcal{C}(g, s), q \in (0, 1), \epsilon \in (0, 1)$, then $\hat{\epsilon}_s(g', q) \le (\ge) \hat{\epsilon}_s(g, q)$ for all $s \in \mathcal{N}$. Furthermore, if there is $s \in \mathcal{N}$ and $q \in (0, 1)$ such that $\hat{\epsilon}_s(g', q) < (>) \hat{\epsilon}_s(g, q)$, then there is $a \in \{1, ..., |\mathcal{N}|\}$ such that for all minimal coalition $c \in \mathcal{M}_a(g, s)$, $j \in c$ and $i \notin c \cup \mathcal{W}_{cg}$.*

Proof of Proposition B1. Let $u_i(c, g, q, \epsilon)$ be the payoff to i from coalition c *on graph* g with monitoring q and scale ϵ. By Lemma A1, we have $\mathcal{C}(g', s) = \mathcal{C}(g, s)$ and for any $c \in \mathcal{C}(g, s)$, $w_{cg} \leq w_{cg'} \leq w_{cg} + 1$. This implies $u_s(c, g', q, \epsilon) \leq u_s(c, g, q, \epsilon)$. Using Lemma 2.1 gives that if $\rho'(\epsilon) \leq (>) \frac{a_c}{p(c,g,q)}$ for any $c \in \mathcal{C}(g, s), q \in (0, 1), \epsilon \in (0, 1)$, then $\hat{\epsilon}_s(g', q) \leq (\geq) \hat{\epsilon}_s(g, q)$. We prove the rest of the proposition as in the original proof. □

Proposition 1.5 becomes:

Proposition B2. *Suppose* $g' = g + ij$. *If* $\rho'(\epsilon) \leq (>) \frac{a_c}{p(c,g,q)}$ *for any* $s \in \mathcal{N}, c \in \mathcal{C}(g, s), q \in (0, 1), \epsilon \in (0, 1)$, *then* $\hat{\epsilon}_s(g', q) \geq (\leq) \hat{\epsilon}_s(g, q)$ *for all* $s \in \mathcal{N}$.

Proof of Proposition B2. Let $u_i(c, g, q, \epsilon)$ be the payoff to i from coalition c *on graph* g with monitoring q and scale ϵ. By Lemma A1 we have that for any $c \in \mathcal{C}(g, s)$, $u(c, g, q, \epsilon) = u(c, g', q, \epsilon)$. Using Lemma 2.1, this implies that $\hat{\epsilon}(g', q) \geq (\leq) \hat{\epsilon}(g, q)$ if $\rho'(\epsilon) \leq (>) \frac{a_c}{p(c,g,q)}$ for any $c \in \mathcal{C}(g, s), q \in (0, 1), \epsilon \in (0, 1)$. □

B.2. Proofs and additional results of section 2.2

Proof of Proposition 2.2. Recall that the highest payoff from not cooperating is $u(c, g, q) = \frac{p(a_c, w_{cg}, q)}{a_c}$ in the case in which none of the remaining $a_c - 1$ agents cooperate. If $\tau_i = H$, Assumption 2.1 implies that i always prefers cooperating, since $b - \kappa_H > u(c, g)$. As such, if $\tau_i = L$ and $\Pr_i(\tau_j = L) = 0$ for some $j \neq i \in c$, the payoff from not cooperating is $-\epsilon < b - \kappa_L$, the payoff from cooperating. Suppose now that $\Pr_i(\tau_j = L) > 0$ for all $j \neq i \in c$. Suppose that agents of type L cooperate, and that l_{-i} members of c different from i have not revealed their type at that information set h – i.e. suppose that $|\{j : \Pr_i(\tau_j = L|h) = r\}| = l_{-i}$. i does not have an incentive to deviate if and only if $r^{l}_{-i} u(c, g) \geq b - \kappa_L$, which is true by Assumption 2.1, since $l_{-i} \leq a_c - 1$. □

Proof of Lemma 2.2. We first prove a series of useful lemmas.

Lemma B2. *Consider a non-terminal information set* h *and a terminal information set* h' *on the path of play from* h*, with respective coalitions* c_h *and* c'_h*. Suppose that* h *and* h' *are such that for any* $i \in \mathcal{N}$, $j \neq i$*, we have* $\Pr_i(\tau_j = L|h) > 0$, $\Pr_i(\tau_j = L|h') > 0$ *if* $j \in c_h$, $c_{h'}$ *respectively. Suppose furthermore* $\Pr_i(\tau_j = L|h) = r$ *for any* $j \notin c_h$*, and that* $b - \kappa_H > \epsilon$*. If* σ *is a MPE, then it must be that* $\Pr_i(\tau_j = L) = r$ *for any* $i \in \mathcal{N}$, $j \neq i \notin c_{h'}$*.*

Proof of Lemma B2. Suppose $\Pr_i(\tau_j = L) = 0$. Then there is an information set on the path from h to h' where an offer is extended to j such that j rejects if $\tau_j = H$ and j accepts otherwise. At the last such information set, j has a profitable deviation to accepting when $\tau_j = H$, for her payoff is $b - \kappa_H - \epsilon > 0$.

Suppose instead that $\Pr_i(\tau_j = L) = 1$. Then there is an information set on the path from h to h' where an offer is extended to j such that j rejects if $\tau_j = L$ and j accepts otherwise. Consider the last such information set, and suppose the offer is extended by i. Suppose that under σ, i makes an offer to j and the vector o_h of offers. Suppose furthermore that her expected payoff from this action is u in the state where $\tau_j = L$. Suppose that her payoff from only extending o_h offers is u'. Note furthermore that since this is the last offer that can be extended to j under σ, and since j rejects such offer under σ, j is not payoff-relevant after h. As such, and since σ is a MPE, agents that move after h must take the same actions irrespective of whether i makes an offer to j and $\tau_j = L$ or whether i does not make an offer to j. So $u = u'$. Therefore, if $\tau_i = L$, i's expected payoff from extending an offer to j is $ru + (1 - r)(b - \kappa_L) - \epsilon < u - \epsilon$ by Assumption 2.1. So i has a profitable deviation in not extending an offer to j. Given that at h', $\Pr_k(\tau_i = L) > 0$ for any $k \neq i$, it must be that if $\tau_i = H$, she is pooling with low types. As such, she takes the same action as when $\tau_i = L$. □

Lemma B3. *Suppose agent* i *with* $\tau_i = L$ *moves at information set* h *with action set* $\mathcal{A}_h$*, associated coalition* c_h*, and such that* $\Pr_j(\tau_k = L|h) > 0$ *for any* $j \in \mathcal{N}, k \neq j \in c_h$*. Suppose furthermore that* $b - \kappa_H > \epsilon$*. If* σ *is a MPE, then the expected payoff associated with any action* $m \in \mathcal{A}_h$ *in which* i *accepts the offer is*

$$\mathbb{E}_i[u_i(m)|h, \sigma] = r^{a_{c_m} - k_{ic_m h} - 1} u(c_m) + (1 - r^{a_{c_m} - k_{ic_m h} - l_{im} - 1})\overline{u} - \epsilon,$$

with c_m *the outcome coalition of the path of play associated with action* m *in the state where* $\tau_k = L$ *for any* $k \in \mathcal{N}$, $k_{ic_m h} \in \{0, \ldots, a_{c_h} - 1\}$ *the number of agents* $k \neq i \in c_m$ *such that* $\Pr_i(\tau_k = L|h) = 1$, $l_{im\sigma} \in \{0, \ldots, a_{c_m} - k_{ih} - 1\}$ *the number of agents* $k \neq i \in c_m$ *that do not reveal their type at that terminal history under* σ*.*

Proof of Lemma B3. Let h_m be the history corresponding to the path of play associated with action m in the state in which $\tau_k = L$ for any $k \in \mathcal{N}$, and let H_m be the information set to which h_m pertains.

Lemma B2 implies that the terminal history associated with action m in any state such that $\Pr_j(\tau_k = L) > 0$ for any $j \in \mathcal{N}$, $k \neq j \in c_m$ also belongs to H_m. This information set is reached if all $a_{c_m} - k_{ic_m h} - l_{im} - 1$ agents that will reveal their type between h and H_m under σ are of type L. This occurs with probability $r^{a_{c_m}-k_{ic_m h}-l_{im}-1}$. Furthermore, conditional on reaching H_m, i pockets the benefit u_{c_m} if the remaining l_{im} agents turn out to be of type L. This occurs with probability $r^{l_{im}}$. So overall, i pockets the benefit u_{c_m} with probability $r^{a_{c_m}-k_{ic_m h}-l_{im}-1} r^{l_{im}} = r^{a_{c_m}-k_{ic_m h}-1}$.

Additionally, if any of the $a_{c_m} - k_{ic_m h} - l_{im\sigma} - 1$ agents that are yet to reveal their type at H_m under σ is of type H, then i pockets $\overline{u}$. This event occurs with probability $1 - r^{a_{c_m}-k_{ic_m h}-l_{im}-1}$. □

Lemma B4. *Let*

$$u(c|k,l) \equiv r^{a_c-k-1} u_c + (1 - r^{a_c-k-l-1})\overline{u},$$

with $k \in \{0, \dots, a_c - 1\}$, $l \in \{0, \dots, a_c - k - 1\}$, $u_c = \frac{p(a_c, w_{cg}, q)}{a_c}$, *and* $\overline{u} = b - \kappa_L$ *We have that if* $u(c_1|k,l_1) \geq u(c_2|k,l_2)$, *then* $u(c_1|k',l_1) \geq u(c_2|k',l_2)$ *for any* $k' \geq k$.

Proof. Note that $u(c_1|k,l_1) \geq u(c_2|k,l_2)$ if and only if

$$r^{a_{c_2}-k-1}\left[r^{a_{c_1}-a_{c_2}} u_{c_1} - u_{c_2} + r^{-l_2}(1 - r^{a_{c_1}-l_1-(a_{c_2}-l_2)})\overline{u}\right] \geq 0$$

Therefore,

$$r^{a_{c_2}-k'-1}\left[r^{a_{c_1}-a_{c_2}} u_{c_1} - u_{c_2} + r^{-l_2}(1 - r^{a_{c_1}-l_1-(a_{c_2}-l_2)})\overline{u}\right] \geq 0 \iff u(c_1|k',l_1) \geq u(c_2|k',l_2) \quad \square$$

Suppose that $b - \kappa_H > \epsilon$. If σ is a MPE, then by Lemma B3, s's action m^* at information set h_0 solves

$$\max_{m \in \mathcal{A}_{h_0}} \mathbb{E}_s[u_s(m)|h_0, \sigma] = r^{a_{c_m}-1} u(c_m) + (1 - r^{a_{c_m}-l_{sm}-1})\overline{u} - \epsilon,$$

If $\mathbb{E}_s[u_s(m^*)|h_0,\sigma] \equiv \hat{\epsilon}_s \in (0,1) > \epsilon$, then s rejects the rent. Otherwise, she accepts it and coalition c_{m^*} is realized in any state where $\tau_i = L$ for any $i \in c_{m^*}$ such that i reveals her type under σ on the path of play. Note that if s accepts the rent, then it must be that $b - \kappa_H > 0$.

Let's show that c_{m^*} solves $\max_{c \in \mathcal{C}(g,s)} \mathbb{E}_s[u_s(c,g,q)|h_0,\sigma]$. Suppose not; that is supposed that there is c such that $\mathbb{E}_s[u_s(c,g,q)|h_0,\sigma] > \mathbb{E}_s[u_s(c_{m^*},g,q)|h_0,\sigma]$. Then this implies that there is an information set h on the path of play from h_0 such that player i with $\tau_i = L$ moves, $\Pr_j(\tau_k = L|h) > 0$ for any $j \in \mathcal{N}$, $k \neq j \in c_h$, and

$$\mathbb{E}_i[u_i(c^*_m, g, q)|h,\sigma] \geq \mathbb{E}_i[u_i(c,g,q)|h,\sigma]$$

Suppose for now that $l_{ic\sigma} = l_{sc\sigma}$ and $l_{ic_{m^*}\sigma} = l_{sc^*_m\sigma}$. We have $k_{sc_{m^*}h_0} = k_{sch_0} = 0 \leq k_{ic_{m^*}h}$. Furthermore, Lemma B2 implies that $k_{ic_{m^*}h} = k_{ich}$. As such, Lemma B4 implies that $\mathbb{E}_i[u_i(c^*_m,g,q)|h,\sigma] < \mathbb{E}_i[u_i(c,g,q)|h,\sigma]$.

Suppose now that $l_{ic\sigma} > l_{sc\sigma}$. This implies that s does not reveal her type in coalition c under σ. In other words, $l_{ic\sigma} = l_{sc\sigma} + 1$. Yet, since s moves before i, if she does not reveal her type in coalition c, she must also not reveal it in coalition c_{m^*}. As such, $l_{ic_{m^*}\sigma} = l_{sc_{m^*}\sigma} + 1$.

Note that $\mathbb{E}_s[u_s(c,g,q)|h_0,\sigma] > \mathbb{E}_s[u_s(c_{m^*},g,q)|h_0,\sigma]$ if and only if

$$r^{a_{c_{m^*}}-1}[u(c)^{a_c-a_{c_{m^*}}} u(c) - u(c_{m^*}) + r^{-l_{sc_{m^*}\sigma}}(1 - r^{a_c-a_{c_{m^*}}-(l_{sc\sigma}-l_{sc_{m^*}\sigma})})\overline{u}] > 0 \tag{B1}$$

With $l_{ic_{m^*}\sigma} = l_{sc_{m^*}\sigma} + 1$ and $l_{ic\sigma} = l_{sc\sigma} + 1$, we have

$$r^{a_{c_{m^*}}-1}[u(c,g,q)^{a_c-a_{c_{m^*}}} u(c,g,q) - u(c_{m^*},g,q) + r^{-(l_{sc_{m^*}\sigma}+1)}(1 - r^{a_c-a_{c_{m^*}}-(l_{ic\sigma}-l_{ic_{m^*}\sigma})})\overline{u}]$$

Since $r^{-(l_{sc_{m^*}\sigma}+1)} > r^{-(l_{sc_{m^*}})\sigma}$, equation (B1) implies that

$$r^{a_{c_{m^*}}-1}[u(c,g,q)^{a_c-a_{c_{m^*}}} u(c,g,q) - u(c_{m^*},g,q) + r^{-(l_{sc_{m^*}\sigma}+1)}(1 - r^{a_c-a_{c_{m^*}}-(l_{ic\sigma}-l_{ic_{m^*}\sigma})})\overline{u}] > 0$$

Using Lemma B4 and the fact that $k_{sc_{m^*}h_0} = k_{sch_0} = 0 \leq k_{ic_{m^*}h} = k_{ich}$ on this expression, we get that $\mathbb{E}_i[u_i(c^*_m,g,q)|h,\sigma] < \mathbb{E}_i[u_i(c,g,q)|h,\sigma]$.

Suppose finally that $l_{ic\sigma} < l_{sc\sigma}$. By the same reasoning as in the previous case, it must be that in both c and c_{m^*}, s reveals her type while i does not. As such, $l_{ic_{m^*}\sigma} = l_{sc_{m^*}\sigma} - 1$ and $l_{ic\sigma} = l_{sc\sigma} - 1$, and when i moves, it must be that $k_{ic_{m^*}h} = k_{ich} \geq 1 > k_{sc_{m^*}h_0} = k_{sch_0} = 0$. Substituting into $\mathbb{E}_i[u_i(c,g,q)|h,\sigma] - \mathbb{E}_i[u_i(c^*_m,g,q)|h,\sigma]$, and with $k = k_{ic_{m^*}h} - 1$ we get

$$\begin{aligned} &r^{a_{c_{m^*}}-k-1-1}[u(c,g,q)^{a_c-a_{c_{m^*}}} u(c,g,q) - u(c_{m^*},g,q) + r^{-(l_{sc_{m^*}\sigma}-1)}(1 - r^{a_c-a_{c_{m^*}}-(l_{ic\sigma}-l_{ic_{m^*}\sigma})})\overline{u}] = \\ &r^{a_{c_{m^*}}-k-1}[r^{-1}(u(c,g,q)^{a_c-a_{c_{m^*}}} u(c,g,q) - u(c_{m^*},g,q)) + r^{-(l_{sc_{m^*}\sigma})}(1 - r^{a_c-a_{c_{m^*}}-(l_{ic\sigma}-l_{ic_{m^*}\sigma})})\overline{u}] \end{aligned}$$

Since $r^{-1} > 1$, equation (B1) and Lemma B4 imply that $\mathbb{E}_i[u_i(c^*_m,g,q)|h,\sigma] < \mathbb{E}_i[u_i(c,g,q)|h,\sigma]$. □

Propositions 1.1 becomes:

Proposition B3 *(Essential uniqueness). Suppose Assumption 2.2 holds and consider MPEs σ_1, σ_2 with associated outcomes c_1, c_2 in states where $\tau_s = L$ and all agents $i \in c_1, c_2$ that do not reveal their type on equilibrium path under σ_1, σ_2 respectively have $\tau_i = L$. Coalitions c_1 and c_2 are essentially different for any $(\epsilon, q) \in (0, \infty) \times (0, 1) \setminus U$, where U has measure zero.*

Proof of Proposition B3. The proposition proves as in Appendix A, with the exception that $u_2 - u_1 = (r^{a_{c_2}-1})\frac{p_2}{a_{c_2}} - (r^{a_{c_1}-1})\frac{p_1}{a_{c_1}} + (1 - r^{a_{c_2}-l_{sc_2\sigma}-1})\overline{u} - (1 - r^{a_{c_1}-l_{sc_1\sigma}-1})\overline{u}$. The directional derivative at this point writes

$$\nabla_x u_2 - u_1 = \left[(r^{a_{c_2}-1})\frac{\partial p_2}{\partial q}\Big/a_{c_2} - (r^{a_{c_1}-1})\frac{\partial p_1}{\partial q}\Big/a_{c_1}\right]x_q$$

Assumption 2.2 implies that $\left[(r^{a_{c_2}-1})\frac{\partial p_2}{\partial q}\Big/a_{c_2} - (r^{a_{c_1}-1})\frac{\partial p_1}{\partial q}\Big/a_{c_1}\right] \neq 0$. The rest of the proof proceeds as the proof in Appendix A. □

Proposition 1.2 becomes:

Proposition B4. *Let $\sigma_q^* = \arg\max_{\sigma \in \mathcal{S}(s,g,q)} \hat{\epsilon}_s(g, q)$, $c_1^* \in \mathcal{C}^*(g, s, q_1, \sigma_{q_1}^*)$, $c_2^* \in \mathcal{C}^*(g, s, q_2, \sigma_{q_2}^*)$. If $q_1 < q_2$, then $\hat{\epsilon}_s(g, q_1) \geq \hat{\epsilon}_s(g, q_2)$. If Assumption 1.2 holds, then $q_1 < q_2 \Rightarrow a_{c_1^*} \leq a_{c_2^*}$.*

Proof of Proposition B4. Let's first show that $q_1 < q_2 \Rightarrow \hat{\epsilon}_s(g, q_1) \geq \hat{\epsilon}_s(g, q_2)$. From Lemma 2.2, if $c^* \in \mathcal{C}^*(g, s, q, \sigma_q^*)$, then $\hat{\epsilon}_s(g, q) = r^{a_{c^*}-1}u(c, g, q) + (1 - r^{a_{c^*}-l_{sc^*\sigma}-1}\overline{u})$. With $f(c, q, \sigma) \equiv r^{a_c-1}u(c, g, q) + (1 - r^{a_c-l_{sc\sigma}-1}\overline{u})$, we have $f(c_1^*, q_1, \sigma_1^*) \geq f(c_2^*, q_1, \sigma_2^*)$. Since for a given coalition, f is decreasing in q, we have $f(c_2^*, q_1, \sigma_2^*) \geq f(c_2^*, q_2, \sigma_2^*)$. This implies $f(c_1^*, q_1, \sigma_1^*) \geq f(c_2^*, q_2, \sigma_2^*)$, which implies $\hat{\epsilon}_s(g, q_1) \geq \hat{\epsilon}_s(g, q_2)$.

I now show that $q_1 < q_2 \Rightarrow a_{c_1^*} \leq a_{c_2^*}$. Let c_1^* be the largest coalition in $\bigcup_{\sigma \in \mathcal{S}(g,s,q_1)} \mathcal{C}^*(g, s, q_1, \sigma)$, and c_2^* the smallest in $\bigcup_{\sigma \in \mathcal{S}(g,s,q_2)} \mathcal{C}^*(g, s, q_2, \sigma)$ with sizes $a_{c_1^*}$ and $a_{c_2^*}$. To prove the claim, it suffices to show that $a_{c_1^*} \leq a_{c_2^*}$. Suppose not. Because $c_1^* \in \mathcal{C}^*(g, s, q_1, \sigma_1^*)$ and $c_2^* \in \mathcal{C}^*(g, s, q_2, \sigma_2^*)$ for some $\sigma_1^*, \sigma_2^* \in \mathcal{S}(g, s, q_1), \mathcal{S}(g, s, q_2)$ respectively, we have $f(c_2^*, q_1, \sigma_2^*) - f(c_1^*, q_1, \sigma_1^*) \leq 0$ and $f(c_2^*, q_2, \sigma^*2) - f(c_1^*, q_2, \sigma_1^*) \geq 0$. Since $f(c_2^*, q, \sigma_2^*) - f(c_1^*, q, \sigma_1^*)$ is continuous in q for any σ, there must be some $q \in [q_1, q_2]$ such that $f(c_2^*, q, \sigma_2^*) = f(c_1^*, q, \sigma_1^*)$. Because $a_{c_1^*} > a_{c_2^*}$, Assumption 2.3 implies that $\frac{\partial}{\partial q}\left[f(c_2^*, q, \sigma_2^*) - u_s(c_1^*, q, \sigma_1^*)\right] < 0$. Since $f(c_2^*, q_1, \sigma_2^*) - f(c_1^*, q_1, \sigma_1^*) \leq 0$, then $f(c_2^*, q_2, \sigma_2^*) - f(c_1^*, q_2, \sigma_1^*) < 0$, a contradiction. □

Proposition 1.3 and Corollary 1.3.1 do not hold anymore.
Proposition 1.4 accommodates the fact that Proposition 1.3 does not hold anymore. It becomes

Proposition B5. *If $g' = g + i \to j$, then $\hat{\epsilon}_s(g', q) \leq \hat{\epsilon}_s(g, q)$ for all $s \in \mathcal{N}$.*

Proof of Proposition B5. By Lemma A1, we have $\mathcal{C}(g', s) = \mathcal{C}(g, s)$ and for any $c \in \mathcal{C}(g, s)$, $w_{cg} \leq w_{cg'} \leq w_{cg} + 1$. This implies $v(a_c, w_{cg'}, q, \sigma) \leq v(a_c, w_{cg}, q, \sigma)$ for any $\sigma \in \mathcal{S}(s, g, q)$. So $\hat{\epsilon}_s(g', q) \leq \hat{\epsilon}_s(g, q)$. □

Proposition 1.5 remains unchanged, but its proof changes slightly.

Proof of Proposition 1.5. By Lemma A1, we have that for any $c \in \mathcal{C}(g, s)$, $v(a_c, w_{cg}, q, \sigma) = v(a_c, w_{cg'}, q, \sigma)$. As such, for any $s \in \mathcal{N}$, it cannot be that $\hat{\epsilon}_s(g', q) < \hat{\epsilon}_s(g, q)$. This implies that for all $s \in \mathcal{N}$, we have $\hat{\epsilon}_s(g', q) \geq \hat{\epsilon}_s(g, q)$. □

B.3. Proofs and additional results of section 2.3

In this section, we denote the three environments (equal-sharing, lawlessness, and monopoly) using the subscripts e, l, m respectively. In particular, $u^e(c, q, \epsilon) = \frac{p(a_c, w_{cg}, q)}{a_c} - \epsilon$ is the seed's utility under equal-sharing, while u^l and u^m are her utility under lawlessness and monopoly.

Proof of Proposition 2.3. Suppose coalition $c \in \mathcal{C}(g, s)$ is an equilibrium outcome for some $(\epsilon, q) \in (0, 1)^2$. Then it must be that $u_i(c, q) \geq 0$ for all $i \in c$ for otherwise, i has an incentive to deviate and reject her offer. If i is an operative, then at each of the histories where she moves on equilibrium path, her action space is to accept or reject an offer. Suppose that in equilibrium, i accepted transfer t_{ji} from broker j. If $u_i(c, q) > 0$, then j has an incentive to deviate and set t_{ji} such that $u_i(c, q) = 0$. □

Before proceeding to the main proofs, I prove a useful lemma. In equilibrium, if i moves at history h, there is a mapping between any of her transfers t_i and all outcome coalitions $\bar{\mathcal{C}}(g,h) \subseteq \mathcal{C}(g,s)$ that can be formed from that history. This proof requires a more specific definition of operatives and brokers. The definition is inductive. Node i is an operative at history h if in all of h's children histories where i moves, her action space does not contain any transfers. Node i is a level-1 broker at history h if in all of h's children histories where i moves, her action space only includes transfers to operatives. Node i is a level-n broker if in all children histories where i moves, her action space only includes transfers to operatives and brokers of level $n' < n$. I prove the following:

Lemma B5. *Suppose level-n broker $i \in \mathcal{N}$ moves at history h after transfer t_{ji}. In all equilibria of the subgame that begins at history h, i rejects t_{ji} if $\max_{c\in\bar{\mathcal{C}}(g,h)} u_i^l(c,q,\epsilon) \equiv t_{ji}p(a_c, w_{cg}, q) - \tau_i(c)\epsilon \geq 0$ for some $\tau_i : \bar{\mathcal{C}}(g,h) \to \mathbb{R}^+$ that satisfies $\tau_i(c) \geq a^i + 1$, where a^i is the number of accomplices in c hired in transfers that use resources from i's transfer, and $\frac{\partial \tau_i}{\partial \epsilon} = 0$. Otherwise, i accepts the transfer and some coalition in $\arg\max_{c\in\bar{\mathcal{C}}(g,h)} t_{ji}p(a_c, w_{cg}, q) - \tau_i(c)\epsilon$ is realized.*

Proof. I prove the claim by induction on the level of the broker. Suppose i is a level-1 broker. In equilibrium, her transfers make operatives indifferent. As such, under deference, if transfer t_i has coalition $c \in \bar{\mathcal{C}}(g,h)$ as an outcome, then $t_{ik} = \frac{\epsilon}{p(a_c,w_{cg},q)}$ if $k \in c$, and $t_{ik} = 0$ otherwise. Assuming i makes $a^i \geq 0$ transfers in such coalition, her payoff is $u_i^l(c,q,\epsilon) = (t_{ji} - a^i \frac{\epsilon}{p(a_c,w_{cg},q)})p(a_c, w_{cg}, q) - \epsilon = t_{ji}p(a_c, w_{cg}, q) - (a^i + 1)\epsilon$. Setting $\tau_i(c) \equiv a^i + 1$ proves the claim. We have $\frac{\partial a^i+1}{\partial \epsilon} = 0$. In equilibrium, i rejects t_{ji} if $\max_{c\in\bar{\mathcal{C}}(g,h)} u_i^l(c,q,\epsilon) < 0$. Otherwise, she accepts and makes the transfers that realize some coalition in $\arg\max_{c\in\bar{\mathcal{C}}(g,h)} u_i^l(c,q,\epsilon)$.

Suppose i is a level-n broker. In equilibrium, her transfers are the cheapest vector of transfers that realize the coalitions in $\bar{\mathcal{C}}(g,h)$. In particular, her transfers make recipients indifferent between their equilibrium move and their best outside option. Suppose recipient k's best outside option is to reject the transfer. Using the inductive hypothesis, in equilibrium, and under deference, t_{ik} solves $t_{ik}p(c,g,q,\epsilon) - \tau_k(c,g,q,\epsilon)\epsilon = 0$ if $k \in c$ and $t_{ik} = 0$ otherwise. That is, $t_{ik} = \frac{\tau_k(c,g,q,\epsilon)}{p(a_c,w_{cg},q)}\epsilon$. Suppose k's best outside option is to accept and make some other transfer resulting in coalition c'. Then t_{ik} solves $t_{ik}p(c,g,q,\epsilon) - \tau_k(c,g,q,\epsilon)\epsilon = t_{ik}p(c',g,q) - \tau_k(c',g,q,\epsilon)\epsilon$, which gives $t_{ik} = \frac{\tau_k(c,g,q,\epsilon)-\tau_k(c',g,q,\epsilon)}{p(a_c,w_{cg},q)-p(c',g,q)}\epsilon$ if $k \in c$. In equilibrium, i's payoff from $c \in \bar{\mathcal{C}}(g,h)$ is $u_i^l(c,g,q,\epsilon) = (t_{ji} - \sum_k t_{ik})p(a_c, w_{cg}, q) - \epsilon = t_{ji}p(a_c, w_{cg}, q) - (1 + \sum_k t_{ik}p(a_c, w_{cg}, q)/\epsilon)\epsilon$. Setting $\tau_i(c) \equiv 1 + \sum_k t_{ik}p(a_c, w_{cg}, q)/\epsilon$ proves the claim. Replacing t_{ik} by their equilibrium values and using the inductive hypothesis on τ_k, it is easy to show that $\frac{\partial}{\partial \epsilon} t_{ik}p(a_c, w_{cg}, q)/\epsilon = 0$, which implies $\frac{\partial \tau_i}{\partial \epsilon} = 0$. Furthermore, in equilibrium, any transfer t_i must make all the accomplices in c hired in transfers using resources from t_i better off than rejecting. For a^i such transfers, it must be that $\sum_k t_{ik} \geq a^i \frac{\epsilon}{p(a_c,w_{cg},q)}$. This implies $\tau_i(c) \geq a^i + 1$. □

Proof of Proposition 2.4. By Lemma B5, i accepts t_{ji} if $\max_{c\in\bar{\mathcal{C}}(g,h)} t_{ij}p(a_c, w_{cg}, q) - \tau_i(c)\epsilon \geq 0$. For any $c \in \bar{\mathcal{C}}(g,h)$, we have

$$t_{ij}p(a_c, w_{cg}, q) - \tau_i(c)\epsilon \geq 0 \iff t_{ij} \geq \frac{\tau_i(c)\epsilon}{p(a_c, w_{cg}, q)} \equiv t_{ji}^*(c) \geq 0$$

Therefore, $t_{ji}^* = \min_{c\in\bar{\mathcal{C}}(g,h)} t_{ji}^*(c)$. The rest of the proposition follows immediately from Lemma B5. □

I then prove that the seed has threshold strategies under both lawlessness and monopoly.

Lemma B6. *Under lawlessness, in equilibrium, s rejects the rent if $\max_{c\in\bar{\mathcal{C}}} u^l(c,q,\epsilon) \equiv p(a_c, w_{cg}, q) - \tau(c)\epsilon < 0$, for some $\bar{\mathcal{C}} \subseteq \mathcal{C}(g,s)$ and some $\tau : \bar{\mathcal{C}} \to \mathbb{R}^+$ that satisfies $\tau(c) \geq a_c$, and $\frac{\partial \tau}{\partial \epsilon} = 0$. Otherwise, s accepts the rent and some coalition in $\mathcal{C}^l(g,s,q,\epsilon) \equiv \arg\max_{c\in\bar{\mathcal{C}}} u^e(c,q,\epsilon)$ is realized.*

Proof. To prove the claim, use Lemma B5 and set $i \equiv s$, $\tau \equiv \tau_i$, define $t_{ji} \equiv 1$, and note that $a_c = a^s + 1$, $\bar{\mathcal{C}}(g,s,h) = \bar{\mathcal{C}}$. □

Lemma B7. *Under monopoly, in equilibrium, s rejects the rent if $\max_{c\in C_g} u^m(c,q,\epsilon) = p(a_c, w_{cg}, q) - a_c\epsilon < 0$. Otherwise, s accepts the rent and some coalition in $\mathcal{C}^m(g,s,q,\epsilon) \equiv \arg\max_{c\in C_g} u^m(c,q,\epsilon)$ is realized.*

Proof. Under monopoly, the seed's share of the rent is $t_m(s_c) = 1 - \sum_{i\in c\setminus\{s\}} \frac{\epsilon}{p(a_c,w_{cg},q)}$. So $u^m(c,q,\epsilon) = t_m(s,c)p(a_c, w_{cg}, q) - \epsilon = p(a_c, w_{cg}, q) - a_c\epsilon$. Consider $c^m \in \mathcal{C}^m(g,s,q,\epsilon)$. If $u^m(c^m,q,\epsilon) \geq 0$, then it is an equilibrium outcome, since non-seed members are indifferent between their equilibrium move and any other move, and c^m is the seed's favorite coalition. Since we consider equilibria with deference, a coalition $c \notin \mathcal{C}^m(g,s,q,\epsilon)$ cannot be an equilibrium outcome. If $u^m(c^m,q,\epsilon) < 0$, then the seed rejects the rent. □

Proof of Proposition 2.5. I first consider monopoly. From Lemma B7, the seed accepts the rent and a coalition in $\mathcal{C}^m(g,s,q,\epsilon)\subseteq\mathcal{C}(g,s)$ is realized whenever $\max_{c\in\mathcal{C}(g,s)} p(a_c,w_{cg},q)-a_c\epsilon\geq 0$. Otherwise, the seed rejects the rent. Therefore, we have $\hat{\epsilon}=\max_{c\in\mathcal{C}(g,s)}\frac{p(a_c,w_{cg},q)}{a_c}>0$. I now consider lawlessness. From Lemma B6, the seed accepts the rent and a coalition in $\bar{\mathcal{C}}\subseteq\mathcal{C}(g,s)$ is realized whenever $\max_{c\in\bar{\mathcal{C}}} p(a_c,w_{cg},q)-\tau(c,g,q,\epsilon)\epsilon\geq 0$. Otherwise, the seed rejects the rent. Therefore, we have $\hat{\epsilon}=\max_{c\in\mathcal{C}(g,s)}\frac{p(a_c,w_{cg},q)}{\tau(c,g,q,\epsilon)}$. Since $p(a_c,w_{cg},q)>0$ and $\tau(c,g,q,\epsilon)\geq a_c>0$, it must be that $\hat{\epsilon}>0$. □

Proof of Proposition 2.6. Showing that monopoly is efficient is a direct corollary of Lemma B7: if $\epsilon\leq\hat{\epsilon}^m$, then all equilibrium outcomes are efficient, since they solve $\max_{c\in\mathcal{C}(g,s)} u^m(c,q,\epsilon)$. Conversely, if $\epsilon>\hat{\epsilon}^m$, then no coalition yields a positive payoff. Corruption is inefficient, and the seed rejects the rent.

That $\hat{\epsilon}^m=\hat{\epsilon}^e\geq\hat{\epsilon}^l$ follows from Lemma 1.1 and the proof of Proposition 2.5: we have $\hat{\epsilon}^m=\hat{\epsilon}^e=\max_{c\in\mathcal{C}(g,s)} p(c,g,q)/a_c$, while $\hat{\epsilon}^l=\max_{c\in\bar{\mathcal{C}}}\frac{p(c,g,q)}{\tau(c,g,q,\epsilon)}$. Lemma B6 tells us that $\tau(c,g,q,\epsilon)\geq a_c$, which implies $\hat{\epsilon}^m\geq\hat{\epsilon}^l$.

Let's show that for any (q,ϵ), $\min_{c\in\mathcal{C}^e} a_c\leq\min_{c\in\mathcal{C}^m} a_c$, and $\max_{c\in\mathcal{C}^e} a_c\leq\max_{c\in\mathcal{C}^m} a_c$. Lemmas 1.1 and B7 tell us that the sets of equilibrium coalitions under equal-sharing and monopoly are, respectively, $\mathcal{C}^e=\mathcal{C}^e(g,s,q,\epsilon)=\arg\max_{c\in\mathcal{C}(g,s)} u^e(c,q,\epsilon)$, and $\mathcal{C}^m=\mathcal{C}^m(g,s,q,\epsilon)=\arg\max_{c\in\mathcal{C}(g,s)} u^m(c,q,\epsilon)$. Note that $u^e(c,q,\epsilon)=u^m(c,q,\epsilon)\iff\epsilon=p(c,g,q)/a_c$. So when $\epsilon=\hat{\epsilon}^m=\max_{c\in\mathcal{C}(g,s)} p(c,g,q)/a_c$, we have $\mathcal{C}^e(g,s,q,\epsilon)=\mathcal{C}^m(g,s,q,\epsilon)$. The claim is trivially true.

Consider the case where $\epsilon<\hat{\epsilon}^m$. Lemma 1.1 tells us that $\mathcal{C}^e(g,s,q,\epsilon)$ does not vary with ϵ. Conversely, the following lemma shows that as ϵ decreases, the coalitions in $\mathcal{C}^m(g,s,q,\epsilon)$ get larger. Using this lemma, it is immediate that $\min_{c\in\mathcal{C}^e} a_c\leq\min_{c\in\mathcal{C}^m} a_c$, and $\max_{c\in\mathcal{C}^e} a_c\leq\max_{c\in\mathcal{C}^m} a_c$.

Lemma B8. *Under monopoly, let $c_1^m\in\mathcal{C}^m(g,s,q,\epsilon_1)$ and $c_2^m\in\mathcal{C}^m(g,s,q,\epsilon_2)$. We have $\epsilon_1<\epsilon_2\Rightarrow a_{c_1^m}\geq a_{c_2^m}$.*

Proof. Let c_1 be the smallest coalition in $\mathcal{C}^m(g,s,q,\epsilon_1)$, and c_2 the largest in $\mathcal{C}^m(g,s,q,\epsilon_2)$ with sizes a_1 and a_2, and probabilities of success p_1 and p_2 respectively, for a given q. To prove the claim, iftsuffices to show that $a_1\geq a_2$. Suppose not. Because $c_1\in\mathcal{C}^m(g,s,q,\epsilon_1)$ and $c_2\in\mathcal{C}^m(g,s,q,\epsilon_2)$, we have $u^m(c_2,q,\epsilon_1)-u^m(c_1,q,\epsilon_1)\leq 0$ and $u^m(c_2,q,\epsilon_2)-u^m(c_1,q,\epsilon_2)\geq 0$. We have $u^m(c_2,g,q,\epsilon)-u^m(c_1,g,q,\epsilon)=(p_2-p_1)-(a_2-a_1)\epsilon$. Then, $\frac{\partial}{\partial q}\left[u^m(c_2,g,q)-u^m(c_1,g,q)\right]=a_1-a_2<0$, since $a_1<a_2$. Since $u^m(c_2,q,\epsilon_1)-u^m(c_1,q,\epsilon_1)\leq 0$, then $u^m(c_2,q,\epsilon_2)-u^m(c_1,q,\epsilon_2)<0$, a contradiction. □ □

I finally prove equilibrium in the repeated game.

Proof of Proposition 2.7. We prove this proposition using the single deviation principle. Note that whenever agents play according to σ_s, the SPNE of the stage game, then there are no profitable deviations, by definition of σ_s. As such, we only have to check for deviations in instances where $c_s^*\neq c_s$. Furthermore, in those stage games, we only have to check for deviations that may occur on the path of the stage game, since agents play according to σ_s at any other subgame.

First, note that the vector of payoffs $u_i(c_s^*,q)$ is feasible. Indeed, it must be that $\Pi(c_s^*)>\Pi(c_s)$, so we can set $u_i(c_s^*,q)=u_i(c_s,q)$ for all $i\in c_s^*$ and redistribute $\Pi(c_s^*)-\Pi(c_s)>0$ among all $i\in c_s^*$ so that $u_i(c_s^*,q)>u_i(c_s,q)$.

Then, suppose that a node i has a profitable deviation on the path of the stage game in period t with seed s_t and outcome $c_{s_t}^*$. Since no offer is rejected, it must be that $i\in c_{s_t}^*$. Let U_{it} be node i's flow of payoff on path, and suppose she has a profitable deviation that leads to the formation of coalition c' such that $u_i(c',q)>U_{it}$. If i deviates to coalition c', then her flow of payoffs from period $t+1$ is $U_{it+1}<U_{it}$. To see why $U_{it+1}<U_{it}$, consider the instantaneous payoff $u_i(c_{t'},q)$ for some period $t'\geq t+1$. If seed $s_{t'}\neq s_t$ gets picked, $u_i(c_{t'},q)$ remains unchanged. However, if $s_{t'}=s_t$, then $c_{t'}=c_s$ and so her instantaneous payoff decreases to $u_i(c_s,q)<u_i(c_{s_t}^*)$. As such, i has no incentive to deviate if and only if:

$$U_{it}\geq(1-\delta)u_i(c',q)+\delta U_{it+1}\iff\delta\geq\frac{u_i(c',q)-U_{it}}{u_i(c',q)-U_{it+1}}\equiv\bar{\delta}_{it}$$

Since $U_{it+1}<U_{it}<u_i(c')$, it must be that $\bar{\delta}_{it}<1$. Since, for any seed $s\in\mathcal{N}$, only two coalitions may be implemented (i.e. c_s^* or c_s), and since there is a finite number of players, there is a finite number of flow payoffs U_{it} and U_{it+1}. Similarly, since the stage game is finite, there is a finite number of profitable deviations c'. As such, there is a finite number of $\bar{\delta}_{it}$. Define $\bar{\delta}$ to be the minimum of the set of such $\bar{\delta}_{it}$'s. Since that set is finite, $\bar{\delta}$ exists. Since all $\bar{\delta}_{it}<1$, then $\bar{\delta}<1$. □

Robustness of findings under monopoly and repeated game

This subsection shows that findings under equal-sharing travel to monopoly and the repeated game. Again, this subsection only considers equilibria that satisfy deference (see Definition 4). I make the argument for monopoly only. Yet, because monopoly and the repeated game both implement the same coalitions – that is, coalitions that maximize the surplus given seed s –, the propositions travel to the repeated game. Specifically, I show that Lemmas 1.1, Propositions 1.1 to 1.5 and Corollary 1.3.1 are left virtually unchanged, under qualitatively similar assumptions. The new proofs address the difference that the set of coalitions that are optimal to the seed, $\mathcal{C}^*(g,s,q)$, now varies with ϵ. Similar to equal-sharing, define

$v^m(a, w, q, \epsilon) \equiv \frac{p(a,w,q)}{a} - [1 - p(a, w, q)]\epsilon$ the valuation of a coalition with a accomplices, w witnesses for monitoring q and cost ϵ. Assumptions 1.1 and 1.2 become, respectively:

Assumption B1. [Assumption 1.1 under monopoly and repeated game] If $v^m(a_1, w_1, q, \epsilon) = v^m(a_2, w_2, q, \epsilon)$ for some $a_1 \leq a_2$, $w_1, w_2 \in [1, N]$, $q \in (0, 1)$, $\epsilon > 0$, then $\frac{\partial p(a_2,w_2,q)}{\partial q} \Big/ \frac{\partial p(a_1,w_1,q)}{\partial q} \neq 1$ for any $q \in (0, 1)$.

Assumption B2. [Assumption 1.2 under monopoly and repeated game] If $v^m(a_1, w_1, q, \epsilon) = v^m(a_2, w_2, q, \epsilon)$ for some $a_1 \leq a_2$, $w_1, w_2 \in [1, N]$, $q \in (0, 1)$, $\epsilon > 0$, then $\frac{\partial p(a_2,w_2,q)}{\partial q} \Big/ \frac{\partial p(a_1,w_1,q)}{\partial q} < 1$ for any $q \in (0, 1)$.

Lemma 1.1 is left largely unchanged, with the exception that the set of optimal coalitions now varies with ϵ. We now have:

Lemma B9 *(Threshold strategy). Let $C^*_{gq\epsilon} = \arg\max_{c \in C_g} u^m(c, q, \epsilon)$. There is a threshold $\hat{\epsilon}(g, q) > 0$ such that all equilibria have the same outcome where s rejects the rent if $\epsilon > \hat{\epsilon}(g, q)$. Otherwise, she accepts it, and some coalition $c \in C^*_{gq\epsilon}$ is realized.*

Proof of Lemma B9. The proof follows directly from Lemma B7 and Proposition 2.5 above. □

Proposition 1.1 is left unchanged. Its proof changes slightly.

Proof of Proposition 1.1. The proof proceeds as the proof of Proposition 1.1 in the simple model (Appendix A), with the exception that equation (A1) now becomes:

$$\nabla_x u_2^m - u_1^m \equiv \frac{\partial u_2^m - u_1}{\partial \epsilon} x_\epsilon + \frac{\partial u_2^m - u_1^m}{\partial q} x_q = \left(\frac{\partial p_2}{\partial q} - \frac{\partial p_1}{\partial q}\right) x_q - (a_{c_2} - a_{c_1}) x_\epsilon \tag{B2}$$

where $x = (x_\epsilon, x_q)$ is a unit-length vector. As in the main specification, I show that equation (B2) has a finite number of solutions. This equation has an infinite number of solutions if and only if the coefficients on x_q and x_ϵ are both zero. Assumption B1 implies that $\frac{\partial p_1}{\partial q} - \frac{\partial p_2}{\partial q} \neq 0$.

The rest of the proof proceeds as the proof of Proposition 1.1 in the main specification of the model (Appendix A). □

Propositions 1.2 and 1.3 change slightly to accommodate the fact that the set of optimal coalitions now varies with ϵ. The proof of Proposition 1.2 changes slightly, while that of 1.3 does not. Corollary 1.3.1 and Propositions 1.4 and 1.5 and their proofs are left unchanged. Propositions 1.2 and 1.3 become, respectively:

Proposition B6. *Let $c_1^* \in \mathcal{C}^*(g, s, q_1, \epsilon)$, $c_2^* \in \mathcal{C}^*(g, s, q_2, \epsilon)$ for some $\epsilon \leq \hat{\epsilon}(g, q_2)$. If $q_1 < q_2$, then $\hat{\epsilon}_s(g, q_1) \geq \hat{\epsilon}_s(g, q_2)$. If Assumption B2 holds, then $q_1 < q_2 \Rightarrow a_{c_1^*} \leq a_{c_2^*}$.*

Proposition B7. *If $c \in \mathcal{C}^*(g, s, q, \epsilon)$ for some $q \in (0, 1)$, then c is minimal.*

Proof of Proposition B6. Let's first show that $q_1 < q_2 \Rightarrow \hat{\epsilon}(g, q_1) \geq \hat{\epsilon}(g, q_2)$. Suppose not. Let $\hat{\epsilon}_1 = \hat{\epsilon}(g, q_1)$ and $\hat{\epsilon}_2 = \hat{\epsilon}(g, q_2)$, and consider $c_1 \in \mathcal{C}^*(g, q_1, \hat{\epsilon}_1)$, $c_2 \in \mathcal{C}^*(g, q_2, \hat{\epsilon}_2)$. By Lemma B9, we have $u^m(c_1, q, \hat{\epsilon}_1) = u^m(c_2, q, \hat{\epsilon}_2) = 0$. Since u^m is strictly decreasing in ϵ, it must be that $u^m(c_2, q, \hat{\epsilon}_2) < u^m(c_2, q, \hat{\epsilon}_1)$. By Lemma B9, it must be that $u^m(c_2, q, \hat{\epsilon}_1) \leq u(c_1, q, \hat{\epsilon}_1)$. Therefore, $u(c_2, q, \hat{\epsilon}_2) < u(c_1, q, \hat{\epsilon}_1)$, a contradiction.

The rest of the proof proves as in Proposition 1.2 in the simple model. □

Proof of Proposition B7. The proof proves as Proposition 1.3 in the simple model. □

Appendix C. Simulations and experiment

In this section, I prove that the functional form for the probability of success p in equation (2) used in simulations and in the experiment satisfies Assumptions 1.1 and 1.2. Since Assumption 1.2 implies Assumption 1.1, I only prove that p satisfies Assumption 1.2.

***p* satisfies Assumption 1.2.** We have $\frac{\partial p(a_2,w_2,q)}{\partial q} \Big/ \frac{\partial p(a_1,w_1,q)}{\partial q} = \frac{N-(a_2+w_2)}{N-(a_1+w_1)}$. This implies

$$\frac{\partial p(a_2, w_2, q)}{\partial q} \Big/ \frac{\partial p(a_1, w_1, q)}{\partial q} \leq \frac{a_2}{a_1} \iff (N - w_1)a_2 - (N - w_2)a_1 \geq 0$$

Suppose first that $a_1 = a_2$. Since $\frac{\partial p}{\partial w} < 0$, it must be that $v(a_2, w_2, q) - v(a_1, w_1, q) \neq 0$ for any $q \in (0, 1)$. Suppose now that $a_1 < a_2$. We have $v(a_2, w_2, q) - v(a_1, w_1, q) \propto (a_2 - a_1) - (1 - q)[(N - w_1)a_2 - (N - w_2)a_1]$. Since $(a_2 - a_1)$ and $(1 - q) > 0$, $v(a_2, w_2, q) - v(a_1, w_1, q) = 0$ requires $(N - w_1)a_2 - (N - w_2)a_1 > 0$. □

Appendix D. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.geb.2020.08.013.

References

Abbink, Klaus, 2004. Staff rotation as an anti-corruption policy: an experimental study. Eur. J. Polit. Econ. 20 (4), 887–906. http://www.sciencedirect.com/science/article/pii/S0176268003000843.

Abbink, Klaus, Irlenbusch, Bernd, Renner, Elke, 2002. An experimental bribery game. J. Law Econ. Organ. 18 (2), 428–454. http://www.jstor.org/stable/3555050.

Abbink, Klaus, Hennig-Schmidt, Heike, 2006. Neutral versus loaded instructions in a bribery experiment. Exp. Econ. 9 (2), 103–121.

Amir, Rabah, Burr, Chrystie, 2015. Corruption and socially optimal entry. J. Public Econ. 123, 30–41.

Andvig, Chr Jens, Moene, Karl Ove, 1990. How corruption may corrupt. J. Econ. Behav. Organ. 13 (1), 63–76.

Aven, Brandy, 2015. The paradox of corrupt networks: an analysis of organizational crime at enron. Organ. Sci. 26 (4), 980–996.

Azfar, Omar, Nelson, William R., 2007. Transparency, wages, and the separation of powers: an experimental analysis of corruption. Public Choice 130 (3/4), 471–493.

Baccara, Mariagiovanna, Bar-Isaac, Heski, 2008. How to organize crime. Rev. Econ. Stud. 75 (4), 1039–1067.

Baccara, Mariagiovanna, Bar-Isaac, Heski, 2009. Interrogation methods and terror networks. In: Memon, Nasrullah, Farley, Jonathan D., Hicks, David L., Rosenorn, Torben (Eds.), Mathematical Methods in Counterterrorism. Springer, Vienna, pp. 271–290.

Baker, Wayne E., Faulkner, Robert R., 1993. The social organization of conspiracy: illegal networks in the heavy electrical equipment industry. Am. Sociol. Rev. 58 (6), 837–860.

Ballester, Coralio, Calvó-Armengol, Antoni, Zenou, Yves, 2010. Delinquent networks. J. Eur. Econ. Assoc. 8 (1), 34–61.

Banerjee, Abhijit, Hanna, Rema, Mullainathan, Sendhil, 2013. Corruption. In: Gibbons, Robert, Roberts, John (Eds.), The Handbook of Organizational Economics. Princeton University Press, Princeton, pp. 1109–1147. chapter 27.

Barabási, Aibert László, Albert, Réka, 1998. Emergence of scaling in random networks. Science 286 (5439), 509–512.

Barr, Abigail, Serra, Danila, 2009. The effects of externalities and framing on bribery in a Petty corruption experiment. Exp. Econ. 12 (4), 488–503.

Barr, Abigail, Lindelow, Magnus, Serneels, Pieter, 2009. Corruption in public service delivery: an experimental analysis. J. Econ. Behav. Organ. 72 (1), 225–239.

Bennet, Richard, 2012. A change agent in the tax office: Nigeria's federal inland revenue service, 2004–2009. In: Innovations for Successful Societies Case Study.

Berninghaus, Siegfried K., Haller, Sven, Krüger, Tyll, Neumann, Thomas, Schosser, Stephan, Vogt, Bodo, 2013. Risk attitude, beliefs, and information in a corruption game – an experimental analysis. J. Econ. Psychol. 34, 46–60.

Besley, Timothy, McLaren, John, 1993. Taxes and bribery: the role of wage incentives. Econ. J. 103 (416), 119–141.

Burgess, Robin, Hansen, Matthew, Olken, Benjamin A., Potapov, Peter, Sieber, Stefanie, 2012. The political economy of deforestation in the tropics. Q. J. Econ. 127 (4), 1707–1754.

Calvó-Armengol, Antoni, Zenou, Yves, 2004. Social networks and crime decisions: the role of social structure in facilitating delinquent behavior. Int. Econ. Rev. 45 (3), 939–958.

Dixit, Avinash K., 2004. Lawlessness and Economics. Princeton University Press.

Friedman, Jonathan, 2012. Saving a sinking agency: the national port authority of Liberia, 2006–2011. In: Innovations for Successful Societies Case Study.

Gambetta, Diego, 1996. The Sicilian Mafia: The Business of Private Protection. Harvard University Press, Cambridge, Mass.

Gonzalez, Luis, Güth, Werner, Levati, Maria Vittoria, 2002. Speeding up bureaucrats by greasing them - an experimental study. Technical Report 2002-05. http://ideas.repec.org/p/esi/discus/2002-05.html.

International Monetary Fund - Staff Team from the Fiscal Affairs Department and the Legal Department. Corruption: Costs and mitigating strategies - IMF Staff Discussion Note no. SDN/16/05. Technical report.

Jávor, István, Jancsics, David, 2013. The role of power in organizational corruption: an empirical study. Administration & Society 48 (5), 527–558.

Johnson, Eric J., Camerer, Colin, Sen, Sankar, Rymon, Talia, 2002. Detecting failures of backward induction: monitoring information search in sequential bargaining. Journal of Economic Theory 104 (1), 16–47.

Kaufmann, Daniel, 2004. Corruption, governance and security: challenges for the rich countries and the world. SSRN Electronic Journal, 83–102.

Khanna, Vikramaditya, Kim, E. Han, Lu, Yao, 2015. CEO connectedness and corporate fraud. The Journal of Finance 70 (3), 1203–1252.

Laffont, Jean-Jacques, 1990. Analysis of hidden gaming in a three-level hierarchy. Journal of Law, Economics & Organization 6 (2), 301.

Laffont, Jean-Jacques, Tirole, Jean, 1991. The politics of government decision-making: a theory of regulatory capture. The Quarterly Journal of Economics 106 (4), 1089–1127.

Lambsdorff, Johann G., Frank, Björn, 2010. Bribing versus gift-giving - an experiment. Journal of Economic Psychology 31 (3), 347–357.

Ledeneva, Alena, 1998. Russia's Economy of Favours: Blat, Networking and Informal Exchanges. Cambridge University Press, Cambridge.

Manea, Mihai, 2017. Bottleneck links, essential intermediaries, and competing paths of diffusion in networks. Working paper. https://economics.mit.edu/files/12516.

McAfee , R. Preston, McMillan, John, 1995. Organizational diseconomies of scale. Journal of Economics & Management Strategy 4 (3), 399–426.

Melumad, Nahum D., Mookherjee, Dilip, Reichelstein, Stefan, 1995. Hierarchical decentralization of incentive contracts. The RAND Journal of Economics 26 (4), 654–672.

Mesmer-Magnus, Jessica R., Viswesvaran, Chockalingam, 2005. Whistleblowing in organizations: an examination of correlates of whistleblowing intentions, actions, and retaliation. Journal of Business Ethics 62 (3), 277–297.

Morselli, Carlo, Giguère, Cynthia, Petit, Katia, 2007. The efficiency/security trade-off in criminal networks. Social Networks 29 (1), 143–153.

Morton, Rebecca, Tyran, Jean-Robert, 2015. Corruption in committees: an experimental study of information aggregation through voting. Journal of Public Economic Theory 17 (4), 553–579.

Nyblade, Benjamin, Reed, Steven R., 2012. Who cheats? Who loots? Political competition and corruption in Japan, 1947–1993. American Journal of Political Science 52 (4), 926–941.

Olken, Benjamin A., Pande, Rohini, 2012. Corruption in developing countries. Annual Review of Economics 4 (1), 479–509. http://www.annualreviews.org/doi/10.1146/annurev-economics-080511-110917.

Polanski, Arnold, 2007. A decentralized model of information pricing in networks. Journal of Economic Theory 136 (1), 497–512.

Ryvkin, Dmitry, Serra, Danila, Tremewan, James, 2017. I paid a bribe: an experiment on information sharing and extortionary corruption. European Economic Review 94, 1–22.

Schickora, Jan Theodor, 2011. Bringing the four-eyes-principle to the lab. Working paper. http://epub.ub.uni-muenchen.de/12160/.

Shleifer, Andrei, Vishny, Robert W., 1993. Corruption. The Quarterly Journal of Economics 108 (3), 599–617.

Skibski, Oskar, Rahwan, Talal, Michalak, Tomasz P., Wooldridge, Michael, 2019. Enumerating Connected Subgraphs and Computing the Myerson and Shapley Values in Graph-Restricted Games. ACM Transactions on Intelligent Systems and Technology, vol. 10. Association for Computing Machinery.

Spenkuch, Jorg, Montagnes, Pablo, Magleby, Daniel, 2018. Backward induction in the wild? Evidence from sequential voting in the U.S. Senate. American Economic Review 108 (7), 1971–2013.

Ting, Michael M., 2008. Whistleblowing. American Political Science Review 102 (2), 249–267.

Tirole, Jean, 1986. Hierarchies and bureaucracies: on the role of collusion in organizations. Journal of Law, Economics and Organization 2 (2), 181–214.

US Department of Justice, 2015a. Nine FIFA officials and five corporate executives indicted for racketeering conspiracy and corruption. http://www.justice.gov/opa/pr/nine-fifa-officials-and-five-corporate-executives-indicted-racketeering-conspiracy-and.

US Department of Justice, 2015b. Sixteen additional FIFA officials indicted for racketeering conspiracy and corruption. https://www.justice.gov/opa/pr/sixteen-additional-fifa-officials-indicted-racketeering-conspiracy-and-corruption.

Vannucci, Alberto, Della Porta, Donatella, 2013. The Hidden Order of Corruption: An Institutional Approach. Ashgate Publishing, Ltd., Farnham.

Wade, Robert, 1982. The system of administrative and political corruption: canal irrigation in South India. Journal of Development Studies 18 (3), 287–328.

Watts, Duncan J., Strogatz, Steven H., 1998. Collective dynamics of ‘small-world’ networks. Nature 393, 440–442.