Partners in crime? Corruption as a criminal network

Romain Ferrali *

January 17, 2019
9,479 words

Abstract

Bureaucratic corruption undermines both the legitimacy and capacity of the government. While much of the literature treats corruption as the product of individual miscreants, this paper analyzes a model and that recognizes the coalitional nature of corruption, which is based on networks of accomplices, and tests it in a lab-in-the field experiment. Model predictions and experimental results emend the standard prediction that organizationally isolated bureaucrats are most prone to corruption. We find that corruption will arise in isolated bureaucratic enclaves; that is, sets of jointly isolated bureaucrats, and, surprisingly, that the size of the enclaves will increase when confronted by more effective enforcement effort on the part of the government. Additionally, simulation results suggest that both imperfect hierarchies and imperfectly flat organizations may feature enclaves. This suggests it would be sensible to redesign government agencies to puncture the isolation of enclaves.

JEL Keywords: C92, C93, D73, D85

---

*Social Science Division, New York University Abu Dhabi, rferrali@nyu.edu I thank Michael Bauer, Roland Benabou, Carles Boix, Stefano Caria, Paul DiMaggio, Johannes Haushoffer, Robert Gibbons, Guy Grossman, Matias Iaryczower, Kosuke Imai, Amaney Jamal, Paul Lagunes, Jenn Larson, John Londregan, Joaquin Morales Belpaire, Rebecca Morton, Salvatore Nunnari, Betsy Paluck, Piero Stanig, Bassel Tarbush, Leonard Wantchekon, Viviana Zelizer, Xiang Zhou, as well as seminar participants at CASBS workshop, CSSO workshop, Harvard University, the Imai Research Group, Oxford University, Princeton University and the WESSI workshop for their helpful comments. I gratefully acknowledge support from Amaney Jamal, the Center for the Study of Social Organizations and the Mandouha S. Bobst Center for Peace and Justice.
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 (see Olken and Pande 2011, for a review). 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. 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 collective corruption and the formation of criminal networks. 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 corruption takes place in. I examine a model and a lab experiment that look at corruption as the result of a diffusion process on a network. In this setting, a bureaucrat 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 bureaucrat 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 important substantive implications. First, char-

---

1 See, e.g., Besley and McLaren (1993) and Banerjee, Hanna and Mullainathan (2013) for principal-agent approaches to the impact of wage incentives and monitoring technologies on corruption, respectively. See Abbink, Irlenbusch and Rennert (2002); Barr and Serra (2009); Ryvkin, Serra and Tremewan (2017) for lab experiments on bribery.

2 See Tirole (1986); Laffont (1990); Laffont and Tirole (1991); McAfee and McMillan (1995); Melumad, Mookherjee and Reichelstein (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, Calvó-Armengol and Zenou (2010) for models of peer-effects in crime. Finally, see Gonzalez, Gülth and Levati (2002); Schickora (2011); Barr, Lindelow and Serenel (2009); Azfar and Nelson (2007); Berninghaus et al. (2013); Morton and Tyran (2015) for experiments on collective corruption.
acterizing how corruption is organized, I show that equilibrium coalitions are the most *enclaved* portions of the organization: they jointly minimize exposure to witnesses. 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. Enclaves have few ties to the out-group, but may exhibit plentiful in-group ties; in other words, what matters is not the number of connections held by corrupt individuals, but the number of connections between the coalition and its witnesses.

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 enclaves to additional witnesses. However, these ties have no effect if they do not target enclaves. Worse, they may increase corruption by facilitating access to existing enclaves. I compare across organizational structures using simulations and reproduce the finding that perfectly flat organizations are less corrupt, while perfectly hierarchical organizations are more corrupt, because hierarchies give birth to enclaves. I also show, however, that there is significant noise in between, due to small structural details: relatively flat organizations can be very enclaved, while relatively hierarchical organizations can feature few enclaves. This suggests that while it would be sensible to redesign government agencies to puncture the isolation of enclaves, this must be done carefully. Organizational reforms may be counterproductive, which is worrying, because these reforms are popular, but rarely subjected to careful evaluation.

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.

In a series of extensions, I show that behavior is largely robust to alternative assumptions about the cost of corruption and the ability of corrupt individuals to strike informal contracts that police their interactions. Additionally, the exercise reveals that the lawlessness (Dixit 2004) inherent to criminal activities introduces inefficiencies that benefit brokers – accomplices that recruits other accomplices, – who exploit their control over the diffusion process to extract higher shares of the surplus.

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 manipulation and measurement of network structures and easy measurement of corruption, two constructs that hare 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: an organizational structure represented by

---

Aven (2015) and Morselli, Giguére and Petit (2007) show that more isolated individuals are more corrupt while Nyblade and Reed (2012) and Khanna, Kim and Lu (2015) show that they are less corrupt. See conclusion for an extended discussion.
a social network that conditions offers and monitoring, as well as a lawless environment where corruption is risky and corrupt individuals cannot commit to recruit the agents that their accomplices find desirable. Participants may, however, leverage cheap talk in order to devise informal contracts to alleviate the commitment problem. Being minimal, the design controls for all other potential sources of confounding, such as moral considerations about corruption. To increase ecological validity, I hold the experiment in Morocco, a mid-income country with median levels of corruption 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 enclaves are more corrupt, and only some ties reduce corruption. Consistently with the predictions of the model, the overall incidence of corruption falls when facing better monitoring, but the corruption that does occur takes place on a larger scale, involving more accomplices. The most striking deviation from theoretical predictions is that agents disproportionately fail to realize larger coalitions, because they pose harder coordination problems. Furthermore, behavior does not differ across subject pools and thus appears to be largely robust to factors outside the model.

Treating corruption as a game of strategic diffusion on a network, this paper contributes to our understanding of corruption in organizations by allowing to consider arbitrarily complex forms of collusion in any kind of organization. Most closely related 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). In particular, 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, 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 (section [1]). I then take the model to the lab to test these propositions (section [2]). I conclude by situating results in the literature, and discussing the main model assumptions and design choices made in the experiment.

---

4Morocco is ranked 81 out of 180 countries in the Transparency International Corruption Perception Index 2017.

5Hierarchy 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.
1 Theory

1.1 Model

This subsection describes the model in its main form and in several variants I examine later as extensions.

The most important series of variants examines different assumptions about the informal contracts that govern interactions among corrupt individuals. Like other forms of criminal activity, corruption occurs in a lawless environment. Although corrupt deals cannot be enforced in court, corrupt bureaucrats typically use a wide array of informal contracts to police their interactions. Because of the nebulous nature of informal contracts, I bound behavior by comparing two different environments: a lawless environment $\Gamma_l$, where contracts cannot be enforced, and a contractual environment $\Gamma_c(t)$, where accomplices agree ex-ante to divide the surplus according to some rule $t$. Within the contractual environment, I examine two opposite division rules: an egalitarian rule, *equal-sharing*, where accomplices agree ex-ante to divide the surplus equally, and an unegalitarian rule, *monopoly*, where accomplices the bureaucrat that initiates the diffusion process gets all the surplus.

While, for simplicity, the main model considers equal-sharing and treats the lawless environment and monopoly as extensions, I first introduce the game in the lawless environment $\Gamma_l$ and then detail the changes introduced by the contractual environment $\Gamma_c$. Additionally, I show in Appendices A.4 and A.5 that results are robust to extensions where I relax some assumptions about the cost of corruption.

1.1.1 Game in the lawless environment

I model corruption as a dynamic game of complete information. Bureaucrats are the nodes of the finite exogenous multiplex graph $g = (N, G_c, G_m)$ where $N$ is a set of nodes indexed from 1 to $|N|$, and $G_c$ and $G_m$ are sets of ties. $G_c$ is an undirected communication network, with $ij \in G_c$ 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 $G_c$ is connected. $G_m$ is a directed monitoring network, with $i \rightarrow j \in G_m$ meaning that $i$ monitors $j$; in other words, $i$ would hold incriminating evidence on $j$ (and turn into a witness should $j$ be corrupt). $G_m$ 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: $i \rightarrow j \in G_m \Rightarrow ij \in G_c$.

A randomly drawn node $s \in N$, the *seed*, discovers an illegal stream of rents which 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. Otherwise, she becomes an accomplice, and (1) she pays a cost $\epsilon_s \geq 0$; (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 $\epsilon_s$ represents effort expended by the seed when engaging in corruption; for instance, learning about corrupt practices, or engaging in the criminal activity itself. Conversely, $1 - \epsilon_s$ represents the benefit of corruption relative to its cost; that is, the scale of corruption. When $\epsilon_s$ is low, corruption is very profitable compared to its cost, indicating high-scale, grand corruption. Conversely, high values of $\epsilon_s$ 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 Appendix A.4).

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. They can reject their offer, or accept it. If node $i$ accepts, she becomes an accomplice and holds the transfer $t_{si}$. Like the seed, she pays the cost $\epsilon_i > 0$, her non-pending, non-accomplice in-neighbors on $G_m$ 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 $G_c$. I assume that $\epsilon_i \leq \epsilon_s$, to capture the idea that as the instigator, the seed may expend more effort than her accomplices. Once all pending nodes have moved, the players to whom they have made offers (if any) can act. They face the same action space, 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 (Figure 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 $G_m$. 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 $W_{cg} \equiv \{i \in N : i \rightarrow j \in G_m, i \notin c, j \in c\}$, and $w_{cg} \equiv |W_{cg}|$ the number of witnesses. Let $C$ be the set of coalitions that can be formed on any graph with $N$ nodes. A

---

**Figure 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.
coalition \( c \) is feasible on graph \( g \) if it is consistent with some diffusion process originating from the seed; formally:

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

Once the coalition is formed, an exogenous enforcer detects the coalition with probability \( 1 - p \), where \( p \equiv p(a, w, q) : \{1, ..., N\} \times \{0, ..., N\} \times (0, 1) \rightarrow (0, 1) \) 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 assumptions on \( p \) that reflect what we know about the impact of accomplices, witnesses, and monitoring on the probability of detection. Of course, better monitoring makes detection more likely; as such, \( \frac{dp}{dq} < 0 \). 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 makes additional accomplices pose a tradeoff: they help the coalition, protecting it against detection by “covering up,” \([Wade 1982, Ledeneva 1998]\)\(^9\). On the other hand, they cost fractions of the rent and create additional witnesses among their colleagues, which increases risk \([Baker and Faulkner 1993]\)\(^10\). I use an all-or-nothing probability of detection to model 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. Finally, note that the probability of detection does not depend on the scale of corruption, \( 1 - \epsilon_s \), which may be problematic: grand corruption might be more visible and easier to detect than petty corruption. Conversely, grand corruption might be better able to protect itself against detection. I show in Appendix A.5 that results are robust to relaxing this assumption.

If player \( i \) is not a member of the coalition at a terminal history her payoff is 0. Otherwise, she incurred cost \( \epsilon_i \) and holds some share of the rent \( \pi_i \geq 0 \). Suppose \( i \) accepted offer \( t_{ki} \). Let \( r_{ij} \equiv 1 \) if \( j \) accepted an offer \( t_{ij} \) from \( i \), and \( r_{ij} \equiv 0 \) otherwise; then \( \pi_i = t_{ki} - \sum_{j \in P_i} t_{ij} r_{ij} \), where \( P_i \) is the set of \( i \)'s susceptible neighbors when she made her offers \( t_i \). With probability \( p \), agent \( i \) gets her share \( \pi_i \), and gets 0 otherwise. With risk-neutral agents, the expected utility of agent \( i \) writes

\[
\begin{align*}
    u_i(c, g, q) = \begin{cases} 
    \pi_i p(a_c, w_{cg}, q) - \epsilon_i, & \text{if } i \in c \\
    0, & \text{otherwise}
    \end{cases}
\end{align*}
\]

(1)

The diffusion process leads to the formation of a coalition of accomplices: a corrupt subgraph of \( g \). I describe equilibria 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 $1 - \epsilon_s$.

1.1.2 Game in the contractual environment

The game $\Gamma_l$ described in the previous subsection corresponds to a lawless environment where agents cannot commit to a schedule of transfers: when $j$ receives a transfer from $i$, she cannot commit to implement the transfers that $i$ prefers. I describe here the variant $\Gamma_c(t)$ where agents operate in a contractual environment; that is, agents are able to contract upon a specific schedule of transfers. In this environment, accomplices only offer their neighbors to join the coalition, and some division rule $t : N \times C_g \rightarrow [0, 1]$ assigns a payoff $\pi_i \geq 0$ to each member of the coalition $i \in c$, and a payoff of 0 to each $i \notin C$. The function $t$ satisfies the budget constraint $\sum_{i \in c} t(i, c) \leq 1$.

I consider two such division rules. The first one is equal-sharing, where accomplices split the bribe equally among themselves. Its division rule is $t_e(i, c) = \frac{1}{ac}$. The second one is monopoly, where the seed pockets all the surplus; that is, $t_m(i, c) = \frac{\epsilon_i}{p(a_c, w_c, q)}$ for $i \in c \setminus \{s\}$, and $t_m(s, c) = 1 - \sum_{i \in c \setminus \{s\}} t_m(i, c)$. In this setting, the expected utility $u_i$ of agent $i$ under rule $t$ becomes, from equation 1

$$u_i(c, g, q) = \begin{cases} t(i, c)p(a_c, w_{cg}, q) - \epsilon_i, & \text{if } i \in c \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

1.1.3 Assumptions under equal-sharing

I now make two additional assumptions on the impact of accomplices and witnesses on the probability of success for the game in the contractual environment under equal-sharing, $\Gamma_c(t_e)$. In section 1.3, where I compare equal-sharing to the other division rules, I introduce comparable assumptions for these division rules.

I first introduce additional notation. Under equal-sharing, the expected utility of an accomplice in equation 1 becomes $u_i(c, g, q) = \frac{p(a_c, w_{cg}, q)}{a_c} - \epsilon_i$. Utilities under equal-sharing have a common component $\frac{p(a_c, w_{cg}, q)}{a_c}$, and an individual component, the cost $\epsilon_i$. To distinguish analytical and graphical considerations, I separate the valuation of a coalition and its expected utility. The valuation of a coalition, $v : \{1, ..., N\} \times \{0, ..., N\} \times (0, 1) \rightarrow \mathbb{R}$ is defined for an arbitrary number of accomplices and witnesses, with $v(a, w, q) = \frac{p(a, w, q)}{a}$. The expected utility of a coalition $u : C_g \times (0, 1) \rightarrow \mathbb{R}$ is the valuation of existing coalitions on specific graphs: $u(c, g, q) = v(a_c, w_{cg}, q)$.

I first assume that larger coalitions are sufficiently resistant against better monitoring technologies. The marginal effect of better monitoring might be decreasing in the size of the coalition. For instance, a large number of forged testimonies may prove all the more effective as the amount of hard evidence gathered by the monitoring technology is high. Conversely, it might be increasing. For instance, larger coalitions might prove easier to detect under better monitoring. If the latter is true, I assume that the handicap of large coalitions should not be too high. Formally:
**Assumption 1** (Larger coalitions are sufficiently resistant against better monitoring). If \(v(a_1, w_1, q) = v(a_2, w_2, q)\) for some \(a_1 \leq a_2, \ w_1, w_2 \in \{0, ..., N\}, \ q \in (0, 1)\), then \(\frac{\partial v(a_2, w_2, q)}{\partial q} \leq \frac{\partial v(a_1, w_1, q)}{\partial q} < \frac{a_2}{a_1}\) for any \(q \in (0, 1)\).

The second assumption imposes regularity conditions to the valuation of coalitions \(v\). I fix the number of witnesses and compare coalitions of different sizes. I ensure that the coalition with highest valuation is unique. Additionally, I make assumptions on the probability of success \(p\) that either imply that \(v\) is monotonic in the size of coalitions, or that additional witnesses sufficiently hurt coalitions. The assumption reads

**Assumption 2** (Witnesses are sufficiently consequential). For any \(a \in \{1, ..., N\}, w \in \{0, ..., N\}, q \in (0, 1), p\) is such that \(v\) is quasi-concave in \(a\) and either its growth rate in \(a\), 
\[
\frac{p(a+1, w, q) - p(a, w, q)}{p(a, w, q)}, \n\]
is bounded by \(\frac{1}{a}\) or \(p\) satisfies
\[
|p(a, w + 1, q) - p(a, w, q)| > a \left[v(a^*, w, q) - \max \{v(1, w, q), v(N, w, q)\}\right],
\]
where \(a^* \in \arg \max_{a \in \{1, ..., N\}} v(a, w, q)\).

### 1.2 Results under equal-sharing

This section derives equilibrium and comparative statics for the game in the contractual environment under equal-sharing, \(\Gamma_c(t_e)\). All proofs are in Appendix A.

#### 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, \(c^* \in \arg \max_{c \in C_q} u(c, g, q)\). If \(u(c^*, g, q) < \epsilon_s\), then \(s\) does not take the rent, since her favorite coalition does not cover her cost. If \(u(c^*, g, q) \geq \epsilon_s\), 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** (Threshold strategy). Let \(C^*_{gq} = \arg \max_{c \in C_g} u(c, g, q)\) and \(c^* \in C^*_{gq}\). There is a threshold \(\hat{\epsilon}_s(g, q) = u(c^*, g, q) \in (0, 1)\) such that all equilibria have the same outcome that \(s\) rejects the rent if \(\epsilon_s > \hat{\epsilon}_s(g, q)\). Otherwise, she accepts it, and some coalition \(c \in C^*_{gq}\) is realized.

Saying more about which coalitions are realized in equilibrium requires characterizing the coalitions that the seed prefers. To do so, I characterize the frontier, i.e., the coalitions that lie in \(C^*_{gq}\). Comparing same-sized coalitions, one prefers the one with fewer witnesses. Assumptions [1] and [2] give more traction, allowing comparisons between coalitions of different sizes.

Assumption [1] implies that although there are many equilibria, equilibrium coalitions are *essentially unique*: they have the same counts of accomplices and witnesses. Indeed, the assumption implies that two essentially different coalitions may give the same payoff on at most a curve in \((\epsilon, q)\). Formally:

\[\text{That is, } v \text{ satisfies } v(a, w, q) \geq \min\{v(a_1, w, q), v(a_2, w, q)\} \text{ for any } a_1 \leq a \leq a_2 \in \{1, ..., N\}.\]
Figure 2: **Set of feasible coalitions** \( C_g \) **for some graph** \( g \). Points are sets of essentially unique coalitions. Coalitions in black minimize \( w \) for a fixed \( a \). Because \( c \) is above the dotted line and \( c_2 \) is below, \( c_1 \) and \( c_2 \) beat \( c \) (Lemma 2). Coalitions on the black line are minimal (Definition 2). By lemma 2 they beat the ones above them. One of them is realized in equilibrium (Proposition 2).

**Proposition 1** (Essential uniqueness). Suppose assumption 2 holds. Then equilibrium coalitions are essentially unique for any \((\epsilon_s, q) \in (0, \infty) \times (0, 1) \setminus U \), where \( U \) has measure zero.

Assumption 2 allows ruling out the coalitions that are too exposed, and characterizing the frontier. Consider coalitions \( c_1 \) and \( c_2 \), where \( c_2 \) is larger than \( c_1 \) and has less witnesses. Consider a third coalition \( c \) in between, with more witnesses than either (Figure 2). If assumption 2 holds, then coalitions \( c_1 \) or \( c_2 \) are always preferred to \( c \):

**Lemma 2.** Let \( a_1 < a_c < a_2 \), and \( w_2 \leq w_1 < w_c \). If assumption 2 holds, then

\[
v(a_c, w_c, q) < \max\{v(a_1, w_1, q), v(a_2, w_2, q)\}
\]

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

All the coalitions that lemma 2 does not rule out belong to the frontier because they are minimal, in the sense that no smaller coalition has fewer witnesses (Figure 2).

**Definition 2.** \( M_g \equiv \{ c \in C_g : a_{c'} \leq a_c \implies w_{c'} \geq w_c \text{ for any } c' \in C_g \} \) is the set of minimal coalitions on graph \( g \).

Because they minimize the number of witnesses, minimal coalitions are realized in equilibrium. Of course, the result uses lemma 2 and, as such, hinges on assumption 2. We have:

**Proposition 2** (Equilibrium coalitions are minimal). If assumption 2 holds and \( c \in C_{gq}^* \) for some \( q \in (0, 1) \), then \( c \) is minimal.

Proposition 2 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. The concept shifts the unit of analysis from the individual to the coalition, and is agnostic about tie density within the in-group. I say that a
Partners in crime

0 = bad
0 = high-scale low-scale = 1
monitoring technology, q
scale of corruption, ε
s

good = 1
reject
c1 c2 c3
ε
^s

Figure 3: 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 3).

set of nodes that has few monitoring ties pointing to it from the out-group—although there may be many such ties within the in-group—is relatively enclosed or isolated. Minimal coalitions are the most enclosed coalitions of a graph. I discuss empirical implications in the conclusion.

1.2.2 Comparative statics

Having pinned down equilibrium behavior allows characterizing how corruption varies 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, how does corruption vary as the organizational structure changes?

Examining how corruption varies as organizations adopt better monitoring technologies, I show that corruption is less frequent but has a higher scale and a broader scope under better monitoring (Figure 3). 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 3** (Corruption is less frequent under better monitoring technologies but grows in scope and selects on large-scale). Let \(c^*_1 \in C^*_{gq_1}, c^*_2 \in C^*_{gq_2}\). Suppose assumptions 1 holds. We have \(q_1 < q_2 \Rightarrow \hat{ε}_s(g, q_1) \geq \hat{ε}_s(g, q_2)\) and \(a^*_1 \leq a^*_2\).

I then investigate how corruption varies as the organizational structure changes. Comparing across networks is difficult because two networks can differ in a variety of ways. I conduct two comparisons. I characterize analytically the impact of a marginal change to an existing organization: adding a tie to the network. Second, I compare across a large number of organizational
structures using simulations.

Examining marginal changes to an existing organization 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 decreases 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. There is, however, a limit to what additional communication ties can achieve: the proof hinges on a graphical argument (lemma A2 in Appendix A.2) showing that an additional communication tie may at best create a coalition that is strictly smaller but has just as many witnesses as an existing coalition. Most additional ties, however, have no effect: because corruption only occurs in enclaves (proposition 2), ties may change behavior 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 $g$ ($g' = g + i \rightarrow j$), or a communication tie ($g' = g + ij$). We have

**Proposition 4** (Adding monitoring ties weakly decreases the frequency of corruption). Suppose $g' = g + i \rightarrow j$. Then $\hat{\epsilon}_s(g', q) \leq \hat{\epsilon}_s(g, q)$. For the inequality to hold strictly for some $q \in (0, 1)$, assumption 3 must hold, and it must be that $j \in c$, and $i \notin c \cup W_{cg}$ for some minimal coalition $c \in M_g$ and all coalitions essentially equal to $c$.

**Proposition 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 the inequality to hold strictly for some $q \in (0, 1)$, assumption 3 must hold, and it must be that there is $c^* \in M_g$ and $c' \in C_{g'}$ such that $a_{c'} > a_{c^*}$ and $w_{c'} = w_{c^*}$.

Using simulations, I show that more hierarchical organizations are more corrupt, although there is a lot of noise. Because enclaves are more corrupt, organizations that contain more enclaves are also more corrupt. More hierarchical organizations contain more enclaves than flatter ones, and are therefore more corrupt. To think of hierarchy, I use the network-analytic concept of modularity [Newman 2006], which captures the extent to which the graph can be divided in independent communities. In a graph whose nodes are partitioned in $n$ communities, modularity $m \in [-1, 1]$ is such that when $m = 0$, ties are distributed uniformly across and within communities. As $m$ increases, more ties fall within communities, and fewer across. More modular graphs represent more vertical organizations.

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. For seed $s$ on graph $g$, the AUC writes $AUC_{sg} \equiv \int_0^1 \hat{\epsilon}_s(g, q) dq$ (Figure 3). Its expected value is $AUC_g \equiv E(AUC_{sg})$, 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 is $O(2^N)$. As such, I consider a random sample of 1000 small graphs (16 nodes) in which communication and monitoring

\[ \text{Note that as in proposition 3, propositions 4 and 5 imply that when the frequency of corruption decreases, it selects on higher-scale projects.} \]
ties are collapsed. [Sah et al. (2014)] propose an algorithm that generates connected graphs of specified modularity and least depart from a random graph. I sample graphs of density of .26 with two equal-sized communities, and modularity ranging from 0 to .4. Simulations assume equal-sharing and use the following probability of success:

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

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 \( p(N, 0, q) = 1 \). Symmetrically, detection is certain when the organization is a witness, with \( p(1, N-1, q) = 0 \).

Figure 4 reveals a positive yet noisy correlation between modularity and AUC. The intuition is simple: more hierarchical organizations, as captured by more modular graphs, are more corrupt because communities are more enclaved (e.g., graph 2 in Figure 4). When the organization becomes flatter, as captured by decreasing modularity, communities disenclave. Reducing modularity, cross-community ties make communities more exposed to each other, which reduces corruption (e.g., graph 1 in Figure 4). Yet, Figure 4 reveals a lot of variation among equally modular organizations, showing that within equally hierarchical organizations, small details in organizational structure may have large consequences for the formation of enclaves.

Analytical results and simulations suggest how endogenous network formation may affect the results. Propositions 4 and 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 ties that decrease modularity are more likely to reduce corruption. Corrupt agents should then prefer adding ties...
within communities, while the social planner prefers adding ties across communities.

1.3 Extension to other division rules

Up to this point, the analysis considered the game $\Gamma_c(t_e)$ that assumes that accomplices are able to contract upon the distribution of the rent ex-ante, and that they are maximally egalitarian, since they divide the bribe equally. I consider here the other division rules: the lawless environment $\Gamma_l$ and, within the contractual environment, the monopoly rule $\Gamma_{c(t_m)}$, which are depicted graphically in Figure 5. This raises new interrogations: how would accomplices distribute the bribe endogenously? do different distributions of the rent prompt for different coalitions? The lawless environment allows a discussion of endogenous division of the surplus. Comparing equal-sharing to monopoly allows appraising the robustness of results to different distributions of bargaining power within the coalition, moving from maximally egalitarian to maximally unegalitarian accomplices. Furthermore, because corruption is most extractive under monopoly, this rule represents an efficient benchmark other rules can be evaluated against. In this extension, I assume a constant cost: $\epsilon_i = \epsilon$ for all $i \in N$.

Lawlessness introduces a commitment problem that benefits brokers. Operatives do not
recruit other nodes on equilibrium path (e.g. because they only have one neighbor, the one who recruited them). 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 is larger, to make them indifferent between the equilibrium schedule and these deviations: brokers extract more surplus than operatives because they have more outside options.

Figure 5 exemplifies the commitment problem. Here, node 3 is a broker because she recruits an operative, node 4. At the third history, the broker may have a profitable deviation in keeping her incoming transfer \( b + \frac{\epsilon}{p(3,1,q)} \) to herself. The seed needs to make her indifferent between hiring node 4 and not hiring anyone. While the operative gets \( \frac{\epsilon}{p(3,1,q)} \), the fraction of the rent that gives her a payoff of 0, the broker ends up with a higher share \( b \geq \frac{\epsilon}{p(3,1,q)} \). The following proposition generalizes the intuition:

**Proposition 6** (Brokers extract more surplus). In the lawless environment, if \( c \) is an equilibrium coalition, then \( u_i(c, g, q) \geq 0 \) for any \( i \in c \). If \( i \) is an operative, then \( u_i(c, g, q) = 0 \).

Moving further requires defining efficiency. The surplus is the expected benefit from the rent, net of the cost \( \epsilon \) paid by each accomplice. A coalition is efficient if it maximizes the surplus, and if such surplus is positive. A division rule is efficient if all of its equilibrium outcomes are efficient coalitions when there are any, and if the seed rejects the rent when there are none. Since the diffusion process requires that feasible coalitions \( C_g \) 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:

**Definition 3.** Let \( \Pi(c, g, q) = p(a_c, w_{cg}, q) - a_c \epsilon \) be the surplus of coalition \( c \) on graph \( g \) for monitoring level \( q \). A division rule is efficient if for any seed \( s \), all equilibrium outcomes are coalitions that solve \( \max_{c \in C_g} \Pi(c, g, q) \) when \( \Pi(c, g, q) \geq 0 \) for some \( c \in C_g \), and the seed rejects the rent in all equilibrium outcomes otherwise.

Making efficiency claims requires pinning down equilibrium. In all division rules, the seed has a threshold strategy:

**Proposition 7** (Threshold strategy, extension). Under the monopoly rule and in the lawless environment, the seed has a threshold \( \hat{\epsilon} > 0 \) such that she rejects the rent if \( \epsilon > \hat{\epsilon} \). Otherwise, some coalition in \( C_g \) is realized.

Knowing when the seed takes the rent allows discussing 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. 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
 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.

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:

Proposition 8 (Efficiency). The monopoly rule is efficient. Let \( \hat{\epsilon}_m, \hat{\epsilon}_e, \hat{\epsilon}_l \) be the seed’s thresholds under monopoly, equal-sharing, and lawlessness respectively. We have \( \hat{\epsilon}_m \geq \hat{\epsilon}_l \). Let \( C^m \) and \( C^e \) be sets of equilibrium coalitions under equal-sharing and monopoly for some \((q, \epsilon)\). Then

\[
\min_{c \in C^e} a_c \leq \min_{c \in C^m} a_c, \quad \text{and} \quad \max_{c \in C^e} a_c \leq \max_{c \in C^m} a_c.
\]

I finally examine the robustness of the findings to these new division rules. I show in Appendix A.3 that all the propositions derived under equal-sharing hold in the monopoly rule under similar assumptions. None of the findings travel to the lawless environment, because the cost of a coalition depends on the outside options of all 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.

Figure 6 considers these three division rules jointly to suggest how informal contracts affect cooperation among accomplices (Figure 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 Figure 6. The contracts considered here also align incentives within the coalition, and display the regularities discussed in the previous section. Consequently, in Figure 6, moving from the middle of the simplex to the left or to the right – that is, to perfectly egalitarian or inequalitarian contracts – increases efficiency. Although monopoly is more efficient than equal-sharing, both contracts show that informal contracts can be welfare-enhancing. As such, real corrupt agents
should try to implement some Pareto-improving contract. Contracts that are less binding or that implement distributions of the rent that are less able to align incentives should introduce inefficiencies and upset the regularities seen above.

2 Experiment

This section takes the model to the lab to test its main predictions. Specifically, we saw that as organizations adopt better monitoring technologies, corruption becomes less frequent but increases in scope and selects on high-scale projects (proposition 3). More isolated subgraphs form enclaves that are more corrupt, because they generate fewer witnesses (proposition 2). 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 4 and 5). Lawlessness introduces significant noise and inefficiencies to the way accomplices cooperate (proposition 8), that they would presumably partially solve using informal, self-enforcing contracts.

2.1 Design

In the experiment, subjects are put in groups of four players to play 12 repetitions of the diffusion game under lawlessness $\Gamma_l$ in various treatment conditions. I first describe the basic setup game 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.

2.1.1 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 $\epsilon$ discrete experimental units (EU) and assigned an index number and network positions through a network diagram drawn on a paper handout.

As in $\Gamma_l$, 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. The process ends when either all players are accomplices, or no offer can be made anymore. Then, non-accomplices each earn their endowment $\epsilon$. The enumerator randomly determines whether accomplices earn their share of the rent, using a paper handout listing the probability of success $p$ from equation 3 for all feasible coalitions 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 communicated, they would 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 the informal contracts corresponding to the contractual environment $\Gamma_c$.

### 2.1.2 Treatments and predictions

The main goal of the experiment is to test the main predictions of the model. To do so, I compare a baseline condition to a series of treatments summed up in Table 1. Within each condition, I vary the profitability of corruption by manipulating the endowment $\epsilon$ held by subjects, using either $\epsilon = 4$ EU (petty corruption), or $\epsilon = 2$ EU (grand corruption). An additional goal of the experiment is to examine an important model assumption: the division rule used by accomplices. To examine the division rule used by accomplices without compromising the main goal of testing the main model predictions, I picked experimental parameters such that all three division rules examined in the theoretical section yield the same equilibrium coalition for each treatment condition, and compare observed patterns of transfers to the three division rules.

The first row of Table 1 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 predictions, derived from proposition 3, 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 1 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

---

<table>
<thead>
<tr>
<th>Condition</th>
<th>Capacity ($q$)</th>
<th>Predicted equilibrium coalition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Grand corruption ($\epsilon = 2$)</td>
</tr>
<tr>
<td>Baseline</td>
<td>.1</td>
<td>(48 groups)</td>
</tr>
<tr>
<td>(4 repetitions)</td>
<td></td>
<td>(192 games)</td>
</tr>
<tr>
<td>Hard</td>
<td>.75</td>
<td></td>
</tr>
<tr>
<td>(4 repetitions)</td>
<td></td>
<td>(192 games)</td>
</tr>
<tr>
<td>Exposing tie</td>
<td>.1</td>
<td></td>
</tr>
<tr>
<td>(4 repetitions)</td>
<td></td>
<td>(184 games)</td>
</tr>
</tbody>
</table>

Table 1: **Experimental parameters and equilibria.** $s$ is the seed. Grey nodes are coalition members. Dotted ties are irrelevant ties, and are added and removed within each condition.
petty corruption.

Another series of predictions considered the impact of adding additional ties to the organization. Propositions 4 and 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:

**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 hold the monitoring technology constant but adds a tie that should decrease corruption. The third row of Table 1 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 dashed ties in Table 1. 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 cell, half the games include the irrelevant tie, and half do not.

The last theoretical prediction was that within an organizational structure, more isolated subgraphs form enclaves that are more corrupt, because they generate fewer witnesses (proposition 2):

**Hypothesis 4.** Enclaves 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.

### 2.1.3 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 2 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.

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.

---

16Due to power considerations, I did not test whether some ties may increase corruption (proposition 5).
Partners in crime

Romain Ferrali

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sample</th>
<th>Students</th>
<th>Employees</th>
<th>Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>age</td>
<td>24.85</td>
<td>20.44</td>
<td>26.17</td>
<td>5.73</td>
</tr>
<tr>
<td>% females</td>
<td>0.21</td>
<td>0.16</td>
<td>0.23</td>
<td>0.07</td>
</tr>
<tr>
<td>% secondary education</td>
<td>0.75</td>
<td>0.95</td>
<td>0.69</td>
<td>-0.20*</td>
</tr>
<tr>
<td>income</td>
<td>1.65</td>
<td>1.83</td>
<td>1.59</td>
<td>-0.23**</td>
</tr>
<tr>
<td>% urban</td>
<td>0.94</td>
<td>1.00</td>
<td>0.93</td>
<td>-0.07***</td>
</tr>
<tr>
<td>% Arabs</td>
<td>0.94</td>
<td>0.95</td>
<td>0.94</td>
<td>-0.01</td>
</tr>
<tr>
<td>risk-taking</td>
<td>2.78</td>
<td>2.59</td>
<td>2.83</td>
<td>0.25</td>
</tr>
<tr>
<td>altruism</td>
<td>1.54</td>
<td>1.83</td>
<td>1.45</td>
<td>-0.37**</td>
</tr>
<tr>
<td>extroversion</td>
<td>2.96</td>
<td>2.69</td>
<td>3.04</td>
<td>0.36**</td>
</tr>
<tr>
<td>N</td>
<td>272</td>
<td>63</td>
<td>209</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: 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 1 (introvert) to 4 (extrovert). Tests for differences in means use a t-test; *p<0.05; **p<0.01; ***p<0.001.

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. In Appendix B I show 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. Appendix B 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.

2.2 Results

2.2.1 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, the mean size of realized coalitions, and the distribution of these coalitions. 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.[17] 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.

Figure 7 considers acceptance behavior. The top panel reveals that behavior under grand
corruption matches predictions reasonably well: in all three treatment conditions, the seed ac-
cepts the rent at high rates (around 90 percent). When moving to petty corruption, acceptance
rates go in the same direction as the predictions: they remain high in the baseline, but drop
sharply in the hard and exposing tie treatments. The bottom panel examines differences in
mean acceptance rates across treatment conditions and confirms those intuitions. Acceptance
rates in the hard and exposing tie treatments under grand corruption and in the baseline under
petty corruption are hardly different from acceptance rates in the baseline under grand cor-
ruption. [18] Conversely, moving from grand to petty corruption in the hard and exposing tie
 treatments shows large, significant differences.[19] Together, results show support for hypotheses
1, 2, and 3: both the hard and exposing tie treatment decrease the frequency of corruption by
selecting on grand corruption (hypotheses 1 and 2), while irrelevant ties have no effect on the
frequency of corruption (hypothesis 3).

Figure 8 considers the mean size of realized coalitions under grand corruption. The top panel
shows that while results do not perfectly match predictions, they go in the expected direction:
the mean size of the coalition is smallest under the baseline (1.7 accomplices), increases in
the exposing tie treatment (2.1 accomplices), and is largest under the hard treatment (2.6
accomplices). The bottom panel confirms the intuition: differences in means are large and
highly significant.[20] This result supports hypothesis 1: the hard treatment increases the scope
of corruption. Note, however, that outcomes least align with predictions in this treatment, where
the mean size of the coalition is 2.6 accomplices, a far cry from the expected 4 accomplices. I
show in the next subsection that this is because forming larger coalitions poses more complex
backward induction problems to subjects.

Finally, Figure 9 examines the distribution of realized coalitions under grand corruption.
Observed behavior aligns very well with predictions under the baseline, where the equilibrium
coalition, represented by the grey bar, is realized in more than 70 percent cases. Similarly,
results align reasonably well with predictions in the exposing tie treatment, where seed favors the
equilibrium coalition – which accounts for 45 percent of realized coalitions – over other possible
coalitions. Specifically, the seed overwhelmingly favors the more enclave 2-people coalition

---

17 Models reported in Appendix B.2
18 Effect sizes range between 5 and 8 percentage points with p-values of, respectively, .04, .03, and .21
19 Effect sizes are of about 35 percentage points. The p-values of the associated F-tests are both smaller than
.001.
20 Both effect sizes are larger than .5 accomplices, and both p-values are smaller than .001.
### Differences in means

<table>
<thead>
<tr>
<th>Frequency of corruption: mean probability that the seed takes the rent</th>
</tr>
</thead>
<tbody>
<tr>
<td>baseline, grand</td>
</tr>
<tr>
<td>0.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Differences in means</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.4</td>
</tr>
</tbody>
</table>

- **Observed value**
- **Equilibrium prediction**

**Figure 7: Frequency of corruption.** Bars are semi-parametric bootstrapped 95 percent confidence interval clustered at the group level using 10,000 replicates. All estimates are constructed from model 1, Table B1 in Appendix B.2, save for the bottom estimate (irrelevant ties), constructed from model 2, Table B1. 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.

### Differences in mean size

<table>
<thead>
<tr>
<th>Scope of corruption: mean size of the coalition, grand corruption</th>
</tr>
</thead>
<tbody>
<tr>
<td>baseline, grand</td>
</tr>
<tr>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Differences in mean size</th>
</tr>
</thead>
<tbody>
<tr>
<td>exposing – baseline [grand]</td>
</tr>
<tr>
<td>0.0</td>
</tr>
</tbody>
</table>

- **Observed value**
- **Equilibrium prediction**

**Figure 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. All estimates are constructed from model 4, Table B1 in Appendix B.2. Conditional on corruption occurring, the realized coalition involves more accomplices in the hard treatment. Results are furthest away from predictions in the hard treatment.

### Distribution of realized coalitions under grand corruption

<table>
<thead>
<tr>
<th>Distribution of realized coalitions under grand corruption</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Baseline</td>
</tr>
</tbody>
</table>

- **Equilibrium prediction**

**Figure 9: Distribution of realized coalitions under grand corruption.** Labels are coalitions of black nodes. Enclaves are more corrupt (panel c, second vs. third bar). Realized coalitions overwhelmingly match the prediction, except in the hard treatment.
over the less enclaved 2-person coalition, which supports hypothesis [4] more enclaved coalitions are more likely to be corrupt. Consistently with results on coalition size, the distribution of realized coalitions least aligns with predictions in the hard treatment, where the equilibrium coalition only accounts for 28 percent of realized coalitions. I discuss reasons for this discrepancy in the next subsection. I also compare the distribution of realized coalitions with and without the irrelevant tie within-treatment using Fisher exact tests and find no significant difference, confirming that irrelevant ties have no effect on the scope of corruption (hypothesis [3]).

Finally, I find little evidence of heterogeneous treatment effects (see Appendix B.3). Using multi-level specifications with random effects at the individual and group level, I show that there is little group- and individual-level heterogeneity, and that group-level heterogeneity is larger than individual-level heterogeneity. I also show that 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.

### 2.2.2 Division rule

Having shown support for the main theoretical predictions, I now investigate the division rule used by participants. I consider the division rules examined in section 1.3: bargaining in a lawless environment ($\Gamma_1$), and, within a contractual environment, equal-sharing and monopoly ($\Gamma_c(t_e)$ and $\Gamma_c(t_m)$, respectively). The design makes characterizing the division rule difficult. 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. Furthermore, because treatments were designed to yield similar outcomes irrespective of the rule, predictions are often identical across rules. I show moderate support for monopoly and bargaining, but mostly identify a deviation from rational behavior that has a substantive implication: recipients accept transfers that give them negative surplus early on in the diffusion sequence, because they do not internalize their future transfers.

To assess which division rule best matches the data, Figure 10 examines the seed’s payoff in

---

21 Results reported in Appendix B.2
Partners in crime

Romain Ferrali

<table>
<thead>
<tr>
<th>did share</th>
<th>not share</th>
</tr>
</thead>
<tbody>
<tr>
<td>share</td>
<td>.07</td>
</tr>
<tr>
<td>not share</td>
<td>.42</td>
</tr>
</tbody>
</table>

(a) Sender’s decision

<table>
<thead>
<tr>
<th>did accept</th>
<th>reject</th>
</tr>
</thead>
<tbody>
<tr>
<td>accept</td>
<td>.24</td>
</tr>
<tr>
<td>reject</td>
<td>.01</td>
</tr>
</tbody>
</table>

(b) Recipient’s decision

Table 3: Learning effects for distribution of errors. Numbers denote observed frequencies. Italicized cells denote errors.

treatments where a multi-player coalition was the equilibrium outcome (hard and exposing tie treatments under grand corruption) and was realized. As off-path behavior may not be well-defined in the contractual environment, this approach restricts the analysis to on-path behavior. Only the exposing tie treatment, reported in panel (b), has different predictions for different division rules. The figure shows that in this treatment, the observed distributions align more closely with the prediction under bargaining and equal-sharing than with the prediction under equal-sharing.

More importantly, Figure [10] reveals a systematic deviation from rational behavior: under both treatments, the mass of the distribution lies to the right of the prediction under monopoly. Specifically, the seed’s payoff is strictly above the monopoly allocation in 72 and 40 percent cases in the hard and exposing tie treatments respectively. Yet, that the seed has higher payoff than under monopoly implies that her accomplices end up with negative surplus. In other words, non-seed accomplices accept offers they should reject, or make offers when they should not.

To further examine this deviation from rational behavior, I focus on deviations from bargaining under lawlessness. The previous exercise allowed comparing division rules but was restrictive: it considered a few treatments, focused on the outcome of the diffusion process, and discarded outcomes that failed to match the equilibrium. Under bargaining, off-path behavior is always well-defined, which allows pinning down best responses at any history in a given game. This allows considering deviations from rational behavior at any history; that is, in the entire sequence of offers and responses that led to the outcome observed at the terminal history. Note that under bargaining, both senders and recipients have threshold strategies: sender $i$ should make transfer offers if her holdings $t_i$ are above some threshold $t_{i}^{*}$. Similarly, recipient $j$ should accept transfer $t_{ij}$ from $i$ if it is above some threshold $t_{ij}^{*}$. Comparing observed transfers to those thresholds indicates whether an offer is greedy and leaves recipient $j$ with negative surplus ($t_{ji} - t_{ij}^{*} < 0$), or generous and leaves $j$ with positive surplus ($t_{ji} - t_{ij}^{*} \geq 0$).

Table 3 examines binary decisions – whether a sender makes any offer, and whether a recipient accepts said offer, – and shows that the deviation observed in Figure [10] holds more generally: in all treatments, subjects over-accept offers, and over-share their holdings. The left panel of Table 3 focuses on senders and examines sharing decisions: does sender $i$ make transfers when she should (i.e. when $t_{i}^{*} \geq t_{i}$)? The right panel focuses on recipients and examines transfer acceptance does recipient $j$ accept transfers when she should (i.e. when $t_{ij} \geq t_{ij}^{*}$)? Table 3 shows that for both senders and recipients, about 50 percent decisions are errors. Furthermore, 22For 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 Figure [7] in the previous subsection.
Figure 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.

virtually all errors are false positives: when behaving irrationally, senders tend to make offers when they should not (over-sharing) and consistently, recipients accept offers they should reject (over-acceptance).

Figure 11 examines amounts offered over time, and shows that deviations from rational behavior attenuate in later histories. The left panel 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. 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.

That deviations attenuate at later histories is consistent with the commonly observed fact that backward induction problems are cognitively taxing, especially at early histories (Johnson et al. 2002; Spenkuch, Montagnes and Magleby 2018). At early histories, offerers and recipients both fail to internalize the future transfers required to realize equilibrium. They make and accept offers that only seem generous because they do not take into account these future transfers. At later histories, there are fewer steps to backward-induce, making the problem easier. Recipients are better able to identify greedy offers and accordingly, offers adjust to get closer to the equilibrium prediction.

Problems with backward induction explain why results are furthest away from predictions in the hard treatment, which has substantive implications for corruption under better monitoring. Results are furthest away from predictions in the hard treatment because this treatment
prompts for the largest equilibrium coalition, which poses a more complex induction problem to participants. This suggests an additional reason why better monitoring makes corruption less frequent: better monitoring prompts for larger coalitions, which are more difficult to form for they require that agents solve harder backward induction problems.

3 Conclusion

We began by highlighting that extant approaches to corruption face difficulties answering two questions: when do large networks of corrupt bureaucrats 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 enclaves (proposition 2). The concept shifts the unit of analysis from the individual to the coalition, and reconciles mixed findings about the structure of criminal networks: Aven (2015) and Morselli, Giguère and Petit (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, Kim and Lu 2015). Collectively, enclaves 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.

Enclaves 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, Mookherjee and Reichelstein 1995). Allowing to move beyond contrasting perfect hierarchies to perfectly flat organizations, the proposed approach shows that while the insight holds true at the aggregate level, there is much variation across organizations (see simulation results in section 1.2.2). 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 (propositions 4 and 5).

Findings on monitoring technologies (proposition 3) 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 strong institutions, recruiting accomplices is more desirable. Yet, only grand corruption is profitable enough to afford their additional protection, leading to the disappearance of petty corruption. The lab reveals a behavioral trait that strengthens the result: forming large coalitions poses a challenging backward induction problem, making corruption all the more unlikely under strong
Finally, the results suggest an important policy recommendation: organizational responses to corruption 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 two 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 enclaves without creating new ones.

Yet, coming to definitive policy recommendations requires further research. Although this paper opens 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, and assumes away transaction costs. Transaction costs would make the model more realistic. They would also make forming larger coalitions disproportionately more costly, since those require more transactions, which would likely strengthen the result that corruption decreases under better monitoring. Considering a repeated game would strengthen existing results on informal institutions, allowing to consider a setting where informal contracts emerge endogenously. Finally, introducing incomplete information, alongside with strategic reporting of corruption and heterogeneity through incorruptible types and/or strong ties (e.g., coethnicity) 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 1.3 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 extension with strong ties, another interesting 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.
References


Partners in crime

Romain Ferrali


Dahström, Carl, Jan Teorell, Stefan Dahlberg, Felix Hartmann, Annika Lindberg and Marina Nistotskaya. 2015. “The QoG Expert Survey Dataset II.”


International Monetary Fund - Staff Team from the Fiscal Affairs Department and the Legal Department. 2016. Corruption: Costs and mitigating strategies - IMF Staff Discussion Note no. SDN/16/05. Technical report.


Appendices

Contents

A  Proofs 35
A.1 Section 1.2.1 35
A.2 Section 1.2.2 36
A.3 Section 1.3 38
A.4 Extension: cost of corruption as expected loss 43
A.5 Extension: detection as a function of the scale of corruption 45
A.6 Simulations and experiment 48

B  Additional experimental results 49
B.1 Details about the protocol 49
B.2 Main experimental results 52
B.3 Power analysis 52
B.4 Learning and pooling effects 54
B.5 Individual-level characteristics 57
B.6 Location 63
B.7 Enumerators 63
B.8 Subjects 63
B.9 Prompts and material 64

C  Additional supportive evidence 65
C.1 Cross-country comparisons 66
C.2 US-India comparison 66

List of Figures

B1 Experimental protocol 50
B2 Experiment, comprehension over time 51
B3 Experiment, post-hoc power analysis 54
B4 Experiment, learning effects 56
B5 Experiment, random effects 59
B6 Experiment, students vs. employees 60
B7 Experiment, example comprehension question 64

List of Tables

B1 Experiment, models used in the main text 52
B2 Experiment, Fisher exact tests for the distribution of realized coalitions 52
B3 Experiment, hypotheses for power analysis 53
B4 Experiment, learning effects for distribution of errors 57
Partners in crime

B5 Experiment, pooling effects for distribution of errors ........................................ 57
B6 Experiment, Fisher exact tests for learning and pooling effects on distribution of
   errors ..................................................................................................................... 57
B7 Experiment, learning effects for main hypotheses ................................................... 58
B8 Experiment, pooling effects for main hypotheses ................................................... 58
B9 Experiment, random effects ................................................................................. 61
B10 Experiment, students vs. employees .................................................................. 62
B11 Experiment, probability of success ................................................................... 64
C1 Cross-country comparisons, models .................................................................... 66
C2 India vs. US, descriptive statistics ..................................................................... 67
C3 India vs. US, models ............................................................................................ 67
A Proofs

A.1 Section 1.2.1

Proof of lemma 7 Suppose \( u(c^*, g, q) < \epsilon_s \). Then, no coalition gives \( s \) a positive payoff. She rejects the rent.

Suppose \( u(c^*, g, q) \geq \epsilon_s \). Since \( c^* \in C_g \), there is at least one strategy profile that has \( c^* \) as an outcome. Since all utility functions are the same up to the constant \( \epsilon_i \), \( c^* \) maximizes the utility of any accomplice \( i \) in \( c^* \). Because \( \epsilon_i < \epsilon_s \), we have \( u_i(c^*, g, q) \geq 0 \). 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 \( C_g^* \) 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, \( \epsilon_s(g, q) = u(c^*, g, q) \in (0, 1) \).

Proof of proposition 7 Consider two essentially different coalitions \( c_1, c_2 \in C \) 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 \( u_1 = u_s(c_1, g_1, q) \) and \( u_2 = u_s(c_2, g_2, q) \) 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:

\[
\nabla_u u_2 - u_1 \equiv \frac{\partial u_2 - u_1}{\partial \epsilon_s} x_{\epsilon_s} + \frac{\partial u_2 - u_1}{\partial q} x_q = \left( \frac{\partial p_2}{\partial q} / a_{c_2} - \frac{\partial p_1}{\partial q} / a_{c_1} \right) x_q
\]

where \( x \equiv (x_{\epsilon_s}, x_q) \) is a unit-length vector. If the equation \( \nabla_u u_2 - u_1 = 0 \) has a finite number of solutions in \( x \), then \( U \) has measure 0. Assumption 7 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 lemma 3 Suppose that assumption 2 holds. Because \( w_2 \leq w_1 \), we have \( v(a_2, w_2, q) \geq v(a_2, w_1, q) \), which implies \( \max\{v(a_1, w_1, q), v(a_2, w_1, q)\} \leq \max\{v(a_1, w_1, q), v(a_2, w_2, q)\} \). As such, it suffices to show that \( v(a_c, w_c, q) \leq \max\{v(a_1, w_1, q), v(a_2, w_1, q)\} \).

Suppose that the growth rate of \( p \) in \( a \) is bounded by \( 1/a \); that is, either \( \frac{v(a_{w_1, q}) - p(a, w, q)}{p(a, w, q)} \leq \frac{1}{a} \) (case 1) or \( \frac{v(a_{w_1, q}) - p(a, w, q)}{p(a, w, q)} \geq \frac{1}{a} \) (case 2). Then \( v \) is monotonic in \( a \). Indeed, rearranging the first case gives \( \frac{p(a_{w_1, q})}{a} \leq \frac{p(a, w, q)}{a+1} \Leftrightarrow v(a + 1, w, q) \leq v(a, w, q) \). Similarly, rearranging the second case gives \( v(a + 1, w, q) \geq v(a, w, q) \). Since \( v \) is monotonic, we have \( v(a_c, w_1, q) \leq \max\{v(a_1, w_1, q), v(a_2, w_1, q)\} \). Note that \( v(a_c, w_c, q) < v(a_c, w_1, q) \), so \( v(a_c, w_c, q) < \max\{v(a_1, w_1, q), v(a_2, w_1, q)\} \).
Suppose that $p$ satisfies
\[ a[v(a^*, w_1, q) - \max\{v(1, w_1, q), v(N, w_1, q)\}] < |p(a, w_1 + 1, q) - p(a, w_1, q)|, \]
where $a^* \in \arg\max_{u \in \{1, \ldots, N\}} v(a, w_1, q)$. Dividing both sides by $a$ gives
\[ v(a^*, w_1, q) - \max\{v(1, w_1, q), v(N, w_1, q)\} < |v(a, w_1 + 1, q) - v(a, w_1, q)|, \quad (A2) \]

Since $w_c > w_1$, then $v(a_c, w_c, q) \leq v(a_c, w_1 + 1, q)$. It suffices to show that $v(a_c, w_1 + 1, q) < \max\{v(a_c, w_1, q), v(a_c, w_1, q)\}$. Because $v$ is quasi-concave, it is non-decreasing from $a = 1$ to $a = a^*$, and non-increasing from $a = a^*$ to $a = N$. As a result, with $1 \leq a_1 \leq a_c \leq a_2 \leq N$, we have
\[ v(a_c, w_1, q) - \max\{v(a_1, w_1, q), v(a_2, w_1, q)\} \leq v(a^*, w_1, q) - \max\{v(1, w_1, q), v(N, w_1, q)\} \]
Combining this with equation $\text{(A2)}$, we get
\[ v(a_c, w_1, q) - \max\{v(a_1, w_1, q), v(a_2, w_1, q)\} < |v(a, w_1 + 1, q) - v(a, w_1, q)|. \]

Note that $v(a_c, w_1 + 1, q) = v(a_c, w_1, q) - |v(a_c, w_1 + 1, q) - v(a_c, w_1, q)|$. Rearranging the above inequality and substituting yields $v(a_c, w_1 + 1, q) < \max\{v(a_1, w_1, q), v(a_2, w_1, q)\}$. \hfill \square

**Proof of proposition 2.** I show the contrapositive of $c' \in C^*_g \Rightarrow c' \in M_g$. Suppose $c' \notin M_g$. Then there is $c \in C$ such that $a_c \leq a_{c'}$ and $w_c < w_{c'}$. Suppose $a_{c'} = a_c$, then $v(a_c, w_c, q) > v(a_{c'}, w_{c'}, q)$. Suppose $a_c < a_{c'}$. Then by lemma $2$ $v(a_{c'}, w_{c'}, q) < \max\{v(a_c, w_c, q), v(N, 0, q)\}$ for any $q \in (0, 1)$. \hfill \square

### A.2 Section 1.2.2

**Proof of proposition 3.** 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 $1$ if $c^* \in C^*_{g_1}$, then $\hat{\epsilon}_s(g, q) = u(c^*, g, q)$. We have $u(c_1^*, g, q_1) \geq u(c_2^*, g, q_1)$. Since for a given coalition, $u$ is decreasing in $q$, we have $u(c_2^*, g, q_1) \geq u(c_2^*, g, q_2)$. This implies $u(c_1^*, g, q_1) \geq u(c_2^*, g, q_2)$: that is, $\hat{\epsilon}_s(g, q_1) \geq \hat{\epsilon}_s(g, q_2)$.

I now show that $q_1 < q_2 \Rightarrow a_{c_1^*} \leq a_{c_2^*}$. Let $c_1^*$ be the largest coalition in $C^*_{g_1}$, and $c_2^*$ the smallest in $C^*_{g_2}$, with sizes $a_{c_1^*}$ and $a_{c_2^*}$. To prove the claim, if suffices to show that $a_{c_1^*} \leq a_{c_2^*}$.

Suppose not. Because $c_1^* \in C^*_{g_1}$ and $c_2^* \in C^*_{g_2}$, we have $u(c_2^*, g, q_1) - u(c_1^*, g, q_1) \leq 0$ and $u(c_2^*, g, q_2) - u(c_1^*, g, q_2) \geq 0$. Since $u(c_2^*, g, q) - u(c_1^*, g, q)$ is continuous in $q$, there must be some $q \in [q_1, q_2]$ such that $u(c_2^*, g, q) = u(c_1^*, g, q)$. Because $a_{c_1^*} > a_{c_2^*}$, assumption $1$ implies that $\frac{\partial}{\partial q} [u(c_2^*, g, q) - u(c_1^*, g, q)] < 0$. Since $u(c_2^*, g, q_1) - u(c_1^*, g, q_1) \leq 0$, then $u(c_2^*, g, q_2) - u(c_1^*, g, q_2) < 0$, a contradiction. \hfill \square

**Lemma A1 (Old coalitions are weakly dominated).** We have $C_g = C_{g'}$ if $g' = g + i \rightarrow j$ and
Lemma 2 implies essentially equal to the proposition. Proposition 2 implies that for any $q \in c$ and $i \notin c \cup W_{cg}$

$$w_{cg'} = \begin{cases} w_{cg} + 1, & \text{if } g' = g + i \rightarrow j \text{ and } j \in c \\
w_{cg} & \text{otherwise.} \end{cases} \quad (A3)$$

Proof. The proof is immediate.

Lemma A2 (New coalitions are weakly dominated). For any coalition $c' \in C_{g'} \setminus C_g$, there is $c \in C_g$ such that $a_c < a_{c'}$ and $w_{cg} \leq w_{c'g'}$.

Proof. By lemma A1 we only need to consider adding communication ties, since $g' = g + i \rightarrow j$ implies $C_{g'} = C_g$. Since $c'$ is not feasible on $g$, there exists at least one node $k \in c'$ such that the tie $ij$ is on all paths between $k$ and $s$ in $g_u'$ such that all nodes on that path are in $c'$. Let $c'_{nf}$ be the set of such nodes. Its complement, $c' = c' \setminus c'_{nf}$ is feasible on $g$. I show that $c' = c$. By construction, we have $a_{c'} < a_{c'}$. Note that $i \in c' \iff j \in c_{nf}$. Without loss of generality, suppose that $i \in c'$. Let $c_{g'} = (W_{cg'} \setminus c_{nf}) \cup W_{c_{nf}g'}$, which implies

$$w_{cg'} = |W_{cg'}| = |W_{c'g'} \setminus c_{nf}| + |W_{c_{nf}g'}|\quad (A4)$$

By construction, $W_{c'g'} \setminus c_{nf} = W_{c'g'}$. Indeed, if node $k \in W_{c'g'} \cap c_{nf}$, then $k$ is the neighbor of some $l \in c'$ on $g_u'$. So the path $S, \ldots, l, k$ is such that all nodes between $s$ and $k$ are in $c'$, and does not contain $ij$ which implies, by definition of $c'$, that $k \in c'$. Lemma A1 implies that $|W_{c'g'}| = w_{c'g'}$. Also note that witnesses of coalition $c_{nf}$ cannot be in $c_{nf}$: $c_{nf} \cap W_{c_{nf}g'} = \emptyset$. As such, $(W_{c'g'} \setminus c_{nf}) \cup W_{c_{nf}g'} = W_{c'g'} \setminus W_{c_{nf}g'}$. Plugging back into equation A4, we get

$$w_{cg'} = w_{c'g} + |W_{c_{nf}g'}|\quad (A5)$$

Proof of proposition 2. By lemma A1 we have $C_{g'} = C_g$ and for any $c \in C_g$, $w_{cg} \leq w_{c'g'} \leq w_{cg} + 1$. This implies $u(c, g', q) \leq u(c, g, q)$. So $\hat{c}_s(g', q) \leq \hat{c}_s(g, q)$. Let’s show the second part of the proposition. Proposition 2 implies that for any $q \in (0, 1)$, there is $c^* \in M_g$ that is realized in equilibrium on $g$. I show the contrapositive. Suppose that for all $c^* \in M_g$, there is some $c \in M_g$ essentially equal to $c^*$ such that $j \notin c$ or $i \in c \cup W_{cg}$. Then by lemma A1, $c$ on $g'$ is essentially equal to $c^*$ on $g$. As such, $u(c, g', q) = u(c^*, g, q)$.

Proof of proposition 3. By lemma A1 we have that for any $c \in C_g$, $u(c, g, q) = u(c, g', q)$, which implies that $\hat{c}_s(g', q) \geq \hat{c}_s(g, q)$. Let’s show the second part of the proposition. Suppose that there is $q \in (0, 1)$ such that $\hat{c}_s(g', q) > \hat{c}_s(g, q)$. Using proposition 2, this means that there is $c \in M_g, c' \in M_{g'}$ such that $u(c', g, q) < u(c, g, q)$. For this to be true, it must be that $c' \in C_g \setminus C_{g'}$ for otherwise, lemma A1 implies $u(c', g, q) = u(c', g', q)$. So lemma A2 implies that there is $c^* \in C_g$ such that $a_{c'} < a_{c'}, w_{c^*g} \leq w_{c'g'}$. It must be that $w_{c^*g} = w_{c'g'}$ for otherwise, lemma 2 implies $u(c', g', q) < \max\{u(c^*, g, q), v(N, 0, q)\}$. Suppose $c^* \notin M_g$. Then there is $\bar{c} \in C_g$ such that $a_{\bar{c}} \leq a_c$ and $w_{\bar{c}g} < w_{cg}$. This implies $a_{\bar{c}} \leq a_{c'}$ and $w_{\bar{c}g} < w_{c'g}$. Then lemma 2 implies $u(c', g', q) < \max\{u(\bar{c}, g, q), v(N, 0, q)\}$, a contradiction.
A.3 Section 1.3

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

**Proof of proposition 6.** Suppose coalition \( c \in C_g \) is an equilibrium outcome for some \( (\epsilon, q) \in (0, 1)^2 \). Then it must be that \( u_i(c, g, q) \geq 0 \) 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 \( u_i(c, g, q) > 0 \), then \( j \) has an incentive to deviate and set \( t_{ji} \) such that \( u_i(c, g, q) = 0 \).

Before proving the next propositions, we pin down equilibrium under lawlessness and monopoly. There division rules create multiple equilibria, some of which arise from uninteresting resolution of indifference conditions. In equilibrium, an accomplice is indifferent between her broker’s favorite outcome and her outside option. There is an equilibrium in which she picks her outside option. To rule out this case, I consider equilibria that satisfy deference; that is, equilibria where indifferent nodes defer to their broker’s preference. Formally:

**Definition A1.** A strategy profile \( \sigma \) satisfies deference if, whenever some node \( i \) that responds to an offer from \( j \) is indifferent between two actions, and there are nodes on the path of accepted offers from the seed to \( j \) that are not, then \( i \) takes the action that is preferred by her closest such node on that path.

All the proofs in this subsection only consider equilibria that satisfy deference. They also denote by \( G \) the set of connected multiplex graphs with ties of communication and monitoring ties (as defined in section 1.1) that can be formed with \( N \) nodes.

**Lemma A3.** Under lawlessness, in equilibrium, \( s \) rejects the rent if \( \max_{c \in \tilde{C}} u^l(c, g, q, \epsilon) \equiv p(a_c, w_{cg}, q) - \tau(c, g, q, \epsilon)\epsilon < 0 \), for some \( \tilde{C} \subseteq C_g \) and some \( \tau : \tilde{C} \times G \times (0, 1)^2 \rightarrow \mathbb{R}^+ \) that satisfies \( \tau(c, g, q, \epsilon) \geq a_c \), and \( \frac{\partial \tau}{\partial \epsilon} = 0 \). Otherwise, \( s \) accepts the rent and some coalition in \( C^l_{qge} \equiv \arg \max_{c \in \tilde{C}} u^l(c, g, q, \epsilon) \) is realized.

**Proof.** This proof requires a more specific definition of operators 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 operators and brokers of level \( n' < n \).

In equilibrium, if \( i \) moves at history \( h \), there is a mapping between any of her transfers \( t_i \) and all outcome coalitions \( \tilde{C}_{ih} \subset C_g \) that can be formed from that history. I prove the following:

**Lemma A4.** Suppose level-\( n \) broker moves at history \( h \) after transfer \( t_{ji} \). In equilibrium, \( i \) rejects \( t_{ji} \) if \( \max_{c \in \tilde{C}_{ih}} t_{ji}p(a_c, w_{cg}, q) - \tau_i(c, g, q, \epsilon)\epsilon \geq 0 \) for some \( \tau_i : \tilde{C}_{ih} \times G \times (0, 1)^2 \rightarrow \mathbb{R}^+ \) that satisfies \( \tau_i(c, g, q, \epsilon) \geq a_i^k + 1 \), where \( a_i^k \) 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 \(\arg\max_{c \in \tilde{C}_{ih}} t_{ji}p(a_c, w_{cg}, q) - \tau_i(c, g, q, \epsilon)\) 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 \(c \in \tilde{C}_{ih}\) 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 \(u_i(c, g, q, \epsilon) = (t_{ji} - a^i \frac{\epsilon}{p(a_c, w_{cg}, q)})p(a_c, w_{cg}, q) - \epsilon = t_{ji}p(a_c, w_{cg}, q) - (a^i + 1)\epsilon\). Setting \(\tau_i(c, g, q, \epsilon) = a^i + 1\) proves the claim. We have \(\frac{\partial \tau_i + 1}{\partial \epsilon} = 0\). In equilibrium, \(i\) rejects \(t_{ji}\) if \(\max_{c \in \tilde{C}_{ih}} u_i(c, g, q, \epsilon) < 0\). Otherwise, \(i\) accepts and makes the transfers that realize some coalition in \(\arg\max_{c \in \tilde{C}_{ih}} u_i(c, g, q, \epsilon)\).

Suppose \(i\) is a level-\(n\) broker. In equilibrium, her transfers are the cheapest vector of transfers that realize the coalitions in \(\tilde{C}_{ih}\). 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, suppose \(t_{ik}\) 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, \(t_{ik} = \frac{\tau_k(c, g, q, \epsilon)}{p(a_c, w_{cg}, q)}\). Suppose \(k\)'s best outside option is to accept and make some other transfer resulting in coalition \(c'\). Then \(t_{ik}\) solves \(t_{ik}p(c, g, q, \epsilon) - \tau_k(c, g, q, \epsilon) = t_{ik}p(c', g, q) - \tau_k(c', g, q, \epsilon)\), which gives \(t_{ik} = \frac{\tau_k(c, g, q, \epsilon) - \tau_k(c', g, q, \epsilon)}{p(a_c, w_{cg}, q) - p(c', g, q)}\) if \(k \in c\). In equilibrium, \(i\)'s payoff from \(c \in \tilde{C}_{ih}\) is \(u_i(c, g, q, \epsilon) = (t_{ji} - \sum_k t_{ik})p(a_c, w_{cg}, q) - \epsilon = t_{ji}p(a_c, w_{cg}, q) - (1 + \sum_k t_{ik})p(a_c, w_{cg}, q)/\epsilon\). Setting \(\tau_k(c, g, q, \epsilon) = 1 + \sum_k t_{ik}p(a_c, w_{cg}, q)/\epsilon\) proves the claim. Replacing \(t_{ik}\) by their equilibrium values and using the inductive hypothesis on \(\tau_k\), it is easy to show that \(\frac{\partial t_{ik}}{\partial \epsilon}p(a_c, w_{cg}, q)/\epsilon = 0\), which implies \(\frac{\partial \tau_k}{\partial \epsilon} = 0\). Furthermore, in equilibrium, any transfer \(t_i\) must make all the accomplices in \(c\) hired in transfers using money 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 \(\tau_k(c, g, q, \epsilon) \geq a^i + 1\).

To prove the claim, use lemma A4 and set \(i \equiv s\), \(\tau \equiv \tau_i\), define \(t_{ji} \equiv 1\), and note that \(a_c = a^s + 1, C_{sh} = \tilde{C}\).

**Lemma A5.** Under monopoly, in equilibrium, \(s\) rejects the rent if \(\max_{c \in C_g} u^m(c, g, q, \epsilon) = p(a_c, w_{cg}, q) - a_c \epsilon < 0\). Otherwise, \(s\) accepts the rent and some coalition in \(C_{gq}^m \equiv \arg\max_{c \in C_g} u^m(c, g, q, \epsilon)\) 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 \(u^m(c, g, q, \epsilon) = t_m(s_c)p(a_c, w_{cg}, q) - \epsilon = p(a_c, w_{cg}, q) - a_c \epsilon\). Consider \(c^m \in C_{gq}^m\). If \(u^m(c^m, g, q, \epsilon) \geq 0\), 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 \(s\) consider equilibria with deference, a coalition \(c \notin C_{gq}^m\) cannot be an equilibrium outcome. If \(u^m(c^m, g, q, \epsilon) < 0\), then the seed rejects the rent.

**Proof of proposition 7** I first consider monopoly. From lemma A5, the seed accepts the rent and a coalition in \(C_{gq} \subseteq C_g\) is realized whenever \(\max_{c \in C_g} p(a_c, w_{cg}, q) - a_c \epsilon \geq 0\). Otherwise, the seed rejects the rent. Therefore, we have \(\epsilon = \max_{c \in C_g} \frac{p(a_c, w_{cg}, q)}{a_c} > 0\). Now consider lawlessness. From lemma A3 the seed accepts the rent and a coalition in \(\tilde{C} \subseteq C_g\) is realized whenever \(\max_{c \in C} p(a_c, w_{cg}, q) - \tau(c, g, q, \epsilon) \geq 0\). Otherwise, the seed rejects the rent. Therefore,
we have $\hat{c} = \max_{c \in C_g} \frac{p(a, w_c, q)}{\tau(c, g, q, \epsilon)}$. Since $p(a, w_c, q) > 0$ and $\tau(c, g, q, \epsilon) \geq a_c > 0$, it must be that $\hat{c} > 0$.

**Proof of proposition 8.** Showing that monopoly is efficient is a direct corollary of lemma A5 if $\epsilon \leq \hat{c}^m$, then all equilibrium outcomes are efficient, since they solve $\max_{c \in C_g} u^m(c, g, q, \epsilon)$. Conversely, if $\epsilon > \hat{c}^m$, then no coalition yields a positive payoff. Corruption is inefficient, and the seed rejects the rent.

That $\hat{c}^m = \hat{c}^e \geq \hat{c}^l$ follows from lemma 1 and the proof of proposition 7; we have $\hat{c}^m = \hat{c}^e = \max_{c \in C_g} \frac{p(c, g, q)}{a_c}$, while $\hat{c}^l = \max_{c \in C_g} \frac{p(c, g, q)}{\tau(c, g, q, \epsilon)}$. Lemma A3 tells us that $\tau(c, g, q, \epsilon) \geq a_c$, which implies $\hat{c}^m \geq \hat{c}^l$.

Let’s show that for any $(g, \epsilon)$, $\min_{c \in C^e} a_c \leq \min_{c \in C^m} a_c$, and $\max_{c \in C^e} a_c \leq \max_{c \in C^m} a_c$. Lemmas 1 and A5 tell us that the sets of equilibrium coalitions under equal-sharing and monopoly are, respectively, $C^e = C_{gq}^e = \arg \max_{c \in C_g} u^e(c, g, q, \epsilon)$, and $C^m = C_{gq}^m = \arg \max_{c \in C_g} u^m(c, g, q, \epsilon)$. Note that $u^e(c, g, q, \epsilon) = u^m(c, g, q, \epsilon) \iff \epsilon = p(c, g, q)/a_c$. So when $\epsilon = \hat{c}^m = \max_{c \in C_g} p(c, g, q)/a_c$, we have $C_{gq}^e = C_{gq}^m$. The claim is trivially true.

Consider the case where $\epsilon < \hat{c}^m$. Lemma 1 tells us that $C_{gq}^e$ does not vary with $\epsilon$. Conversely, the following lemma shows that as $\epsilon$ decreases, the coalitions in $C_{gq}^m$ get larger. Using this lemma, it is immediate that $\min_{c \in C^e} a_c \leq \min_{c \in C^m} a_c$, and $\max_{c \in C^e} a_c \leq \max_{c \in C^m} a_c$.

**Lemma A6.** Under monopoly, let $c_1^m \in C_{gq}^m$ and $c_2^m \in C_{gq}^m$. We have $c_1 < c_2 \Rightarrow a_{c_1}^m \geq a_{c_2}^m$.

**Proof.** Let $c_1$ be the smallest coalition in $C_{gq}^m$, and $c_2$ the largest in $C_{gq}^m$ 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, if suffices to show that $a_1 \geq a_2$. Suppose not. Because $c_1 \in C_{gq}^m$ and $c_2 \in C_{gq}^m$, we have $u^m(c_2, g, q, \epsilon) - u^m(c_1, g, q, \epsilon) \leq 0$ and $u^m(c_2, g, q, \epsilon) - u^m(c_1, g, q, \epsilon) \geq 0$. 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, $\frac{\partial}{\partial q} [u^m(c_2, g, q) - u^m(c_1, g, q)] = \alpha_1 - \alpha_2 < 0$, since $a_1 < a_2$. Since $u^m(c_2, g, q, \epsilon) - u^m(c_1, g, q, \epsilon) \geq 0$, then $u^m(c_2, g, q, \epsilon) - u^m(c_1, g, q, \epsilon) \leq 0$, a contradiction.

**A.3.1 Robustness of findings under monopoly**

This subsection shows that findings under equal-sharing travel to monopoly. Again, this subsection only considers equilibria that satisfy deference (see definition A1). Specifically, I show that lemmas 1 and 2 and propositions 1 to 5 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, $C_{gq}^*$ now varies with $\epsilon$.

Assumption 1 and 2 are replaced by qualitatively similar assumptions. The new assumption still requires the ratio of partial derivatives of $p$ with respect to $q$ for any two coalitions to be bounded above by some quantity greater than 1. The new assumption 2 requires either the variation rate of $p$ in $a$ to be bounded (a regularity condition that ensures that $v$ is monotonic in $a$), or that $p$ is sufficiently decreasing in $w$. Similar to equal-sharing, define $v^m(a, w, q, \epsilon) \equiv \frac{p(a, w, q)}{a} - [1 - p(a, w, q)]$ the valuation of a coalition with $a$ accomplices, $w$ witnesses for monitoring $q$ and cost $\epsilon$. Assumptions 1 and 2 become, respectively:
**Assumption A1.** [Assumption 1 under extension with monopoly] 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 \( \frac{\partial p(a_2, w_2, q)}{\partial q} / \frac{\partial p(a_1, w_1, q)}{\partial q} < 1 \) for any \( q \in (0, 1). \)

**Assumption A2.** [Assumption 2 under extension with monopoly] For any \( a \in \{0, ..., N\}, w \in \{1, ..., N\}, q, \epsilon \in (0, 1)^2, p \) is such that \( v^m \) is quasi-concave in \( a \) and either its variation rate in \( a, p(a + 1, w, q) - p(a, w, q) \) is bounded by \( \epsilon, \) or \( p \) satisfies

\[
|p(a, w + 1, q, \epsilon) - p(a, w, q, \epsilon)| > v^m(a^*, w, q, \epsilon) - \max\{v^m(1, w, q, \epsilon), v^m(N, w, q, \epsilon)\},
\]

where \( a^* \in \arg \max_{a \in [1, N]} v^m(a, w, q, \epsilon). \)

Lemma 1 is left largely unchanged, with the exception that the set of optimal coalitions now varies with \( \epsilon. \) We now have:

**Lemma A7 (Threshold strategy).** Let \( C^{*}_{gq\epsilon} = \arg \max_{c \in C_g} u^m(c, g, q, \epsilon). \) There is a threshold \( \hat{\epsilon}(g, q) > 0 \) such that all equilibria have the same outcome that \( s \) 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 A7. The proof follows directly from lemma A5 and proposition 7 in Appendix A.3.

Proposition 1 is left unchanged. Its proof changes slightly.

Proof of proposition 7. The proof proceeds as the proof of proposition 1 in the main specification of the model (Appendix A.1), with the exception that equation A5 now becomes:

\[
\nabla_x u^m_2 - u^m_1 \equiv \frac{\partial u^m_2 - u_1}{\partial \epsilon} x_\epsilon + \frac{\partial u^m_2 - u^m_1}{\partial q} x_q = \left[ \frac{\partial p_2}{\partial q} - \frac{\partial p_1}{\partial q} \right] x_q - (a_{c_2} - a_{c_1}) x_\epsilon \tag{A5}
\]

where \( x = (x_\epsilon, x_q) \) is a unit-length vector. As in the main specification, I show that equation A5 has a finite number of solutions. This equation has an infinite number of solutions if and only if the coefficients on \( x_q \) and \( x_\epsilon \) are both zero. Assumption A1 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 in the main specification of the model (Appendix A.1).

Lemma 2 changes in its formulation to accommodate the fact that the valuation of coalitions now varies with \( \epsilon. \)

**Lemma A8.** Let \( a_1 < a_c < a_2, \) and \( w_2 \leq w_1 < w_c. \) If assumption A1 holds, then

\[
v(a_c, w_c, q, \epsilon) < \max\{v(a_1, w_1, q, \epsilon), v(a_2, w_2, q, \epsilon)\}
\]

for any \( q, \epsilon \in (0, 1)^2. \)

Proof of lemma A8. The proof proceeds as the proof of lemma 2 in the main specification of the model (Appendix A.1), with one exception. The original proof showed that if the growth rate of \( p \) in \( a \) is bounded by \( 1/a \) as per assumption 2 then \( v \) is monotonic in \( a. \) Here note that
\[ v^m(a+1, w, q) - v^m(a, w, q) = p(a+1, w, q) - p(a, w, q) + \epsilon. \]

Immediately, if \( p(a+1, w, q) - p(a, w, q) \) is bounded by \( \epsilon \) as per assumption A2, then \( v^m \) is monotonic in \( a \).

The rest of the proof proceeds as the proof of lemma 2 in the main specification of the model (SI, section A.1).

Propositions 2 and 3 change slightly to accommodate the fact that the set of optimal coalitions now varies with \( \epsilon \). Their proofs are also largely similar. Propositions 2 and 3 become, respectively:

**Proposition A1.** If assumption A1 holds and \( c \in C^*_g \) for some \( q \in (0, 1) \), then \( c \) is minimal.

**Proposition A2.** Let \( c_1^* \in C^*_g \), \( c_2^* \in C^*_g \) for some \( \epsilon \leq \hat{\epsilon}(g, q_2) \). Suppose assumption A1 holds. We have \( q_1 < q_2 \Rightarrow \hat{\epsilon}(g, q_1) \geq \hat{\epsilon}(g, q_2) \) and \( a_{c_1^*} \leq a_{c_2^*} \).

**Proof of proposition A1.** I show the contrapositive of \( c' \in C^*_g \Rightarrow c' \in M_g \). Suppose \( c' \notin M_g \). Then there is \( c \in C_g \) such that \( a_c \leq a_{c'} \) and \( w_c < w_{c'} \). Suppose \( a_{c'} = a_c \), then \( v^m(a_c, w_c, q, \epsilon) > v^m(a_{c'}, w_{c'}, q, \epsilon) \). Suppose \( a_c < a_{c'} \). Then by lemma A8 \( v^m(a_{c'}, w_{c'}, q, \epsilon) < \max\{v^m(a_c, w_c, q, \epsilon), v^m(N, 0, q, \epsilon)\} \) for any \( q \in (0, 1) \).

**Proof of proposition A2.** 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 \( c^* \in C^*_g \). By lemma A7 we have \( u^m(c_1, g, q_1, \hat{\epsilon}_1) = u^m(c_2, g, q_1, \hat{\epsilon}_2) = 0 \). Since \( u^m \) is strictly decreasing in \( \epsilon \), it must be that \( u^m(c_2, g, q, \hat{\epsilon}_1) > u^m(c_2, g, q, \hat{\epsilon}_2) \). By lemma A7 it must be that \( u^m(c_2, g, q, \hat{\epsilon}_1) \leq u(c_1, g, q, \hat{\epsilon}_1) \). Therefore, \( u(c_2, g, q, \hat{\epsilon}_2) < u(c_1, g, q, \hat{\epsilon}_1) \), a contradiction.

I now show that \( q_1 < q_2 \Rightarrow a_{c_1^*} \leq a_{c_2^*} \). Let \( c_1^* \) be the largest coalition in \( C^*_g \) with sizes \( a_{c_1^*} \) and \( a_{c_2^*} \). To prove the claim, if suffices to show that \( a_{c_1^*} \leq a_{c_2^*} \). Suppose not. Because \( c_1^* \in C^*_g \) and \( c_2^* \in C^*_g \), we have \( u^m(c_2^*, g, q_1, \epsilon) - u^m(c_1^*, g, q_1, \epsilon) < 0 \) and \( u^m(c_1^*, g, q_2, \epsilon) - u^m(c_1^*, g, q_1, \epsilon) \geq 0 \). Since \( u^m(c_2^*, g, q_1, \epsilon) - u^m(c_1^*, g, q_1, \epsilon) \) is continuous in \( q \), there must be some \( q \in [q_1, q_2) \) such that \( u^m(c_2^*, g, q, \epsilon) = u^m(c_1^*, g, q, \epsilon) \). Because \( a_{c_1^*} > a_{c_2^*} \) assumption A1 implies that \( \frac{\partial}{\partial \epsilon} [u(c_2^*, g, q, \epsilon) - u(c_1^*, g, q, \epsilon)] \) is zero. Then \( u^m(c_1^*, g, q_1, \epsilon) - u^m(c_2^*, g, q_2, \epsilon) = u^m(c_1^*, g, q_1, \epsilon) - u^m(c_2^*, g, q_2, \epsilon) < 0 \), a contradiction.

Finally, propositions 4 and 5 are left unchanged but their proofs change slightly, again to accommodate the fact that the set of optimal coalitions now varies with \( \epsilon \). Note that these proofs rely on lemmas A1 and A2 proved in Appendix A.2. These lemmas are left unchanged, as they are graphical arguments that do not rely on the new payoff function. The proofs become:

**Proof of proposition 4.** By lemma A1 we have \( C' = C_g \) and for any \( c \in C_g \), \( w_{cg} \leq w_{cg'} \leq w_{cg} + 1 \). This implies \( u^m(c, g', q, \epsilon) \leq u^m(c, g, q, \epsilon) \). So \( \hat{\epsilon}_s(g', q) \leq \hat{\epsilon}_s(g, q) \). Let’s show the second part of the proposition. Proposition A1 implies that for any \( q \in (0, 1) \), there is \( c^* \in M_g \) that is realized in equilibrium on \( g \). I show the contrapositive. Suppose that for all \( c^* \in M_g \), there is some \( c \in M_g \) essentially equal to \( c^* \) such that \( j \notin c \) or \( i \in c \cup W_{cg} \). Then by lemma A1, \( c \) on \( g' \) is essentially equal to \( c^* \) on \( g \). As such, \( u^m(c, g', q, \epsilon) = u^m(c^*, g, q, \epsilon) \).

**Proof of proposition 5.** By lemma A1 we have that for any \( c \in C_g \), \( u^m(c, g, q, \epsilon) = u^m(c, g', q, \epsilon) \), which implies that \( \hat{\epsilon}_s(g', q) \geq \hat{\epsilon}_s(g, q) \). Let’s show the second part of the proposition. Suppose that there is \( q \in (0, 1) \) such that \( \hat{\epsilon}_s(g', q) \geq \hat{\epsilon}_s(g, q) \). Using proposition A1 this implies
that there is \( c \in M_g, c' \in M_{g'} \) such that \( u^m(c', g', q, \hat{\epsilon}_s(g, q)) > u(c, g, q, \hat{\epsilon}_s(g, q)) \). For this to be true, it must be that \( c' \in C_{g'} \setminus C_g \) for otherwise, lemma \( A1 \) implies \( u(c', g, q, \hat{\epsilon}_s(g, q)) = u^m(c', g', q, \hat{\epsilon}_s(g, q)) \). So lemma \( A2 \) implies that there is \( c^* \in C_g \) such that \( a_{c^*} < a_c, w_{c^*} \leq w_{c'}. \) It must be that \( w_{c^*} = w_{c'} \) for otherwise, lemma \( A8 \) implies \( u^m(c', g', q, \hat{\epsilon}_s(g, q)) < \max\{u^m(c^*, g, q, \hat{\epsilon}_s(g, q)), u^m(N, 0, q, \hat{\epsilon}_s(g, q))\} \). Suppose \( c^* \notin M_g \). Then there is \( c \in C_g \) such that \( a_{c^*} \leq a_c \) and \( w_{c^*} < w_{c'}. \) Then lemma \( A8 \) implies \( u^m(c', g', q, \hat{\epsilon}_s(g, q)) < \max\{u^m(c, g, q, \hat{\epsilon}_s(g, q)), u^m(N, 0, q, \hat{\epsilon}_s(g, q))\} \), a contradiction. 

A.4 Extension: cost of corruption as expected loss

In this extension, I consider a different formulation for the cost of corruption. While the main specification assumes that \( \epsilon \) is a sunk cost that accomplices pay upfront, this extension assumes instead that agents pay the cost \( \epsilon \) if and only if they get caught, which occurs with probability \( 1 - p(a, w, q) \). Furthermore, I assume for simplicity that this cost is constant across agents; that is, \( \epsilon_i = \epsilon \in (0, 1) \) for any \( i \in N \). Under equal-sharing, the payoff function in equation 1 and valuation functions become:

\[
 u(c, g, q, \epsilon) = v(a_c, w_{cg}, q, \epsilon) = \frac{p(a_c, w_{cg}, q)}{a_c} - [1 - p(a_c, w_{cg}, q)]\epsilon \quad (A6)
\]

if \( i \in c, \) and 0 otherwise.

I show in the next subsection that all the results from the main specification; that is, lemmas 1 and 2 and propositions 1 to 5 are left virtually unchanged, under qualitatively similar assumptions. Similar to the monopoly rule, the only noticeable difference is that the set of coalitions that are optimal to the seed, \( C^*_{ggq} \), now varies with \( \epsilon \). This is because in the main specification, the difference in payoffs from any two essentially different coalitions did not vary with \( \epsilon \). Under the payoff function in equation A6, this difference in payoffs from any two essentially different coalitions varies with \( \epsilon \).

Assumption 1 and 2 are replaced by qualitatively similar assumptions. The new assumption 1 still requires the ratio of partial derivatives of \( p \) with respect to \( q \) for any two coalitions to be bounded above by some quantity greater than 1. The new assumption 2 still requires either the growth rate of \( p \) in \( a \) to be bounded, or that \( p \) is sufficiently decreasing in \( w \). Assumptions 1 and 2 become, respectively:

**Assumption A3** (Assumption 1 under extension with cost as expected loss). If \( v(a_1, w_1, q, \epsilon) = v(a_2, w_2, q, \epsilon) \) for some \( a_1 \leq a_2, w_1, w_2 \in \{0, ..., N\}, q \in (0, 1), \epsilon > 0 \), then \( \frac{\partial p(a_2, w_2, q, \epsilon)}{\partial q} / \frac{\partial p(a_1, w_1, q, \epsilon)}{\partial q} < \frac{a_1 a_2 + a_2^2}{a_1 a_2 + a_1^2} \) for any \( q \in (0, 1) \).

**Assumption A4** (Assumption 2 under extension with cost as expected loss). For any \( a \in \{0, ..., N\}, w \in \{1, ..., N\}, q, \epsilon \in (0, 1)^2, p \) is such that \( v \) is quasi-concave\(^{25} \) in \( a \) and either its growth rate in \( a, \frac{p(a+1, w, q) - p(a, w, q)}{p(a, w, q)} \), is bounded below by \( \frac{1}{a} \) or bounded above by \( \frac{1}{a(a+2)} \), or \( p \) satisfies

\[
 |p(a, w + 1, q) - p(a, w, q)| > a \left[ v(a^*, w, q, \epsilon) - \max \{v(1, w, q, \epsilon), v(N, w, q, \epsilon)\} \right]
\]

\(^{25}\)That is, \( v \) satisfies \( v(a, w, q) \geq \min\{v(a_1, w, q), v(a_2, w, q)\} \) for any \( a_1 \leq a \leq a_2 \in \{1, ..., N\} \).
where \(a^* \in \arg \max_{a \in \{1, \ldots, N\}} v(a, w, q, \epsilon)\).

### A.4.1 Results

Many of the proofs and propositions in this subsection are identical to results in the extension with monopoly rule. I detail the relevant changes here, and refer to the relevant results in Appendix A.3.1.

As in the extension under monopoly, Lemma 2 now must accommodate the fact that the set of optimal coalitions varies with \(\epsilon\). Its formulation and proof become as in lemma A8 in the extension under monopoly.

Proposition 1 is left unchanged. Its proof changes slightly.

#### Proof of proposition 1

The proof proceeds as the proof of proposition 1 in the main specification of the model (Appendix A.1), with the exception that equation A1 now becomes:

\[
\nabla_x u_2 - u_1 = \left[ \frac{\partial p_2}{\partial q} \left( \epsilon + \frac{1}{a_{c_2}} \right) - \frac{\partial p_1}{\partial q} \left( \epsilon + \frac{1}{a_{c_1}} \right) \right] x_q + (p_2 - p_1)x_\epsilon
\]

where \(x = (x_\epsilon, x_q)\) is a unit-length vector. As in the main specification, I show that equation A7 has a finite number of solutions. This equation has an infinite number of solutions if and only if both coefficients on \(x_q\) and \(x_\epsilon\) are zero.

I show that coefficient on \(x_q\) is non-zero. We have \(\frac{\partial p_2}{\partial q} \left( \epsilon + \frac{1}{a_{c_2}} \right) - \frac{\partial p_1}{\partial q} \left( \epsilon + \frac{1}{a_{c_1}} \right) = 0 \iff \frac{\partial p_2}{\partial q} / \frac{\partial p_1}{\partial q} = \left( \epsilon + \frac{1}{a_{c_1}} \right) / \left( \epsilon + \frac{1}{a_{c_2}} \right)\). Define \(h(\epsilon) \equiv \left( \epsilon + \frac{1}{a_{c_1}} \right) / \left( \epsilon + \frac{1}{a_{c_2}} \right)\). We have \(h'(\epsilon) \propto \frac{a_{c_1} - a_{c_2}}{a_{c_1} a_{c_2}-a_{c_1}} \leq 0\), since \(a_{c_1} \leq a_{c_2}\). Since \(\epsilon \in (0, 1)\), this implies \(h(\epsilon) \leq h(1) = \frac{a_{c_1} a_{c_2} + a_{c_2}}{a_{c_1} a_{c_2} + a_{c_1}}\). Using assumption A3 \(\frac{\partial p_2}{\partial q} - \frac{\partial p_1}{\partial q} < h(1) \leq h(\epsilon)\). As such, \(\frac{\partial p_2}{\partial q} \left( \epsilon + \frac{1}{a_{c_1}} \right) - \frac{\partial p_1}{\partial q} \left( \epsilon + \frac{1}{a_{c_1}} \right) \neq 0\).

The rest of the proof proceeds as the proof of proposition 1 in the main specification of the model (Appendix A.1).

As in the extension under monopoly, lemma 2 changes in its formulation to accommodate the fact that the valuation of coalitions now varies with \(\epsilon\). Its formulation becomes identical to that of lemma A8. Its proof relies on the new assumptions and becomes the following:

#### Proof of lemma A8

The proof proceeds as the proof of lemma 2 in the main specification of the model (Appendix A.1), with one exception. The original proof showed that if the growth rate of \(p\) in \(a\) is bounded by \(1/a\) as per assumption 2 then \(v\) is monotonic in \(a\). This proof shows instead if \(p\) is bounded by the bounds defined by assumption A4 then \(v\) is monotonic in \(a\).

Note that

\[
v(a+1, w, q) - v(a, w, q) \geq 0 \iff \frac{p(a+1, w, q) - p(a, w, q)}{p(a, w, q)} \geq \frac{1}{a(1+(a+1)\epsilon)}
\]

Define \(h(\epsilon) \equiv \frac{1}{a(1+(a+1)\epsilon)}\), and note that \(h'(\epsilon) \leq 0\). So if the growth rate of \(p\) in \(a\) is bounded below by \(1/a = h(0) \geq h(\epsilon)\), then \(v\) is non-decreasing in \(a\). Conversely, if the growth rate of \(p\) in \(a\) is bounded below by \(\frac{1}{a(a+2)} = h(1) \leq h(\epsilon)\), then \(v\) is non-increasing in \(a\).
The rest of the proof proceeds as the proof of lemma 2 in the main specification of the model (Appendix A.1).

As in the extension under monopoly, propositions 2 and 3 change slightly to accommodate the fact that the set of optimal coalitions now varies with \( \epsilon \). Their formulation and proof become as in propositions A1 and A2 in the extension under monopoly.

Finally, propositions 4 and 5 are left unchanged. Their proofs change slightly, again to accommodate the fact that the set of optimal coalitions now varies with \( \epsilon \). Proofs are the same as in the extension under monopoly (Appendix A.3.1).

A.5 Extension: detection as a function of the scale of corruption

A.5.1 Results

The model assumes that the probability of detection is independent of the scale of corruption. This assumption simplifies the analysis, but may be unrealistic. The effect is unclear. On the one hand, more profitable, grand corruption could be more egregious, hence more likely to be detected. On the other hand, grand corruption might allow agents to spend some of the additional profit to increase protection. I consider both cases and show that 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, I show that results do not change if the effect is sufficiently small. If the effect is sufficiently large, then one result changes: as capacity increases, 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 opposite, I favor the original assumption that the probability of detection either decreases for more profitable schemes, or does not increases by much. The rest of this subsection details changes in the setting, and the intuition behind changes in the results. The next subsection rewrites and proves the propositions that change under this extension.

In this extension, I assume again a constant cost: \( \epsilon_i = \epsilon \) for all \( i \). I also assume that \( \epsilon \in (0, 1) \), and make the probability of success dependent on \( \epsilon \) by amending our original probability of success as follows:

\[
\tilde{p}(a, w, q, \epsilon) = \rho(\epsilon)p(a, w, q),
\]

where \( \rho : (0, 1) \to (0, 1) \) is twice-differentiable and rescales the probability of success according to \( \epsilon \). Recall that \( 1 - \epsilon \) measures the scale of corruption, with large values of \( \epsilon \) indicating petty corruption. Assuming that \( \rho'(\cdot) > 0 \) makes grand corruption more likely to be detected, while assuming \( \rho'(\cdot) \leq 0 \) 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, \textit{independently} of the composition of the supporting coalition? The utility function now writes:

\[
u_i(c, g, q, \epsilon) = \begin{cases} 
  u(c, g, q, \epsilon) - \epsilon = \frac{\tilde{p}(a, w, q, \epsilon)}{a} - \epsilon, & \text{if } i \in c \\
  0, & \text{otherwise}
\end{cases}
\]
The valuation function $v(a, w, q, \epsilon) \equiv \frac{\hat{p}(a,c,w,q,\epsilon)}{a}$ is defined analogously. Assumption 1 is maintained, and assumption 2 changes in its formulation, to accommodate that $v$ now varies with $\epsilon$. It becomes:

**Assumption A5** (Assumption 2 under extension with detection as a function of scale of corruption). For any $a \in \{0, \ldots, N\}$, $w \in \{1, \ldots, N\}$, $q, \epsilon \in (0,1)^2$, $p$ is such that $v$ is quasi-concave in $a$ and either its growth rate in $a$, $\frac{p(a+1,w,q) - p(a,w,q)}{p(a,w,q)}$, is bounded by $\frac{1}{a}$ or $p$ satisfies

$$|p(a, w + 1, q) - p(a, w, q)| > a \left[v(a^*, w, q, \epsilon) - \max \{v(1, w, q, \epsilon), v(N, w, q, \epsilon)\}\right],$$

where $a^* \in \arg \max_{a \in \{1, \ldots, N\}} v(a, w, q, \epsilon)$.

For this entire extension, I also assume a regularity condition on $p$:

**Assumption A6** (Regularity condition on $p$). Either $p'(\epsilon) \leq \frac{a_c}{p(c,g,q,\epsilon)}$ for any $c \in C_g$, $q \in (0,1), \epsilon \in (0,1)$, or $p'(\epsilon) > \frac{a_c}{p(c,g,q,\epsilon)}$ for any $c \in C_g$, $q \in (0,1), \epsilon \in (0,1)$.

Most of the intuition does not change. For a fixed level of capacity $q$, when considering whether to accept the bribe, the seed looks at $C_g$ 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 $\epsilon$. When grand corruption is less likely to be detected than petty corruption, then that coalition becomes increasingly profitable as $\epsilon$ 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.

### A.5.2 Results

Lemma 1 changes slightly. The new proof shows that the set of optimal coalitions does not vary with $\epsilon$ and that the threshold strategy flips when grand corruption is too likely to be detected: in this case, the seed now rejects projects whose scale is too large. The lemma reads:

**Lemma A9.** Let $C_g^* \equiv \arg \max_{c \in C_g} u_s(c, q, \epsilon)$. We have that $c \in C_g^*$ for some $\epsilon > 0$ if and only if $c \in C_g^*$ for any $\epsilon \in (0,1)$.

There is a threshold $\epsilon_s(g, q) \in (0,1)$ such that:

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

- If $\rho'(\epsilon) > \frac{a_c}{p(c,g,q,\epsilon)}$ for any $c \in C_g$, $q \in (0,1), \epsilon \in (0,1)$, then all equilibria have the same outcome that $s$ rejects the bribe if $\epsilon < \epsilon_s(g, q)$. Otherwise, she accepts it, and some coalition $c^* \in C_g^*$ is realized.

26 That is, $v$ satisfies $v(a, w, q) \geq \min\{v(a_1, w, q), v(a_2, w, q)\}$ for any $a_1 \leq a \leq a_2 \in \{1, \ldots, N\}$. 

---

46
Proof of lemma A9. Let’s first show that \( \arg \max_{c \in C_g} u_s(c, g, q, \epsilon) = \arg \max_{c \in C_g} u_s(c, g, q, \epsilon') \). Consider \( c, c' \in C_g \) such that \( u_s(c, g, q, \epsilon) \leq u_s(c', g, q, \epsilon') \). This implies \( \rho(\epsilon)p(a_c, w_c, q, \epsilon) \leq \rho(\epsilon)p(a_{c'}, w_{c'}, q, \epsilon') \). As such, \( u_s(c, g, q, \epsilon) \leq u_s(c', g, q, \epsilon') \), proving the point.

Let’s show the rest of the lemma. Note that \( \frac{\partial u}{\partial \epsilon} = \rho'(\epsilon)p(c, g, q, \epsilon) - 1 \). So \( \frac{\partial u}{\partial \epsilon} \geq 0 \iff \rho'(\epsilon) \geq \frac{\rho(\epsilon)p(a_c, w_c, q, \epsilon)}{a_c} \).

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

whenever \( u_s(c^*, g, q, \epsilon) \geq 0 \), we show as in lemma A10 that \( s \) accepts the bribe and \( c^* \) is an equilibrium outcome. □

Proposition A1 is left unchanged. Its proof changes slightly.

Proof of proposition A1. The proof proceeds as the proof of proposition A1 in the main specification of the model (Appendix A.1), with the exception that equation A1 now becomes:

\[
\nabla_x u_2 - u_1 = \rho(\epsilon) \left[ \frac{\partial p_2}{\partial q} \bigg/ a_{c_2} - \frac{\partial p_1}{\partial q} \bigg/ a_{c_1} \right] x_q + \rho'(\epsilon) \left[ \frac{p_2}{a_2} - \frac{p_1}{a_1} \right] x_c
\]

where \( x = (x_c, x_q) \) is a unit-length vector. As in the main specification, I show that equation A10 has a finite number of solutions. This equation has an infinite number of solutions if and only if the coefficients on \( x_q \) and \( x_c \) are both zero. Note that \( \rho(\epsilon) \neq 0 \), and that assumption A1 implies that \( \frac{\partial q_1}{\partial q} - \frac{\partial q_2}{\partial q} \neq 0 \). So the coefficient on \( x_q \) is non-zero.

The rest of the proof proceeds as the proof of proposition A1 in the main specification of the model (SI, section A.1). □

Lemma 2 changes in its formulation to accommodate the fact that the valuation of coalitions now varies with \( \epsilon \), but proves exactly as lemma 2.

Lemma A10. Let \( a_1 < a_c < a_2 \), and \( w_2 \leq w_1 < w_c \). Then

\[
v(a_c, w_2, q, \epsilon) < \max\{v(a_1, w_1, q, \epsilon), v(a_2, w_2, q, \epsilon)\}
\]

for any \( q, \epsilon \in (0, 1)^2 \).

Proposition 2 is left unchanged by the extension and proves as in the main specification (Appendix A.2).

Proposition 3 becomes:

Proposition A3. Let \( c_1^* \in C_{g_1}^*, c_2^* \in C_{g_2}^* \). If \( \rho'(\epsilon) < \frac{a_c}{p(c, g, q, \epsilon)} \) for any \( c, g, q, \epsilon \), we have \( q_1 < q_2 \Rightarrow \epsilon(g, q_1) \geq \epsilon(g, q_2) \). If \( \rho'(\epsilon) > \frac{a_c}{p(c, g, q, \epsilon)} \) for any \( c, g, q, \epsilon \), we have \( q_1 < q_2 \Rightarrow \epsilon(g, q_1) \leq \epsilon(g, q_2) \). For any \( \rho'(\epsilon) \), we have \( a_{c_1} \leq a_{c_2} \).

Proof of proposition A3. From lemma A9, \( \epsilon(g, q) \) is one of the bounds of the interval in \( \epsilon \) such that \( s \) accepts the bribe; that is, such that \( u(c^*, g, q, \epsilon) \geq \epsilon \) for some \( c^* \in \arg \max_{c \in C_g} u(c, g, q, \epsilon) \).
Pick \( q_1 < q_2 \), and their associated equilibrium coalitions, \( c_1, c_2 \). Coalition \( c_1 \) satisfies \( u(c_1, g, q_1, \epsilon) \geq u(c_2, g, q_1, \epsilon) \). Since for a given coalition, \( u \) is decreasing in \( q \), we have \( u(c_2, g, q_1, \epsilon) \geq u(c_2, g, q_2, \epsilon) \). This implies \( u(c_1, g, q_1, \epsilon) \geq u(c_2, g, q_2, \epsilon) \). As such, the interval such that \( u(c_1, g, q_1, \epsilon) \geq \epsilon \) has a weakly greater range than the interval such that \( u(c_2, g, q_2, \epsilon) \geq \epsilon \). That is, \( \epsilon(g, q_1) \geq (\leq) \epsilon(g, q_2) \) if \( \rho'(\epsilon) \leq (>) \frac{a_{c}+1}{p(c,g,q,\epsilon)} \). The rest of the proposition proves as in proposition 5.

Proposition 4 becomes:

**Proposition A4.** Suppose \( g' = g + i \rightarrow j \). Then if \( \rho'(\epsilon) \leq (>) \frac{a_{c}}{p(c,g,q,\epsilon)} \) for any \( c, g, q, \epsilon \), we have \( \epsilon(g',q) \leq (\geq) \epsilon(g,q) \). For the inequality to hold strictly for some \( q \in (0,1) \), it must \( j \in c \), and \( i \notin c \cup W_{cg} \) for some minimal coalition \( c \in M_{g} \) and all coalitions essentially equal to \( c \).

**Proof of proposition A4.** By lemma A1 we have \( C_{g'} = C_{g} \) and for any \( c \in C_{g} \), \( w_{cg} \leq w_{cg'} \leq w_{cg} + 1 \). This implies \( u(c, g', q, \epsilon) \leq u(c, g, q, \epsilon) \). Using lemma A9 gives that if \( \rho'(\epsilon) \leq (>) \frac{a_{c}}{p(c,g,q,\epsilon)} \) for any \( c, g, q, \epsilon \), then \( \epsilon(g',q) \leq (\geq) \epsilon(g,q) \). We prove the rest of the proposition as in the original proof.

Proposition 5 becomes:

**Proposition A5.** Suppose \( g' = g + ij \). Then if \( \rho'(\epsilon) \leq (>) \frac{a_{c}}{p(c,g,q,\epsilon)} \) for any \( c, g, q, \epsilon \), we have \( \epsilon(g',q) \geq (\leq) \epsilon(g,q) \). For the inequality to hold strictly for some \( q \in (0,1) \), it must be that there is \( c' \in M_{g} \) and \( c' \in C_{g'} \) such that \( a_{c'} > a_{c} \) and \( w_{c'} = w_{c} \).

**Proof of proposition A5.** By lemma A1 we have that for any \( c \in C_{g} \) \( u(c, g, q) = u(c, g', q) \), which implies that \( \epsilon(g',q) \geq (\leq) \epsilon(g,q) \) if \( \rho'(\epsilon) \leq (>) \frac{a_{c}}{p(c,g,q,\epsilon)} \) for any \( c, g, q, \epsilon \). The second part of the proposition proves as in proposition 5 if \( \rho'(\epsilon) \geq (>) \frac{a_{c}}{p(c,g,q,\epsilon)} \) for any \( c, g, q, \epsilon \). If \( \rho'(\epsilon) \geq (>) \frac{a_{c}}{p(c,g,q,\epsilon)} \) for any \( c, g, q, \epsilon \), then suppose that there is \( q \in (0,1) \) such that \( \epsilon(g',q) \leq \epsilon(g,q) \) and use the same argument as in proposition 5.

### A.6 Simulations and experiment

In this section, I prove that the functional form for the probability of success \( p \) in equation 3 used in simulations and in the experiment satisfies assumptions 1 and 2.

**p satisfies assumption 2.** We have \( \frac{\partial p(a_2, w_2, q)}{\partial q} / \frac{\partial p(a_1, w_1, q)}{\partial q} = \frac{N - (a_2 + w_2)}{N - (a_1 + w_1)} \). This implies

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

Note that since \( \frac{\partial p}{\partial q} < 0 \), \( a_1 = a_2 \) implies that \( v(a_1, w_1, q) \neq 0 \) for any \( q \in (0,1) \). Suppose \( a_1 < a_2 \). 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 \((a_2 - a_1) > 0\) and \((1 - q) > 0\), \( v(a_2, w_2, q) - v(a_1, w_1, q) = 0 \) requires \((N - w_1)a_2 - (N - w_2)a_1 > 0\).

**p satisfies assumption 3.** We have \( \frac{\partial p(a_2, w_2, q)}{\partial q} / \frac{\partial p(a_1, w_1, q)}{\partial q} \geq \frac{1}{a} \iff q \geq 1 - \frac{1}{N - a} \). So, if \( q \geq 1 - \frac{1}{N - a} \), then the growth rate of \( p \) in \( a \) is bounded below by \( \frac{1}{a} \). Otherwise, it is strictly bounded above by \( \frac{1}{a} \).
B Additional experimental results

This section provides additional experimental results. I first give additional details about the protocol, and show that respondents displayed satisfactory levels of comprehension (section B.1), then report the models used to estimate the main results (section B.2). I report post-hoc power calculations, and show that the design was sufficiently well-powered (section B.3). I then test for pooling or learning effects, and find little evidence for either (section B.4). Accounting for group- and respondent-level effects, and comparing between employees and students, I also show that subject characteristics had little impact on the results (section B.5). Finally, I provide details about recruitment, prompts, and other experimental material (sections B.6 to B.9).

B.1 Details about the protocol

I held 17 sessions of 16 respondents each in Mohammedia, Morocco, a mid-income country with median levels of corruption. This led to a sample of 272 subjects, and 808 games. I used a convenience sample that maximized sample size given existing budget constraints, and comprised of one quarter undergraduate students, and three quarters employees of the service industry (see next subsection for power analysis). Subjects were compensated, with the average payment amounting to daily minimum wage ($2.6 average gain and $5 show-up fee).

Figure B1 details the experimental protocol of a session. Before a session, I randomly decided whether it would be played with petty or grand corruption. Subjects entered the lab, and took a short pre-experiment survey. Subjects were then randomly assigned to groups of four. Each group had an enumerator that conducted the session. The enumerator read the instructions aloud, and conducted 12 repetitions of the game. The rent amounted to about $1 and was symbolized by 12 red cards. The cost $\epsilon$, called the “salary” in the experiment, was represented by 2 or 4 blue cards held by the subjects. For each round, the network was drawn on a board that was placed on the table. The probability of success $p$ associated to each coalition was communicated on a paper handout placed on the table.

In order to mimic the interpersonal interactions that arise in organizations, the game used face-to-face interactions. However, to implement take-it-or-leave-it, sequential offers, the enumerator would mediate communication around offers, asking subject $i$ if she wished to offer some amount to $j$, then asking $j$ whether she accepted $i$’s transfer, and if so, have $j$ give up her salary. Cheap talk was otherwise allowed. The outcome was drawn by rolling a hundred-sided dice. To prevent framing effects from biasing the results, the experiment used a neutral framing.

The twelve repetitions were divided in three parts of four repetitions each, corresponding to the baseline and the two treatments. Within each block, each subject got to be the seed once, and to occupy each of the other network positions once according to the ordering in figure B1. Within each block, two repetitions include the irrelevant tie, and two do not. The ordering was designed such that the same two subjects were always assigned to be the seed with the irrelevant tie, while the other two never did. After playing twelve repetitions, subjects took a post-experiment survey. They were paid their earnings, which amounted to about daily minimum wage ($2.6 average gain and $5 show-up fee).

Since subjects played the game several times, the design incorporates several features to test
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 game to behavior in another game (pooling). Learning and pooling effects pose a tradeoff. On the one hand, the game is cognitively taxing, and playing it repeatedly gives room for convergence to some equilibrium (which may be different from the one predicted by the theory). On the other hand, repeating the game might bias the results by (1) incentivizing subjects to pool across games, and (2) getting subjects to learn other players’ idiosyncratic strategies over time, making results diverge from the prediction in later repetitions. To discourage adverse learning and pooling effects, enumerators did not tell respondents how many repetitions of the game they would play, and did not allow them to keep track of their gains. In order to compare earlier and later games and test for potential learning and pooling effects, I randomized the order of the games within block, and randomly permuted the first two blocks (baseline and hard). I kept
the exposing tie block last, because it was more cognitively demanding. I show in section B.4 that learning effects are insignificant and mixed, and that there are no pooling effects.

Figure B2: **Average comprehension over time.** Questions 1, 2, and 3 correspond to the questions asked before the beginning of the first, second and third blocks of the experiment respectively. Question 4 was asked in the post-experiment survey.

For comprehension, subjects played practice repetitions before each block until the enumerator was confident that at least two out of four understood the rules. In practice, the enumerator usually gave one to two practice repetitions, and never more than three. I measured understanding before each block and at the end of the experiment through comprehension quizzes. For all but the last question, the enumerator would first record the subject’s answer, then correct her publicly so all could learn from her mistake. During a session, mean comprehension was above 80 percent, and reached 94 percent by the end of a session (Figure B2).
B.2 Main experimental results

<table>
<thead>
<tr>
<th></th>
<th>Baseline (1)</th>
<th>Accept Irrelevant (2)</th>
<th>GAM (3)</th>
<th>N accomplices Size (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hard, Grand</td>
<td>−0.089**</td>
<td>−0.089**</td>
<td></td>
<td>1.023***</td>
</tr>
<tr>
<td></td>
<td>(0.044)</td>
<td>(0.044)</td>
<td></td>
<td>(0.143)</td>
</tr>
<tr>
<td>Exposing, Grand</td>
<td>0.056**</td>
<td>0.056**</td>
<td></td>
<td>0.612**</td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td>(0.025)</td>
<td></td>
<td>(0.114)</td>
</tr>
<tr>
<td>Baseline, Petty</td>
<td>−0.081</td>
<td>−0.081</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.064)</td>
<td>(0.064)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hard, Petty</td>
<td>−0.444***</td>
<td>−0.444***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.090)</td>
<td>(0.090)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exposing, Petty</td>
<td>−0.306***</td>
<td>−0.306***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.088)</td>
<td>(0.088)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>With irrelevant tie</td>
<td>−0.052**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>History</td>
<td></td>
<td>−0.433***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.133)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.906***</td>
<td>0.932***</td>
<td>2.888***</td>
<td>1.563***</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.023)</td>
<td>(0.397)</td>
<td>(0.101)</td>
</tr>
<tr>
<td>Observations</td>
<td>808</td>
<td>808</td>
<td>732</td>
<td>508</td>
</tr>
<tr>
<td>R²</td>
<td>0.158</td>
<td>0.163</td>
<td></td>
<td>0.153</td>
</tr>
<tr>
<td>Log Likelihood</td>
<td></td>
<td>−324.154</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table B1: **Models used in the main text.** Clustered standard errors at the group-level in parentheses. Models 1, 2, 4 use use OLS, and all their variables are binary. Model 3 is a logistic generalized additive model (see footnote 23 in the main text for details about estimation). The variable history ranges from 2 to 4. Models 1 and 2 are used to construct Figure 7. Model 3 is used to construct figure 11, panel b. Model 4 is used to construct figure 8.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>baseline, grand</td>
<td>1.00</td>
</tr>
<tr>
<td>hard, grand</td>
<td>1.00</td>
</tr>
<tr>
<td>exposing, grand</td>
<td>1.00</td>
</tr>
<tr>
<td>baseline, petty</td>
<td>1.00</td>
</tr>
<tr>
<td>hard, petty</td>
<td>1.00</td>
</tr>
<tr>
<td>exposing, petty</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Table B2: **Fisher exact tests for differences in the distribution of realized coalitions with and without the irrelevant tie.** The p-value column reports the p-value of these tests. Adding the irrelevant tie never significantly alters the distribution of realized coalitions.

B.3 Power analysis

This section conducts post-hoc power analysis to show that the design was sufficiently well-powered to guard against the risk of false discovery. The experiment relies on a fairly complex design, such that:
• Subjects are clustered in groups of 4 participants.
• Each group plays 12 iterations of the game, with 4 iterations of each main treatment (Baseline, Hard, Exposing tie).
• Within each main treatment, 2 conditions include the irrelevant tie, and 2 do not include it.
• Two thirds of groups play under grand corruption, and one third play under petty corruption.

With these constraints, an experiment requires at least 3 groups, giving a minimum sample size $3 \times 4 = 12$ participants, and $3 \times 12 = 36$ games. Note that the design has the added difficulty that hypotheses on the size of the coalition can only be tested if the seed takes the rent. As such, I conduct power analyses using simulations, with sample sizes ranging from 36 groups (144 subjects) to 108 groups (432 subjects), and the conventional significance threshold of 5%. I evaluate the power underlying the main tests conducted in the paper (bottom panels of figures 7 and 8). Because sample size was determined so as to maximize the number of subjects within existing budget constraints, I did not conduct power analysis ex-ante. As a result, I report a post-hoc power analysis exercise and evaluate the statistical power associated with effects of the magnitude observed in the experiment. I report below the models used to conduct the main tests in the paper, and report in Table B3 the hypotheses that these models evaluate as well as the observed effect size.

$$accept = \beta_0 + \beta_1 \text{Hard, Grand} + \beta_2 \text{Exposing, Grand} + \beta_3 \text{Baseline, Petty} + \beta_4 \text{Hard, Petty} + \beta_5 \text{Exposing, Petty} + \epsilon \quad (B1)$$

$$N = \gamma_0 + \gamma_1 \text{Hard, Grand} + \gamma_2 \text{Exposing, Grand} + \epsilon \quad (B2)$$

$$accept = \delta_0 + \delta_1 \text{Hard, Grand} + \delta_2 \text{Exposing, Grand} + \delta_3 \text{Baseline, Petty} + \delta_4 \text{Hard, Petty} + \delta_5 \text{Exposing, Petty} + \delta_6 \text{With irrelevant tie} + \epsilon \quad (B3)$$

<table>
<thead>
<tr>
<th>Hypothesis no.</th>
<th>Null hypothesis</th>
<th>Expected result</th>
<th>Observed effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>$\beta_1 = 0$</td>
<td>Fail to reject</td>
<td>$-0.09$</td>
</tr>
<tr>
<td>H2</td>
<td>$\beta_2 = 0$</td>
<td>Fail to reject</td>
<td>0.06</td>
</tr>
<tr>
<td>H3</td>
<td>$\beta_3 = 0$</td>
<td>Fail to reject</td>
<td>$-0.08$</td>
</tr>
<tr>
<td>H4</td>
<td>$\beta_4 - \beta_2 = 0$</td>
<td>Reject</td>
<td>$-0.36$</td>
</tr>
<tr>
<td>H5</td>
<td>$\beta_5 - \beta_1 = 0$</td>
<td>Reject</td>
<td>$-0.36$</td>
</tr>
<tr>
<td>H6</td>
<td>$\gamma_1 = 0$</td>
<td>Reject</td>
<td>1.02</td>
</tr>
<tr>
<td>H7</td>
<td>$\gamma_2 = 0$</td>
<td>Reject</td>
<td>0.61</td>
</tr>
<tr>
<td>H8</td>
<td>$\delta_6 = 0$</td>
<td>Fail to reject</td>
<td>$-0.05$</td>
</tr>
</tbody>
</table>

Table B3: **Hypotheses for power analysis.** Variables for null hypotheses are defined in equations (B1), (B2), and (B3). The last column reports the effect size observed in the experiment.
Figure B3 shows that the design is sufficiently well-powered. Even the smallest sample size (144 subjects) is sufficient to pick up effects as large as the ones observed for the hypotheses where it is expected to reject the null. Conversely, the effect sizes observed for effects where it is expected to fail to reject the null are so small that even doubling the sample size would not allow detecting those effects at the conventional significance threshold of 5%.

Figure B3: **Post-hoc power analysis.** Power analysis conducted using 10,000 simulations per sample size. Each simulation draws error terms for the models reported in equations B1 to B3, taking as many draws as required by the sample size under consideration, and considers a significance threshold of 5% with effect sizes matching the ones observed during the experiment. The design is sufficiently well-powered. Even the smallest sample sizes are sufficient to pick up effects of that magnitude for hypotheses where it is expected to reject the null. Conversely, even doubling the sample size would not allow detect effects as small as the ones observed for hypotheses where it is expected to fail to reject the null. See Table B3 for details about the hypotheses.

### B.4 Learning and pooling effects

This section tests for potential learning and pooling effects. *Learning* effects refer to whether subjects converge to or diverge from equilibrium predictions over time, while *pooling* effects are the act of tying behavior in a game to behavior in another game. Learning and pooling effects are challenging because they pose a tradeoff. On the one hand, the game is cognitively taxing, and playing it repeatedly gives room for convergence to some equilibrium, which may be different from the one predicted by the theory. On the other hand, repeating the game might bias the results by (1) incentivizing subjects to pool across games, and (2) getting subjects to learn other players’ idiosyncratic strategies over time, making results diverge from the prediction in later repetitions.
The experimental design detailed in section section B.1 incorporates several features to discourage adverse learning and pooling effects, and measure their magnitude. To minimize these effects, enumerators did not tell respondents how many repetitions of the game they would play, and did not allow them to keep track of their gains. To evaluate these effects, I randomized the order of the games. The experiment was divided into three parts of four repetitions each, corresponding to the main treatment conditions (baseline, hard, and exposing tie). I randomized the order of the games within part, and randomly permuted the first two parts (baseline and hard). I kept the exposing tie part last, because it was more cognitively demanding.

Learning effects might go two ways. On the one hand, learning might have the expected effect: subjects may converge to the equilibrium strategy over time. Learning could also have an unexpected effect: subjects might learn about each others’ types, and further diverge from equilibrium strategy. Suppose that a group contains a subject who never accepts the bribe. Over time, other subjects may progressively learn about this and adjust their strategies accordingly, hence deviating from equilibrium strategy over time.

Comparing games that were played early and late in the first two blocks, I show that learning effects are insignificant and mixed. In early or late repetitions, results never vary significantly. Over time, some results converge to the equilibrium prediction, while others diverge. I use the first two blocks of the experiment to estimate variation in the effect of increasing capacity over time between the early and the late block using a difference in difference strategy. I estimate similarly variation in the the effect of adding non-exposing ties. Figure B4 shows that both in early and late repetitions, results go in the expected direction. Effect size never varies significantly. The effect of monitoring on size gets marginally closer to the prediction, while its effect on coalition size gets further away from it. The effect of adding non-exposing ties is minuscule. Finally, I compare the distribution of errors, measured as deviations from predictions under bargaining (see section 2.2.2), in early and late blocks. Table B4 shows that errors follow a very similar distribution in early and late repetitions. I test for differences in the distribution of errors in offering and accepting behaviors between early and late repetitions within treatment using Fisher exact tests (Table B6). Both tests fail to reject the null that errors are distributed similarly in early and late treatments.

Pooling effects mean that participants may tie their behavior in one game to behavior in another game. This is problematic because the model analyzes a one shot game, and because pooling may explain why participants engaged in greedy bargaining, with offers that leave recipients with negative surplus, yet end up being accepted (Section 2.2.2). When pooling, subjects tacitly agree on reciprocal exploitation. Recipient \( i \) accepts to be exploited by offerer \( j \) in some round of the game because she knows that she will exploit \( j \) when she will get to be an offerer in a later round. Pooling effects should imply an end-game effect; that is differences in behavior in the very last repetitions of the game. Specifically, recipients in the last repetitions would be less inclined to accept greedy bargaining because there is no further opportunity to reciprocate.

Comparing the first two repetitions of the last part (exposing tie) to the last two repetitions of that part, I show no evidence for pooling effects. In particular, I look at the distribution of
Figure B4: **Learning effects.** Bars are semi-parametric bootstrapped 95 percent confidence interval clustered at the group level using 10,000 replicates. The top three panel report estimates for early and late iterations of the first two blocks. The bottom panel reports estimates for early and late games within the last blocks. There is little evidence for learning and pooling effects: behavior never differ significantly between the late and early blocks. The models used to construct this figure are reported in tables **B7** and **B8**.

offers as deviation from predictions under bargaining. Table **B5** shows that these distributions are very similar in early and late repetitions. I test for differences in the distribution of errors in offering and accepting behaviors between early and late repetitions within treatment using Fisher exact tests (Table **B6**). Both tests fail to reject the null that errors are distributed similarly in early and late treatments. Figure **B4** also shows that facing an equally greedy offer (the median offer, which is greedy by about 1 credit), recipients are equally likely to accept that offer in early and in late repetitions.
Partners in crime

Table B4: **Learning effects for distribution of errors relative to the bargaining rule.**
This table reproduces Table 3 separately for early and late rounds. Numbers denote observed frequencies. Subjects over-share and over-accept: errors (italicized cells) are overwhelmingly false positives. Trends are comparable in early and late round (Fisher exact tests not significant, table B6).

<table>
<thead>
<tr>
<th></th>
<th>did share</th>
<th>not share</th>
<th>did accept</th>
<th>reject</th>
</tr>
</thead>
<tbody>
<tr>
<td>should</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>share</td>
<td>.07</td>
<td>.42</td>
<td>.46</td>
<td>.55</td>
</tr>
<tr>
<td>not share</td>
<td>.31</td>
<td>.42</td>
<td>.55</td>
<td>.29</td>
</tr>
</tbody>
</table>

(a) Sender’s decision, early rounds

(b) Recipient’s decision, early rounds

(c) Sender’s decision, late rounds

(d) Recipient’s decision, late rounds

Table B5: **Pooling effects for distribution of errors relative to the bargaining rule.**
This table reproduces table 3. Numbers denote observed frequencies. Subjects over-share and over-accept: errors (italicized cells) are overwhelmingly false positives. Trends are comparable in early and late round (Fisher exact tests not significant, table B6).

<table>
<thead>
<tr>
<th></th>
<th>did share</th>
<th>not share</th>
<th>did accept</th>
<th>reject</th>
</tr>
</thead>
<tbody>
<tr>
<td>should</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>share</td>
<td>.12</td>
<td>.01</td>
<td>.24</td>
<td>.01</td>
</tr>
<tr>
<td>not share</td>
<td>.56</td>
<td>.31</td>
<td>.56</td>
<td>.19</td>
</tr>
</tbody>
</table>

(a) Sender’s decision, early rounds

(b) Recipient’s decision, early rounds

(c) Sender’s decision, late rounds

(d) Recipient’s decision, late rounds

Table B6: **Fisher exact tests for differences in the distribution of errors relative to the bargaining rule in early and late rounds.** The rows on learning effect compare the distributions reported in Table B4. The rows on pooling compare the distributions reported in Table B5. The distributions are never significantly different between early and late rounds.

<table>
<thead>
<tr>
<th>Effect</th>
<th>Decision</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning Sender</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Learning Recipient</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Pooling Sender</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Pooling Recipient</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

B.5 **Individual-level characteristics**

This section shows that individual-level characteristics have little effect on the main results. I first show that group- and individual-level heterogeneity have little influence, and that group-level heterogeneity is larger than individual-level heterogeneity. I re-estimate our quantities of interest, using linear mixed models with individual-level random effects, group-level effects, and
### Table B7: Learning effects for main hypotheses.
Clustered standard errors at the group-level in parentheses. Analysis is subsetted to the first two blocks of the experiment. All models use OLS, and all variables are binary. Late is a binary variable equal to 1 if that block was played second in the experiment, and equal to 0 if it was played first. Most learning effects are not statistically different from zero. These models are used to construct Figure [B4](#).

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Irrelevant</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td></td>
</tr>
<tr>
<td>Hard, Grand</td>
<td>−0.156**</td>
<td>−0.156**</td>
<td>1.250***</td>
</tr>
<tr>
<td>(0.073)</td>
<td>(0.073)</td>
<td>(0.190)</td>
<td></td>
</tr>
<tr>
<td>Baseline, Petty</td>
<td>−0.052</td>
<td>−0.052</td>
<td></td>
</tr>
<tr>
<td>(0.093)</td>
<td>(0.093)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hard, Petty</td>
<td>−0.375***</td>
<td>−0.375***</td>
<td></td>
</tr>
<tr>
<td>(0.134)</td>
<td>(0.134)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>With irrelevant tie</td>
<td></td>
<td>−0.051</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.034)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Late</td>
<td>0.000</td>
<td>0.000</td>
<td>0.460**</td>
</tr>
<tr>
<td>(0.049)</td>
<td>(0.053)</td>
<td>(0.188)</td>
<td></td>
</tr>
<tr>
<td>Late × (Hard, Grand)</td>
<td>0.135</td>
<td>0.135</td>
<td>−0.455</td>
</tr>
<tr>
<td>(0.106)</td>
<td>(0.106)</td>
<td>(0.334)</td>
<td></td>
</tr>
<tr>
<td>Late × (Baseline, Petty)</td>
<td>−0.073</td>
<td>−0.073</td>
<td></td>
</tr>
<tr>
<td>(0.122)</td>
<td>(0.122)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Late × (Hard, Petty)</td>
<td>−0.115</td>
<td>−0.115</td>
<td></td>
</tr>
<tr>
<td>(0.179)</td>
<td>(0.180)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Late × With irrelevant tie</td>
<td>−0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.045)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.906***</td>
<td>0.932***</td>
<td>1.333***</td>
</tr>
<tr>
<td>(0.033)</td>
<td>(0.032)</td>
<td>(0.091)</td>
<td></td>
</tr>
</tbody>
</table>

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

### Table B8: Pooling effects for main hypotheses.
Clustered standard errors at the group-level in parentheses. Analysis is subsetted to the last block of the experiment. The model uses OLS. Late is a dummy variable equal to 1 for the last two games, and 0 otherwise. There is no pooling effect: the effect of deviations from the equilibrium offer does not vary between early and late games. This model is used to construct Figure [B4](#).

<p>| |</p>
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Accept</td>
</tr>
<tr>
<td>(b_i - b_i^*)</td>
</tr>
<tr>
<td>(0.015)</td>
</tr>
<tr>
<td>Late</td>
</tr>
<tr>
<td>(0.046)</td>
</tr>
<tr>
<td>((b_i - b_i^*)\times Late)</td>
</tr>
<tr>
<td>(0.018)</td>
</tr>
<tr>
<td>Constant</td>
</tr>
<tr>
<td>(0.028)</td>
</tr>
</tbody>
</table>

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

58
Figure B5: **Random effect specifications.** The specifications without random effects is estimated using a Gaussian GLM with errors clustered at the group-level; RE specifications use linear mixed models. Bars are semi-parametric bootstrapped 95 percent confidence interval clustered at the group level using 10,000 replicates. The main quantities of interest are robust to adding random effects. The models used to construct this figure are reported in Table B.5 both. Figure B5 shows that the quantities of interest are virtually unchanged. Table B.5 shows that random effect specifications fit the data marginally better than a specification without pooling, suggesting that there is little heterogeneity across groups, or across groups. Furthermore, individual-level effects add virtually no predictive power. This shows that individual-level effects are very small compared to group-level effects, and further justifies our decision to cluster errors at the group level.

Second, I show that although they have very different characteristics, students and employees have very similar behavior. Table 2 in the main paper showed that employees are poorer, less educated, more rural, less altruistic, and more extroverted than students. Yet, their behavior is very similar in the lab. I reestimate the quantities of interest separately for students and
employees (Figure B6): the predictions for students are more noisy because of the smaller sample size, but they largely overlap with that of employees.

Figure B6: **Students vs. employees.** “All” reports the estimates from the main specification (Table B1). The specifications for students and employees are estimated using OLS with boot-strap errors clustered at the group-level. Bars are semi-parametric bootstrapped 95 percent confidence interval clustered at the group level using 10,000 replicates. Students and employees have similar behavior. The models used to construct this figure are reported in Table B10.
Table B9: Random effect specifications. Clustered standard errors at the group-level in parentheses. The specifications without random effects is estimated using a Gaussian GLM with errors clustered at the group-level; RE specifications use linear mixed models. All variables are binary. Models have virtually identical point estimates. Random effects have little impact on model fit (AIC), but group effects reduce it more than individual effects. These models are used to construct Figure B5.
### Table B10: Students vs. employees.

Standard errors are in parentheses, and errors are clustered at the group level. All models use OLS, and all variables are binary. Effects for students and employees are comparable. These models are used to construct Figure B6.

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Accept</th>
<th>Irrelevant</th>
<th>Irrelevant</th>
<th>N accomplices</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5) Size</td>
</tr>
<tr>
<td>Hard, Grand</td>
<td>−0.172</td>
<td>−0.074*</td>
<td>−0.172</td>
<td>−0.074</td>
<td>1.224***</td>
</tr>
<tr>
<td></td>
<td>(0.107)</td>
<td>(0.043)</td>
<td>(0.107)</td>
<td>(0.043)</td>
<td>(0.274)</td>
</tr>
<tr>
<td>Exposing, Grand</td>
<td>−0.034</td>
<td>0.072***</td>
<td>−0.034</td>
<td>0.072***</td>
<td>0.942***</td>
</tr>
<tr>
<td></td>
<td>(0.061)</td>
<td>(0.027)</td>
<td>(0.061)</td>
<td>(0.027)</td>
<td>(0.223)</td>
</tr>
<tr>
<td>Baseline, Petty</td>
<td>−0.142</td>
<td>−0.070</td>
<td>−0.148</td>
<td>−0.068</td>
<td>0.986***</td>
</tr>
<tr>
<td></td>
<td>(0.105)</td>
<td>(0.065)</td>
<td>(0.106)</td>
<td>(0.065)</td>
<td>(0.125)</td>
</tr>
<tr>
<td>Hard, Petty</td>
<td>−0.377***</td>
<td>−0.526***</td>
<td>−0.383***</td>
<td>−0.525***</td>
<td>1.224***</td>
</tr>
<tr>
<td></td>
<td>(0.103)</td>
<td>(0.097)</td>
<td>(0.104)</td>
<td>(0.098)</td>
<td>(0.274)</td>
</tr>
<tr>
<td>Exposing, Petty</td>
<td>−0.348***</td>
<td>−0.309***</td>
<td>−0.354***</td>
<td>−0.307***</td>
<td>1.224***</td>
</tr>
<tr>
<td></td>
<td>(0.110)</td>
<td>(0.099)</td>
<td>(0.111)</td>
<td>(0.099)</td>
<td>(0.125)</td>
</tr>
<tr>
<td>With irrelevant tie</td>
<td>−0.075</td>
<td>−0.044*</td>
<td>(0.061)</td>
<td>(0.025)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.966***</td>
<td>0.896***</td>
<td>1.007***</td>
<td>0.917***</td>
<td>1.429***</td>
</tr>
<tr>
<td></td>
<td>(0.034)</td>
<td>(0.029)</td>
<td>(0.054)</td>
<td>(0.026)</td>
<td>(0.133)</td>
</tr>
</tbody>
</table>

|                  | (6) Size |
|                  | (5) Size |
| Hard, Grand      | 0.986*** |
|                  | (0.154)  |
| Exposing, Grand  | 0.551*** |
|                  | (0.125)  |
| Baseline, Petty  | 1.007*** |
|                  | (0.115)  |
| Hard, Petty      | 1.429*** |
|                  | (0.133)  |
| Exposing, Petty  | 1.589*** |
|                  | (0.115)  |

Sample Observations

<table>
<thead>
<tr>
<th></th>
<th>Students</th>
<th>Employees</th>
<th>Students</th>
<th>Employees</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations</td>
<td>189</td>
<td>619</td>
<td>189</td>
<td>619</td>
</tr>
<tr>
<td>R²</td>
<td>0.119</td>
<td>0.182</td>
<td>0.127</td>
<td>0.185</td>
</tr>
</tbody>
</table>

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

---

Partners in crime

Romain Ferrali
B.6 Location

The experiment was held in Mohammedia, Morocco from September 9-21, 2015. Working with our local partner, Mhammed Abderebbi of “MEDA Solutions” firm, we rented an apartment in Mohammedia appropriate for our lab. The apartment featured a large salon that we converted into a waiting room, and two bedrooms that we converted into a survey room and an experiment room. The survey room contained a bed, a couch, and a table, thus allowing three surveys to take place simultaneously with relative privacy. The experiment room contained two circular tables, each with five chairs for the enumerator and four subjects to play the game.

Due to unforeseen security threats (local youth demanding to participate in the experiment), we temporarily relocated sessions on September 16, 17, and 21 to our partner’s office, which similarly contained a waiting, survey, and experiment room.

B.7 Enumerators

Our partner selected two male and two female enumerators, three of them students from the Hassan II University in Mohammedia and one from our partner’s company. Having uploaded our pre-experiment survey, experiment survey, and post-experiment survey to Qualtrics, we trained our enumerators to administer the Qualtrics surveys on handheld tablets. We trained all four enumerators to administer the pre- and post-experiment surveys, and trained three of them (one male, and two females) to administer the experiment as well. Enumerators received 200 dirhams per day.

Training was held on Tuesday, September 8, 2015 and lasted half a day. It consisted in having the enumerators administer the pre- and post-experiment surveys to each other, under the author’s supervision. Similarly, they administered the diffusion game to each other, under the author’s supervision.

B.8 Subjects

Recruiters solicited subjects from public squares in Mohammedia, presenting them with flyers with the address, time, and following description:

Invitation to participate in a study session

The company MEDA Solutions has the honor of inviting you to a study session that will last about an hour. The topic is one’s financial behavior.

Day: XXX
Time: XXX
Address: Lotissement de la gare. Villa Mounia, no. 82. El Alia, Mohammedia
Phone number: 0668775219

Note: this invitation is personal and cannot be transferred to anyone else. You will not be allowed to participate without this invitation.

In recruiting subjects, we explicitly blocked on occupation, asking recruiters to recruit employees of the service industry, and, if necessary, completing with university students. Recruiters were told to select a diverse range of ages and occupations. Recruiters mentioned that all participants would receive 50 dirhams for their time plus any gains they won in the behavioral
Partners in crime

B.9 Prompts and material

![Table B11: Document displaying the probability of success in each treatment condition, for coalitions of 1 to 4 accomplices. In the exposing tie condition, the one-eyed cell denotes the coalition including the seed and the more isolated node, while the two-eyed cell denotes the coalition including the seed and the more exposed node.]

![Figure B7: Example comprehension question under grand corruption. Red and blue rectangles correspond to the bribe and salaries, respectively. Number 62 represents the outcome of the die. The question asked was: “How much as player 1 won?” [Answer: 0]]

**Prompt of the first block (control or hard)**

You are about to participate in an experiment on behavior in uncertain situations. The experiment looks like a game in which you will have to make several decisions that may make you win money. We will count the money in credits. One credit is worth a bit less than one dirham.

The experiment is very short. We will repeat it several times. Sometimes, we will change a few details. It is very important that you remain silent during the experiment. You will be able to talk only when I will allow you.
During the experiment, each of you will have a salary of \(2 \text{[4]}\) credits, represented by the \(2 \text{[4]}\) blue cards. You will have to decide between winning your salary with certainty, or taking a risk to maybe win a higher amount. You will be assigned to positions on a network [draw the star network on the board]. If two people are connected, they are “neighbors,” which allows them to communicate.

I will pick one of you and offer him \(12\) credits, represented by the \(12\) red cards. This person will have to decide between taking this sum and giving up her salary, or refusing this sum and keeping her salary. If he refuses it, the experiment is over, and you will all win your salary. If he takes it, I will offer him to share this amount with her neighbors. He will announce how much he wishes to offer to each. I will then allow the neighbors to accept or refuse. If they refuse, they keep their salary. If they accept, they give up their salary. The neighbors that have accepted will then be able to share the amount they have at hand with their neighbors that do not have pending offers and have not given up their salary. The experiment is over when no further offer can be made.

In the end, the ones that have held on to their salary win it. The ones that have given up their salary form a team. I will throw a dice. If the score is below some threshold, team members win their credits. Otherwise, they lose them. The threshold is written on this document [show the document]. It depends on the amount of team members.

Prompt of the second block (hard or control)

I will now change the probabilities of victory a bit. Note that now, it is more difficult [easier] for a player on his own to win.

Prompt of the third block (dense)

I will now change the network you are playing on [draw the line network on the board]. I will also change the probabilities of victory. They now not only depend on the amount of people in the team, but also on the amount of neighbors of the team that have held on to their salary. Now, sharing with the left hand side player only is better than sharing with the right hand side player only because the latter has one extra neighbor.

C Additional supportive evidence

This section presents preliminary evidence supporting an implication of proposition \(^3\) as institutional strength increases, corruption has a broader scope. Using cross-country comparisons and a comparison of 110 cases of corruption in India and the US. I show that controlling for the scale of corruption, instances of corruption in the US involve more accomplices than in India, and that accomplices in corruption schemes are better-paid in developed countries.
C.1 Cross-country comparisons

Cross-country comparisons use the 2017 Quality of Government dataset (QoG) and its related 2015 Expert Survey (QoGEx, Dahlström et al. 2015), to show that accomplices are better paid in developed countries (Table C1). Institutional strength captures the extent to which formal institutions allow detecting and punishing corruption. I proxy for this concept using a broad indicator for development: GDP per capita in PPP $ 2011, from the World Development Indicators (World Bank 2016). I construct the share of the rent of accomplices from the following question in QoGEx: “Hypothetically, let’s say that a typical public sector employee was given the task to distribute an amount equivalent to 1000 USD per capita to the needy poor in your country. According to your judgment, please state the percentage that would reach” (a) The needy poor, (b) People with kinship ties to the employee, (c) Middlemen/consultants, (d) The public employee’s own pocket, (e) The superiors of the public employee, (f) Others. The variable amounts to \((b+d)/(b+c+d+e)\).

<table>
<thead>
<tr>
<th>Dependent variable: Share of accomplices</th>
</tr>
</thead>
<tbody>
<tr>
<td>log(GDP p/c), $PPP 2011</td>
</tr>
<tr>
<td>Constant</td>
</tr>
</tbody>
</table>

Table C1: Cross-country comparisons, accomplices’ share vs. GDP p/c. Heteroskedastic-robust standard errors in parenthesis. The model is estimated using OLS. Accomplices pocket higher shares of the rent in more developed countries.

C.2 US-India comparison

I collected data on corruption cases in the US and India by searching for the words “arrest” and “corruption,” “fraud,” “bribery,” “embezzlement,” or “graft” (as well as their variants, such as “arrested” or “corrupt”) in the National Desk of the New York Times (NYT) and the Times of India (TOI) using Factiva. I then went through each article to identify the ones actually covering corruption cases. For each selected article, I collected the amount stolen and the number of accomplices. While the latter measures the scope of corruption, I normalize the amount stolen by Gross National Income (GNI) per capita to obtain a measure of the scale of corruption indicating its profitability relative to average income. In the NYT data, I covered the 2000-2014 time period and ended up with 55 cases. For TOI, I started at December 31, 2014 and stopped collecting data when I obtained a sample of the same size.

I compare the US and India because the former has stronger institutions than the latter. I picked the NYT and the TOI because they are both major national dailies in two large democracies with a vivid free press. This lends confidence that both newspapers will cover corruption cases to a similar extent. I ran the above query because using a large vocabulary for
corruption would select many articles while looking for the word “arrest” would select the first article on the case to appear in the newspaper, which would usually be the most detailed.

Using newspaper data on corruption is not uncommon (see, for instance, Glaeser and Goldin 2008). This data has several pitfalls. Most importantly, different newspapers may select differently on the types of cases they cover. The fact that both newspapers are national, generalist dailies should alleviate this concern. Furthermore, corruption being more widespread in India than in the US should push the TOI to select against petty corruption, which would be less interesting to its readers. As such, selection would only dampen the finding that petty corruption is more prevalent in India. In any case, the stylized facts below should only be taken as tentative evidence.

Table C2 provides a few descriptive statistics, and table C3 shows the finding: controlling for scale, corruption has a broader scope in the US than in India.

<table>
<thead>
<tr>
<th></th>
<th>India</th>
<th>USA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median amount stolen, fraction GNI p/c</td>
<td>0.31</td>
<td>1.7</td>
</tr>
<tr>
<td>Mean N accomplices</td>
<td>3.02</td>
<td>10.79</td>
</tr>
<tr>
<td>Percent cases with strong ties</td>
<td>0.192</td>
<td>0.209</td>
</tr>
<tr>
<td>First case</td>
<td>2014-11-04</td>
<td>2000-03-18</td>
</tr>
<tr>
<td>Last case</td>
<td>2014-12-31</td>
<td>2014-10-21</td>
</tr>
<tr>
<td>N</td>
<td>55</td>
<td>55</td>
</tr>
</tbody>
</table>

Table C2: Descriptive statistics on corruption in India and the US.

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent variable:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N accomplices</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poisson</td>
<td></td>
<td></td>
</tr>
<tr>
<td>negative binomial</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>log(amount)</td>
<td>0.197###</td>
<td>0.237###</td>
</tr>
<tr>
<td>(0.031)</td>
<td>(0.051)</td>
<td>(0.304)</td>
</tr>
<tr>
<td>USA</td>
<td>1.102###</td>
<td>1.182###</td>
</tr>
<tr>
<td>(0.199)</td>
<td>(0.304)</td>
<td>(0.234)</td>
</tr>
<tr>
<td>Constant</td>
<td>−0.380**</td>
<td>−0.488**</td>
</tr>
<tr>
<td>(0.176)</td>
<td>(0.234)</td>
<td>(0.234)</td>
</tr>
<tr>
<td>Observations</td>
<td>71</td>
<td>71</td>
</tr>
<tr>
<td>Akaike Inf. Crit.</td>
<td>253.305</td>
<td>216.467</td>
</tr>
</tbody>
</table>

Table C3: **Count regressions for number of accomplices.** Amount is measured as a fraction of GNI p/c. Controlling for the amount stolen, corruption in the US involves more accomplices than in India.