

How to Organize Crime*

Mariagiovanna Baccara and Heski Bar-Isaac[†]

October 2007

Abstract

In criminal organizations, diffusing information widely throughout the organization might lead to greater internal efficiency (in particular, since these organizations are self-sustaining, through enhancing trust). However, this may come at a cost of leaving the organization more vulnerable to external threats such as law enforcement. We consider the implications of this trade-off and characterize the optimal information structure, rationalizing both “hierarchical” structures and organization in cells. Then, we focus on the role of the external authority, characterize optimal detection strategies, and discuss the implications of different forms of enforcement on the internal structure of the organization and policy. Finally, we discuss a number of applications and extensions.

1 Introduction

In this paper, we study the interplay between trust within an illegal organization and its vulnerability to the authorities. Illegal organizations function more effectively when the people who constitute them trust each other. We argue that information sharing is an

*Previous drafts have circulated under the titles “Power and organizational structure under an external threat” and “Crime, Punishment and Organizational Structure.” For their useful and stimulating comments, we thank the editor, two anonymous referees, Pierpaolo Battigalli, Jean-Pierre Benoit, Sandro Brusco, Patrick Bolton, Vicente Cuñat, Alessandro Lizzeri, Wolfgang Pesendorfer, Ronny Razin, Ariel Rubinstein, Joel Shapiro, Marciano Siniscalchi, Vasiliki Skreta, Alexander Stille, Larry White, and audiences at Bocconi University, Caltech, IDEI Toulouse, INSEAD, NYU, Paris-Jourdan Sciences Economiques, Pennsylvania State, Tel Aviv University, UCLA, University of Pennsylvania, USC, University of Venice, the American Economic Association 2007 annual meeting, the 2005 EEA, the 2006 ESSET, the 2005 IIOC, the Third Conference on Mathematical Methods in Counterterrorism, and the 2nd-UTCC Annual International Conference.

[†]Department of Economics, Stern School of Business, New York University, 44 W Fourth Street, New York, NY 10012. Email addresses: mbaccara@stern.nyu.edu and heski@nyu.edu.

important factor in building internal trust and cohesion, but it can leave the organization vulnerable. Understanding how this trade-off affects the information structure of an organization and its productivity can allow us to assess detection policies designed to destabilize such organizations.

Since 9/11, a \$400 billion annual budget has been passed in the US for the war on terror; new domestic institutions have been created or enhanced (among many others, the Counterterrorism section in the Criminal division at the U.S. Department of Justice); international intelligence cooperation has been strengthened; and new protocols and controversial legal tools such as the Patriot Act have been developed. However, these new agencies and institutions face the same basic questions that have challenged the prosecutors fighting organized crime in Italy, South America, and Eastern Asia, as well as authorities fighting terrorism all over the world now and in the past. How can we learn about the internal structure of criminal organizations? How should we go about investigating a criminal organization in order to break its internal cohesion? How does a criminal organization react to investigation policies? We highlight that simply understanding the information links within an organization gives us insight into addressing these questions.

The anecdotal evidence suggests that there is a wide heterogeneity across the information structures of different criminal organizations. The most credited theory about the Mafia, developed in the early '90s, has identified the so-called “Cupola” as the highest level of the organization—supposedly consisting of agents who hold large amounts of information about the organization itself and carry out the enforcement needed for the organization to function.¹ These crucial agents are shielded from the authorities since they are typically not directly involved in criminal activities. This theory suggests a centralized information and enforcement structure.

However, recent studies about modern terrorism suggest a decentralized organization characterized by the presence of independent “cells.” These cells consist of agents who know each other and enforce each other’s actions, but who have a very vague idea of how the organization looks outside the cell boundaries. Thus, even if authorities detect a cell, it is difficult to expand the detection further. This structure seems to resemble that of other organizations observed in history, including the anarchist and revolutionary organizations in the late 19th century in Europe and the communist organization in the early 20th

¹Another famous but less substantiated theory is the so-called “Third Level Theory,” which refers to a level of enforcement higher than the Cupola itself. The expression was first used by Falcone and Turone (1982).

century.²

Focusing simply on information allows us to understand and formally rationalize these different structures and thereby to take a step towards implications for policy-makers and other interested observers. We study the optimal information structure of organization where trust is critical. We represent the need for trust with an infinitely-repeated multi-person prisoners' dilemma, augmented with the possibility of additional punishment, which can help encourage "good" behaviour. We consider the trade-off between the enhancement in internal cohesion derived by exchanging internal information and the increase in vulnerability to detection that this exchange implies.

Of course, information exchange in an organization may have several distinct roles. In this paper, we focus on information exchange that is related to the enforcement of cooperation within the organization. In particular, we focus on information about a specific person that makes this person vulnerable to the individuals who know it. Examples of this kind of information include identity, whereabouts, or some incriminating evidence about a person.³ On the other hand, we abstract from information exchange that is related to increases in productivity, coordination, or benefit distribution in the organization. While in Section 7 we discuss the extension of our results to information with a directly productive role, one of our goals in this paper is to isolate and highlight how this specific kind of information, which is particularly relevant for the investigative authorities, is shared within the organization. Note that, after the information exchange takes place, the resulting information structure can coexist and may interact with other underlying organization structures, such as communication, production and decision-making structures.⁴ Our view is that the different roles of information exchange in an organization should be disentangled because they do not necessarily imply each other. To see this, observe that coordinating and planning activities can be conducted using only code-names.⁵ In that case, coordinating actions can

²Among the first revolutionaries to organize conspiracies into secret cells was Louis Auguste Blanqui, a socialist of the Napoleonic and post-Napoleonic eras. The cell organization structure was also largely used by the European partisan organization in WWII. See Anselmi (2003), who describes how the partisans in the Italian resistance "...knew each other by 'cells, which were typically small, only two or three individuals...".

³Thompson (2005), for example, describes that in his role as a journalist reporting on organized crime, he had to divulge his address and that of his close family members. Charlie (2002) describes how committing a murder in front of peers is often an initiation ritual used by US gang members.

⁴Even when considering operational decisions, a number of different structures might be relevant. For example Tucker (2007) highlights that informal network position in addition to formal reporting roles can play an important role in adoption decisions.

⁵Indeed this was exactly what was done, for example, by the Italian resistance in WWII. Moreover, the

be quite independent of who within the organization knows the real name or whereabouts of other members of the organization. However, the diffusion of this kind of information clearly has implications both for the extent of trust (and the efficient functioning of the organization) and the organization’s vulnerability with respect to its enemies.⁶

We consider an organization of N agents and characterize its optimal information structure. When we address information structure, we have in mind a directional graph describing which members of the organization know, for example, the real name rather than nickname of some other member of the organization, or hold some incriminating evidence or other detailed information that would harm him if it came to light. We provide a model which rationalizes the benefit of information links—essentially arguing it leads to greater trust that is crucial in organizations that cannot rely on externally enforced contracts.

We also assume that there exists an external authority whose goal is to minimize the cooperation of the organization. It does so by allocating resources to detect the agents; further, by accessing information that they hold about other agents it can (indirectly) detect these further agents. The focus of our analysis is the information structure that optimizes the organization’s trade-off between productive efficiency and vulnerability, and in particular how this information structure reacts to alternative policies of the external agent.

We consider a general model of detection available to the external authority, where an agent’s probability of getting detected directly depends on the extent of his cooperation with the organization. This model includes two particularly interesting special cases. In the first—*agent-based detection*—the likelihood that the external authority detects each agent is independent of his level of cooperation. For example, regardless of the activities he is currently engaged in, authorities are actively seeking Osama Bin Laden and his lieutenants. More generally, standards of proof and the ability to extract information in some jurisdictions and for some crimes may be milder than for others. In the other extreme case—*cooperation-based detection*—an agent is never directly detected when not cooperating. For instance, if the organization members are drug dealers, a possible policy

investigations on Al Queda suggest that the same Al Queda member often uses different names depending on whom he is interacting with (see for instance the account of the investigation following the Pearl’s murder in Pearl (2003)).

⁶Even when organization members know each others’ names more “trust” can be attained by exchanging further information. For example, an agent j may possess incriminating evidence about another agent i , may know where i ’s family reside, or hold other information that can help force cooperation without enhancing productivity *per se* and yet is likely to leave i more vulnerable if j is caught.

for the authority is to look for drug exchanges. Then, if a member is more active he will be detected more often.

We give some general results and then fully characterize the optimal information structure within the organization in these two models of detection and compare them. In the agent-based detection model we find that if the probabilities of detection are sufficiently similar either it is optimal to create no information links or the optimal structure consists of “binary cells” (pairs of agents with information about each other, but with no information links with other members of the organization). We are also able to provide a full characterization of the structure for any probabilities of detection.

Given this characterization, we go on to consider the optimal budget allocation for an external authority who is trying to minimize cooperation within the organization. There are circumstances in which allocating the budget symmetrically induces the organization to exchange no information. In these cases, a symmetric allocation is optimal. However, sometimes a symmetric allocation induces the agents to form a binary cell structure. We show that in this case the authority optimizes by not investigating one of the agents at all, while investigating the others equally.

In the cooperation-based detection model, since each agent’s probability of detection is a function of the level of cooperation within the organization, an optimal information structure may require lower levels of cooperation from some of the agents to keep them relatively shielded from detection. Even if all agents are ex-ante symmetric, we show that the optimal information structure can be asymmetric, resembling a hierarchy with an agent who acts as an information hub, does not cooperate at all, and thus remains undetected. If each individual agent’s contribution to the organization is sufficiently high, the optimal organization can also be a binary cell structure. Moreover, the optimal strategy of the external agent is quite different under cooperation-based detection. For example, devoting considerable resources to scrutinizing a single agent makes that agent relatively likely to be detected whether linked or not under the agent-based cooperation model and so makes it cheap for the organization to link the agent and induce him to cooperate. In contrast, under cooperation-based detection, it is costly to make such a scrutinized agent cooperate (and thereby increase considerably the probability that he is detected).

The driving forces in these two models are somewhat different. In the agent-based detection model the authority chooses a strategy that seeks to make it unappealing for an organization to allow one agent to be vulnerable to another. In the cooperation-detection model, however, the external authority’s strategy is driven not so much by an attempt

to make it unappealing to make one agent vulnerable to another, but rather to make it unappealing to cooperate and be vulnerable directly.

Although the principal motivation in writing this paper has been consideration of illegal organizations and criminal activity, the trade-off and considerations outlined above may play a role in legitimate organizations as well. In particular, many firms might gain some kind of internal efficiency by widely diffusing information within the organization, but might be concerned that this leaves the firm vulnerable to rival firms poaching informed staff.⁷ Thus, our results can shed some light on the optimal information sharing protocols of these organizations.

1.1 Related Literature

To our knowledge, this is among the first papers addressing the optimal information structure in organizations subject to an external threat.

Several papers have some elements that relate to our work. Work by Farley (2003, 2006) considers the robustness of a terrorist cell. In that work robustness is with regard to maintaining a chain of command in a hierarchy. Ben Porath and Kahneman (1996) study a repeated game model in which only a subset of the other agents can monitor an agent's actions but communication is allowed at the end of every period. They show that having two other agents observing each agent's actions is sufficient to implement efficient outcomes as the discount factor tends to 1. Although the role of the information structure is clearly different from ours, there is an interesting link between these results and the benefit side of our model that we discuss in Section 7.1.2.

Garoupa (2007) looks at the organizational problem of an illegal activity and at the trade-off between enhancing internal productivity and leaving members of the organization more exposed to detection. He takes a different approach, focusing on the optimal size of the criminal organization and taking its internal structure as given.

This paper is also related to the literature on social networks. In particular, Ballester et al. (2006), under the assumptions that the network structure is exogenously given and observed, characterize the “key player”—the player who, once removed, leads to the optimal change in aggregate activity. Reinterpreting networks as trust-building structures, in this paper, we ask how a network can be (endogenously) built to make criminal activity as efficient as possible.

⁷For instance, consider secrecy issues in patent races and R&D departments.

There is a wide literature on organization structure, though it has focused on somewhat different concerns to those raised in this paper. For example work by Radner (1992, 1993) and Van Zandt (1998, 1999) has highlighted the role of hierarchy in organizations—in particular, where agents have limitations on their abilities to process information—and Maskin, Qian and Xu (2000) have studied the impact of the organizational form on the incentives given to managers. Whereas these papers, in a sense, are concerned with the internal efficiency of the organization, the work of Waldman (1984) and Ricart-I-Costa (1988), which abstracts from considering what affects internal efficiency, highlights that external considerations (in their paper, the information transmitted to other potential employers and so affecting employee wages) might lead to distortions with respect to the structure that is most internally efficient.⁸ At the heart of this paper, by contrast, is the trade-off between particular internal and external efficiencies, specifically the allocation of information that gives the power to punish and, thereby, facilitates cooperative behaviour within the organization, but renders agents more vulnerable to an external threat. Note that while we focus on the structure of information in an organization, communication structure, formal decision-making hierarchies, networks of influence, and many other characterizations of information structures might coexist and, indeed, interact simultaneously. We abstract from all these latter considerations, which have been the focus of the work discussed above as well as of a wide literature in sociology (see, for example, Wasserman and Faust (1994)).

Recent contributions to the literature on cartels deal with the impact of an external authority on the cartel's behaviour. In particular, Harrington (2003), (2004) and (2005) looks at the impact of an external authority on the cartel's optimal pricing behaviour, and Spagnolo (2003) and Aubert *et al.* (2006) study the effect of leniency programs on cartels' stability.

This paper is also related to the literature on organized crime, though this literature has concentrated on the role of organized crime in providing a mechanism for governance or private contract enforcement. For such analyses of the organized crime phenomenon, see Gambetta (1993), Smith and Varese (2001), Anderson and Bandiera (2006), Bandiera (2003), Bueno de Mesquita and Hafer (2007), and Dixit (2004).⁹ Other papers that use

⁸Also, Zabochnik (2002) focuses on a situation in which a firm decides how to optimally distribute some (common) private information given an external threat—that is, the risk of employees leaving and joining competitors.

⁹For insightful and less formal accounts of the organized crime phenomenon, we refer the interested reader to Stille (1995) and Falcone (1991).

rational frameworks to model the behaviour of terrorist groups include Berman (2003), Berman and Laitin (2005) and Benmelech and Berrebi (2006).

2 Model

Suppose that there are $N > 2$ risk-neutral players, with N an even number and one additional player who we will refer to as the “external agent” or the “external authority”.¹⁰

The authority moves first and sets a given detection strategy as specified in Section 2.1. Then, the N players have the possibility of forming an information structure by exchanging information among themselves as specified below in Section 2.2.¹¹ After forming an information structure, the N agents start playing an infinitely repeated stage game as described in Section 2.3.

2.1 External Authority

At each period of the repeated stage game, each agent could be detected by the external authority. If the agent is detected then he must pay an amount $b > 0$. This payment may represent a punishment such as some time in prison, reduction in consumption or productivity.

There are two ways for an agent to be detected, a *direct* way and an *indirect* one. In particular, at each period, an independent Bernoulli random draw determines whether a particular agent is detected *directly*. The direct detection of a particular agent at each period is *independent* of other agents’ detection and detection in previous periods. Thus, at every period t an agent i can be detected directly by the authority according to some probability α_{it} . This probability depends both on whether or not the agent cooperates with the organization at period t and on the extent of the external authority’s scrutiny β_i which is determined by the authority at the beginning of the game. Specifically the probability with which agent i is detected directly in period t is $\alpha_{it} = \beta_i$ if the agent cooperates in

¹⁰Allowing N to be an odd number presents no conceptual difficulties, but adds to the number of cases that need be considered in some of the results with regard to how to treat the last odd agent, with no real gain in insight.

¹¹We assume that the N agents constitute an organization through some other production structure that is independent of the information structure. Although we do not explicitly model the formation process (see footnote 20 for a further discussion), one could assume that the information structure is determined by a “benevolent” third party. Indeed, this is essentially the approach advocated by Mustafa Setmariam Nasar, an Al-Qaeda strategist, who suggested that cell-builders be from outside the locale or immediately go on suicide missions after building cells.

period t , while it is $\alpha_{it} = \gamma\beta_i$ where $\gamma \in [0, 1]$ if he does not cooperate in period t . In particular, when $\gamma = 1$ the probability with which an agent is detected is independent of his behaviour. We term this case *agent-based detection*. When $\gamma = 0$ then the agent cannot be directly detected unless he is cooperating. We term this case *cooperation-based detection*.¹²

Second, the external authority might also detect agents indirectly. Indeed, we assume that when the external authority detects an agent who has information about other members of the organization (see below for the details on information exchange), the external authority detects these agents as well with probability one. Thus the external authority's ability to detect agents indirectly depends on the information structure.

The external agent has a budget $B \in [0, N]$ to allocate for detecting the N members of the organization and devotes $\beta_i \in [0, 1]$ to detecting member i where $\sum_{i=1}^N \beta_i \leq B$. Without loss of generality we label agents so that $\beta_1 \leq \beta_2 \leq \dots \leq \beta_N$. We refer to β_i as the external authority's level of scrutiny of agent i .

2.2 Information Structure

We assume that each of the agents has a piece of *private* and *verifiable* information about himself, and can decide to disclose this information to any of the other agents. Examples of such information could be the identity of the player, his whereabouts, some incriminating evidence, etc. We formalize the fact that player j discloses his information to player i by an indicator variable μ_{ij} , such that $\mu_{ij} = 1$ if and only if player i knows the information regarding player j ($\mu_{ij} = 0$ otherwise). We also use the notation $j \rightarrow i$ to represent $\mu_{ij} = 1$ (and, similarly, for instance $i, j, k \rightarrow l$ to represent $\mu_{li} = \mu_{lj} = \mu_{lk} = 1$).¹³ The set \mathcal{I} of all the possible organization (or "information") structures among N people is a subset of the set $\{0, 1\}^{N^2}$ of values of the indicator variables, and we denote by μ its generic element.

An agent i is indirectly linked to an agent j if there is a path of direct links that connect i to j . That is, if there is a set of agents $\{h_1, \dots, h_n\}$ such that $i \rightarrow h_1, h_1 \rightarrow h_2, \dots, h_n \rightarrow j$. Thus, given an information structure μ , for each agent i we can identify the set of agents *including i himself* and all those whom i is, directly or indirectly, linked to. We refer to

¹²In Section 6 we address circumstances in which the agent-based or cooperation-based detection model might be more appropriate. Broadly, we see this as arising from institutional constraints, for example, on the burden of proof that the external authority requires to apprehend a member of the organization.

¹³Note that $\mu_{ii} = 1$ for all i .

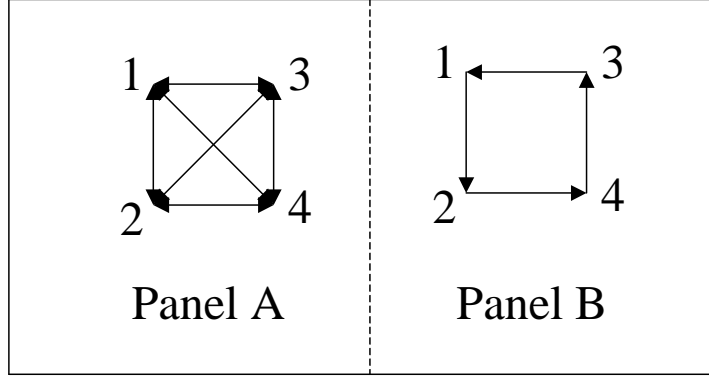


Figure 1: Examples of equivalent organizations

this set as V_i . Note that the sets V_i are induced by the choice of the information structure.

The information structure affects the agents' probabilities of detection by the external authority. Specifically, if i has information about another player j and i is detected (either directly or indirectly), player j is detected as well.¹⁴ Given an information structure μ and independent detection probabilities $\{\alpha_1, \dots, \alpha_N\}$, agent i is detected in one period if and only if at least one agent in V_i is detected.

Observe that, under the assumptions made so far, given an information structure μ , each agent is detected by the external agent with probability $1 - \prod_{j \in V_i} (1 - \alpha_j)$.

Note that different information structures μ might lead to identical $\{V_i\}_{i=1}^N$. For example, if $N = 4$, an information structure in which $\mu_{ij} = 1$ for all $i, j = 1, 2, 3, 4$, represented by Panel A in Figure 1, is equivalent to a structure in which $\mu_{12} = \mu_{23} = \mu_{34} = \mu_{41} = 1$ and $\mu_{ij} = 0$ otherwise, a structure represented by Panel B in Figure 1. In fact, $V_i = \{1, 2, 3, 4\}$ for all i , and the probability of detection is $1 - \prod_{j=1}^4 (1 - \alpha_j)$ for each player in both cases.

2.3 The Stage Game

After exchanging information about each other, the agents play an infinitely repeated stage game. In effect, we represent “trust” in the organization with an infinitely repeated multi-

¹⁴Allowing for “detection decay” (that is, supposing that if agent i has information about agent j and agent i is detected, then agent j is detected with probability less than 1) would not change the qualitative results of this paper. We discuss this extension in Section 7.1.7.

person prisoners’ dilemma game, augmented with the possibility of additional punishment.

In every period, each agent can either “cooperate” (C) or “not cooperate” (NC) and each agent i who has direct information over another agent j can also decide to make agent j suffer a punishment (“ P ”) or not (“ NP ”). The cooperation choice and the punishment choice are made simultaneously.

We assume that the full history of the game is perfectly observed by all members of the organization.¹⁵ Note that in our context, it is sufficient to suppose that everyone in the group may know that someone neglected their duties without necessarily knowing who. For example, if agents observed that a planned operation failed (and that the per-period payoff is lower than expected). Such an observation could sow distrust within the organization so that all agents would stop cooperating.

2.3.1 Cooperation

We focus on the cooperation choice first. The action sets of the stage-game associated with the cooperation choice for player i is $A_i = \{C, NC\}$ for $i = 1, \dots, N$. Cooperation is a productive action that increases the sum of the resources available to the agents, but it is costly for the agent who cooperates. In particular, if $n - 1$ other agents cooperate, the payoff of an agent is $\lambda n - c$ if he cooperates and $\lambda(n - 1)$ if he does not, with $\lambda, c > 0$.

We assume that $c > \lambda$, which implies that not cooperating is a dominant strategy in the stage game, and that $\lambda N > c$, which implies that full cooperation is the most efficient outcome of the stage game.

2.3.2 Punishment technology

Suppose that player i has revealed his information to player j (that is $\mu_{ji} = 1$). This revelation makes player i vulnerable to player j . In fact, we assume that if i reveals his information to j , then in every period of the stage game, player j can decide whether to “punish” (“ P ”) player i by making him pay a cost $k > 0$ or not to punish him (“ NP ”). Note that while being punished is costly, we treat the act of punishing as costless. This is a simplifying assumption intended to reflect that it is relatively easy to ensure appropriate

¹⁵ Assuming imperfect public monitoring (all members of the organization observe imperfect signals of each others’ actions) merely changes threshold values but not qualitative results. A more substantive extension is to suppose that the information structure affects agent’s ability to monitor each other’s actions. We discuss this extension extensively in Section 7.1.2.

punishment.¹⁶

The action set of player i associated with the punishment choices is $A'_i = \{P, NP\}^{\{j|\mu_{ij}=1\}}$ (if $\mu_{ij} = 0$ for all $j \neq i$, then player i cannot punish anybody, and $A'_i = \emptyset$).

We assume that every agent i can pay the cost k at most once at every period. This means that if two or more players know his information and they all decide to punish him, only one of the punishments has an effect.

2.4 The Game

2.4.1 Timing

The timing of the game is as follows:

(1) The external agent chooses the allocation $\{\beta_1, \dots, \beta_N\}$. The N agents perfectly observe this allocation.

(2) Each agent may or may not reveal his/her information to one or more of the others. An information structure $\mu \in \mathcal{I}$ arises.

(3) For a given information structure μ and implied sets V_1, \dots, V_N , the stage game described above is played an infinite number of times at periods $t = 1, 2, \dots$. At every period, agents simultaneously choose whether to cooperate or not cooperate and, if they have information on some other agents, whether to punish them or not. Moreover, at every period, each agent is directly detected in accordance with the external authority's scrutiny and the agent's level of cooperation (i.e., the induced $\{\alpha_1, \dots, \alpha_N\}$) and, if detected (either directly or indirectly) he has to pay a cost b .

Note that the external agent has to set its policy between the organization forms. Though we discuss (and relax) the timing assumption of the game in Section 7.1.1 we think that it is appropriate in situations in which the external agent represents public law enforcement, which may be fairly inflexible in setting its policies and strategies with respect to a criminal organization.

¹⁶For instance, the tasks of having someone beaten once you know his whereabouts, hurting his family once you know his identity, or stealing some of his belonging once you know his address are relatively cheap to accomplish. Moreover, a player may even benefit from choosing to punish another player. For example if punishing is achieved by whistle-blowing and is rewarded by authorities. This analysis lies beyond the scope of this paper. Alternatively, if one considers "punishing" as among the cooperative actions for which a threat of punishment is required, we are quickly led to consider structures in which every agent in the organization must be linked—a special case which is analyzed within our more general characterization.

2.4.2 Payoffs

The Agents and the Organization Let h^t denote a period $t \geq 1$ history in the repeated game.¹⁷ Let \mathcal{H} denote the set of histories. Then, player i 's (pure) strategy is denoted as $s_i : \mathcal{H} \rightarrow A_i \times A_i'$. As discussed above, the history is perfectly observed by all members of the organization.

Given the description of the agents' behaviour $s(h^t)$ at period t given history h^t , player i 's payoff in that period is

$$\pi_i^t(s(h^t)) = \lambda n(s(h^t)) - c1_{s(h^t)}^A(i) - k1_{s(h^t)}^B(i) - b \left[1 - \prod_{j \in V_i} (1 - \alpha_j(s(h^t))) \right], \quad (1)$$

where $n(s(h^t))$ denotes the number of players cooperating at time t under $s(h^t)$, $1_{s(h^t)}^A(i)$ is an indicator variable which takes the value 1 if agent i cooperates at history h^t under $s(h^t)$ and $1_{s(h^t)}^B(i)$ takes the value 1 if anyone with information about i chooses to punish him at history h^t and 0 otherwise. Note that, as specified above, the probability of direct detection α_i depends on the agent's strategy since it depends on whether or not the agent is cooperating. In the agent based detection model (where $\gamma = 1$) then the probability of detection is not affected by the agent's behaviour and determined solely by the authority's scrutiny).

The per-period payoff of agent i , $\pi_i^t(s(h^t))$, can be decomposed into an "internal" component and an "external" component. In particular, $\lambda n(s(h^t)) - c1_{s(h^t)}^A(i) - k1_{s(h^t)}^B(i)$ is the payoff coming from the interaction among the N agents in the stage game. Note that it is independent of the information structure μ (recall that μ determines the set V_i) and the costs associated with detection. In contrast, with respect to external vulnerability, we refer to $b \left[1 - \prod_{j \in V_i} (1 - \alpha_j) \right]$ as the per-period "*information leakage cost*" for agent i associated with the information structure μ .

We suppose that agents discount the future in accordance with a discount factor $\delta \in (0, 1)$, and we write $\pi_i(s) = \sum_{t=0}^{\infty} \delta^t \pi_i^t(s(h^t))$, where h^t is the history in period t induced by the strategy profile s . Finally, we can write down the overall payoff for the organization as $\Pi(s) = \sum_{i=1}^N \pi_i(s)$.

¹⁷ h^t contains information on the allocation $\{\beta_1, \dots, \beta_N\}$, on the organization structure μ , and on the previous decisions to cooperate or not and to punish or not by all players.

The External Agent We assume that the goal of the external agent is to minimize the cooperation among the N other agents. In other words, given that at each period t the production of the cooperation is $\lambda n(s(h^t))$ (where h^t is the history in period t induced by the strategy profile s), the external agent aims to *minimize* $\sum_{t=0}^{\infty} \delta^t \lambda n(s(h^t))$. For simplicity, we assume that the authority gets no utility from saving part of the budget B . Also, the external authority does not benefit from the the payments b incurred by the detected agents.¹⁸

2.4.3 Efficient Information Structure

For each information structure μ and for any δ , it is possible to identify a set of Subgame Perfect Nash Equilibrium (SPNE hereafter) in the repeated game. In the analysis of the game, to compare alternative information structures, for every information structure μ and for any δ , *we identify the most efficient SPNE outcome achievable under μ* (the SPNE that maximizes $\Pi(s)$) *when the discount factor is equal to δ* . Let us refer to such an outcome as $\Pi^*(\mu, \delta)$.¹⁹

For a given δ , we say that one information structure μ is *strictly more efficient* than another information structure μ' if we have $\Pi^*(\mu, \delta) > \Pi^*(\mu', \delta)$. Then, we assume that, once the external agent chooses the allocation $\{\alpha_1, \dots, \alpha_N\}$, the organization μ that will be formed is the most efficient one—that is, one that achieves the highest $\Pi^*(\mu, \delta)$. In other words, we assume that the N agents select $\mu^* \in \arg \max_{\mu} \Pi^*(\mu, \delta)$. Note this is a finite optimization problem and so has a well-defined solution.²⁰

¹⁸Indeed, these payments may be costly for the external authority. For example, they may consist of detention in prison facilities.

¹⁹Note that standard compactness and convexity results about SPNE outcomes in repeated games guarantee that the most efficient SPNE achievable under any μ is well-defined. However, note that the strategy $s^*(\mu, \delta)$ does not need to be uniquely defined.

²⁰We do not explicitly model the process of the formation of the organization. However, note that the information exchange is a one-time act that can be performed in a controlled environment in which it is easier to enforce efficient behaviour from the agents (in particular, it can involve the exchange of side-payments or hostages to be completed (see Williamson (1983)). After that, the agents move on to play the infinitely repeated game in which day-to-day cooperation is harder to sustain without punishments. Note also, that it is always possible to sustain this behaviour in equilibrium. To see this, assume that the agents decide simultaneously and *non-cooperatively* whether to reveal their information to other agents. In a game like this, it is always possible to obtain the most efficient organizational structure as an equilibrium outcome by imposing that if agents do not exchange information as prescribed by the most efficient organizational structure, no agents will ever cooperate in the repeated game.

3 The Optimal Information Structure: General Results

In this section, we study the optimal information structure problem from the criminal organization’s point of view. We take the authority’s scrutiny $\{\beta_1, \dots, \beta_N\}$ as given and we study the most efficient information structure that the other N agents can form.

First, we present some general results that apply to any $\gamma \in [0, 1]$. Then, we proceed to full characterizations of the agent-based detection and cooperation-based detection models ($\gamma = 1$ and $\gamma = 0$, respectively). These characterizations will allow us to tackle the problem of the external agent’s optimal behaviour.²¹

As is usual in repeated games, the threat of punishment helps to sustain cooperation. In our model, exchanging information modifies the threat of punishment for some of the agents. This could lead to higher cooperation within the organization. However, such information exchanges come at the cost of increasing the information leakage cost of the organization because they may expand the sets $\{V_i\}_{i=1}^N$. In this section, we study how this trade-off affects the organization’s optimal structure.

In our characterizations of the optimal information structures, it is often useful to begin by focusing on optimal structures *given a fixed number of agents “linked” to other agents*—that is, a fixed number of agents that disclose their information to at least one other agent. Note that the benefits of having agents linked depend only on their number rather than on the structure of the organization. In particular, since an agent cannot be punished more harshly by revealing his information to more than one agent (see the assumptions in Section 2.3.2), the potential benefit that the links can yield to the organization is constant with respect to all the information structures with the same number of agents linked to someone else. As a consequence, we obtain the following Lemma, the proof of which appears in the Appendix.

Lemma 1 *(i) Any optimal information structure is equivalent to another organization in which each agent reveals his information to at most one other agent. (ii) If the number of linked agents is fixed, an optimal organization minimizes the cost of information leakage.*

Using this result, we argue that the optimal information structure can take one of only

²¹Note that even though we assume that the external authority determines the level of scrutiny, these probabilities could also be exogenously given and due to some intrinsic characteristics of the agents. For example, some agents may be more talented in evading detection (some may have a cleaner criminal record or simply might be able to run faster). If this is the case, the optimal organization characterizations we provide in the next sections can be seen as self-contained.

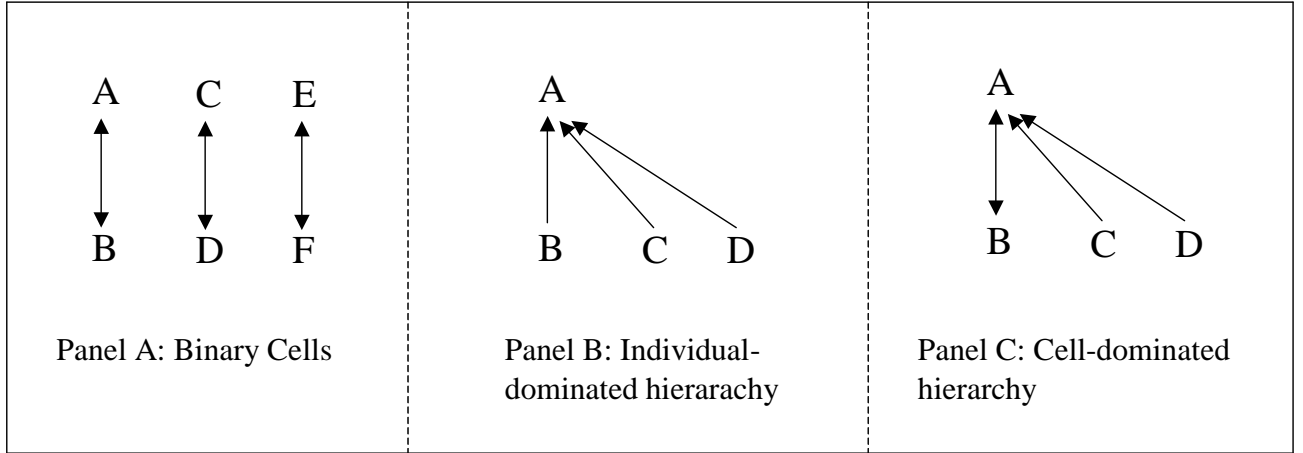


Figure 2: Optimal Structures

a small number of possible forms, as described in Proposition 1.

Proposition 1 *An optimal information structure can consist only of (i) unlinked agents (ii) pairs of agents linked to each other but to no others in the organization (binary cells) and (iii) at most one hierarchy where the information hub is either a single agent or a binary cell.*

Proposition 1 guarantees that, besides the possibility of including some agents not linked to anybody else (as in (i) of Proposition 1), an optimal information structure comprises structures of the forms illustrated in Panels A, B and C of Figure 2, or a combination of them (including at most one of the forms illustrated in Panel B or C).

Panel A illustrates a set of agents organized in binary cells. Both Panel B and Panel C illustrate hierarchical structures. Panel B illustrates a hierarchy dominated by a single individual, while Panel C illustrates a hierarchy dominated by a binary cell. In a hierarchy, we refer to the agent(s) who holds information about the other agents as the “information hub”. For example, in the hierarchy $i, j, k \rightarrow l$, the information hub is l . In a cell-dominated hierarchy, the pair who hold information are the information hub. Note that there are many equivalent representations for a cell-dominated hub, for example the three structures in Figure 3 allow for identical payoffs, and where 1 and 2 constitute the information hub.

Let us proceed to the intuition of Proposition 1. First of all, note that if δ is very high, cooperation in the stage game can be sustained without the threat of additional

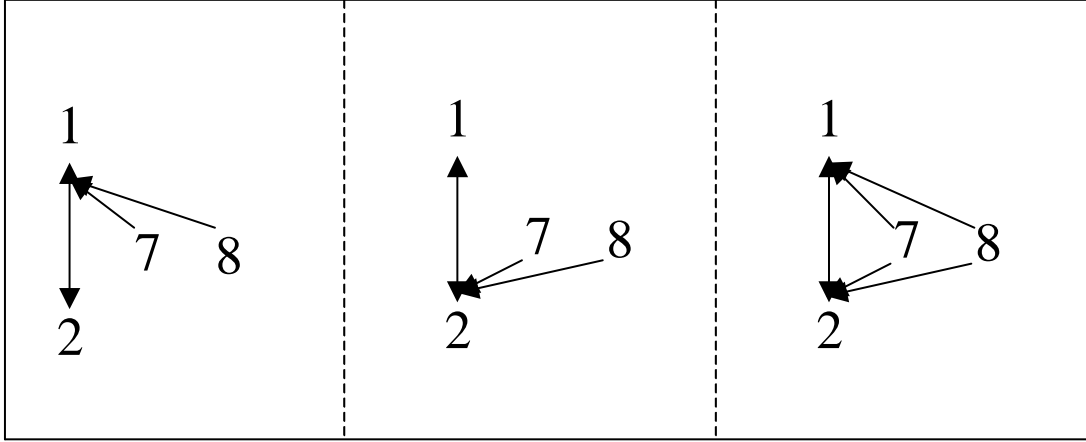


Figure 3: Cell-dominated hierarchy: equivalent representations.

punishments. Thus, introducing links in the organization does not yield any benefit but may have a positive information leakage cost. This implies that for sufficiently high δ , an anarchy (an organization with no information links) is the optimal information structure. Similarly, if δ is very low, the additional punishment threat k is not enough to induce agents to cooperate. Thus, as information links yield no benefit and (weakly) positive costs, an anarchy is the optimal information structure.

Note that it is easy to rule out the optimality of a number of structures. In particular, it is clear that any structure that contains a chain of the form $i \rightarrow j \rightarrow k$ is dominated by replacing the link $i \rightarrow j$ with a link $i \rightarrow k$. Since under $i \rightarrow j \rightarrow k$, agent i is vulnerable to k in any case, the only effect of such a change is to ensure that agent i (and any agents vulnerable to i) are no longer vulnerable to agent j getting detected. Similarly, a structure which contains a chain of the form $i \leftarrow j \leftarrow k$ is dominated by $i \leftarrow j \leftarrow k$. These observations are sufficient to ensure that the information structure can only consist of an “individual-dominated hierarchy” where a number of agents are linked to a single agent as in Panel B of Figure 2, a “cell-dominated hierarchy” as in Panel C and binary cells as in Panel A.

Next, we argue that an optimal information structure cannot include more than one hierarchy. Suppose that there are two hierarchies, where S_1 is the set of subordinates in the first hierarchy and S_2 is the set of subordinates in the second hierarchy, and the sets H_1 and H_2 comprise the information hubs for the two hierarchies (so that H_i may either

be a singleton or a binary cell). We can consider the total probability of detection of a hub. If $H_1 = \{j\}$ this is simply $\gamma\beta_i$ (note that the singleton hub would not be cooperating), whereas if $H_1 = \{j, k\}$ the total probability of detection is the probability that either of the agents i and j is detected: $2(\beta_i + \beta_j - \beta_i\beta_j)$ —note that the agents in the binary cell at the hub would cooperate. By assigning all the subordinates $S_1 \cup S_2$ to whichever hub has the lower probability of detection, clearly reduces the information leakage costs. It follows that it cannot be strictly optimal for the information structure to include more than one hierarchy, concluding the proof of Proposition 1.

In the case where all N agents are linked and induced to cooperate, then $\alpha_i = \beta_i$ for all i (and in particular α_i does not depend on the information structure). In this case, we can obtain a full characterization of the optimal information structure.

To start the characterization, consider a cell $\{i, j\}$. Let $\rho(i, j) \equiv \frac{2(1-\alpha_i)(1-\alpha_j)}{2-\alpha_i-\alpha_j}$. This is a useful ratio in understanding agents' proclivity to be linked as a binary cell rather than as subordinates to another agent. If two agents $\{i, j\}$ are in a cell, each of them *will not* pay b with probability $(1 - \alpha_i)(1 - \alpha_j)$. On the other hand, if each of them is independently linked to a third agent (the same for both, and who may be linked to others) with *overall probability of avoiding detection* α' , agent i will not pay b with probability $\alpha'(1 - \alpha_i)$, and agent j will not pay b with probability $\alpha'(1 - \alpha_j)$. Then, having the agents $\{i, j\}$ forming an independent cell rather than linking each of them to the third agent minimizes the cost of information leakage if and only if

$$2(1 - \alpha_i)(1 - \alpha_j) > \alpha'(1 - \alpha_i) + \alpha'(1 - \alpha_j), \quad (2)$$

or, equivalently,

$$\rho(i, j) = \frac{2(1 - \alpha_i)(1 - \alpha_j)}{2 - \alpha_i - \alpha_j} > \alpha'. \quad (3)$$

Thus, for any couple of agents, the higher is $\rho(i, j)$, the greater is the advantage of forming a cell rather than being linked to a third agent. *Notice that $\rho(i, j)$ is decreasing in both α_i and α_j* —that is, the higher the probability of detection of an agent, the lower $\rho(i, j)$ of the cell to which he belongs.

We now characterize the optimal information structure with N linked agents in the following proposition, the proof of which appears in the appendix.

Proposition 2 *The optimal information structure with N linked agents is described as follows. Let $i^* \in \{2, \dots, N\}$ be the largest even integer such that $\rho(i-1, i) > (1 - \alpha_1)(1 - \alpha_2)$ (if no such integer exists, set $i^* = 1$): all the agents $i = 1, \dots, i^*$ are arranged in binary cells as $1 \leftrightarrow 2, 3 \longleftrightarrow 4, \dots, i^* - 1 \longleftrightarrow i^*$ and the agents $i = i^* + 1, \dots, N$ all reveal their information to agent 1, that is, $i^* + 1, \dots, N \rightarrow 1$.*

Proposition 2 states that the optimal way to link N agents in an organization is to divide the agents in two groups according to their probabilities of detection: a group comprising the i^* agents with the lowest probabilities of detection, and another group with the $N - i^*$ agents with the highest probability of detection. The agents belonging to the first group are arranged in binary cells formed by agents with adjacent probability of detection (i.e. $1 \leftrightarrow 2, 3 \longleftrightarrow 4, \dots, i^* - 1 \longleftrightarrow i^*$). All the agents belonging to the second group reveal their information to agent 1 ($i^* + 1, \dots, N \rightarrow 1 \leftrightarrow 2$).²²

The number of agents i^* belonging to the independent cell component depends on how steeply the ratio $\rho(i, i + 1)$ of each couple grows. If α_1 and α_2 are very low relative to the other agents' probabilities of detection, it could be the case that $\rho(i - 1, i) < (1 - \alpha_1)(1 - \alpha_2)$ for all $i = 4, \dots, N$. In this case, Proposition 2 requires that an optimizing organization links all the agents $3, \dots, N$ to agent 1 (who remains linked in a cell with agent 2).²³ On the other hand, if α_3 and α_4 are close enough to α_2 , then $\rho(3, 4) > (1 - \alpha_1)(1 - \alpha_2)$, and Proposition 2 prescribes agents 3 and 4 to form a cell rather than being linked to both agents 2 and 1, and so on.

The optimal information structure described in Proposition 2 is illustrated in Figure 4, when there are $N = 8$ agents and $i^* = 6$.

Finally, Proposition 2 implies that if either agent 1 or agent 2 (or both) are detected, the lowest ranks of the organization (i.e., the agents with the highest probabilities of detection) are detected as well but it is possible that relatively high ranks of the organization, organized in cells, remain undetected.

Note that with a full characterization of the form of optimal structures, we could easily calculate the information leakage of linking n agents and inducing them to cooperate. If agents can be induced to cooperate if and only if linked, the per-period benefit of each link is $N\lambda - c$. Comparing costs and benefits of links then allows us to characterize the optimal

²²Note that, because agents 1 and 2 form a cell, agents $i^* + 1, \dots, N$ could equivalently reveal their information to agent 2 or to both agents 1 and 2.

²³In particular, if α_1 and α_2 approach zero, all these links have an arbitrarily small information leakage cost, so the organization's information leakage cost is the same as in the structure with no links.

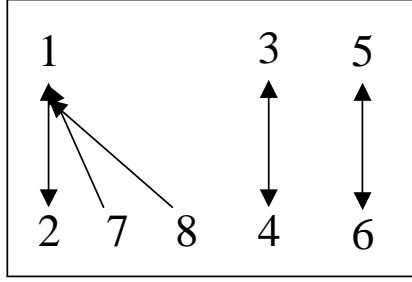


Figure 4: Case with $N = 8$ and $i^* = 6$.

information structure.

Although Proposition 1 establishes the forms that the optimal structure may take, there is still considerable latitude, for example, in establishing which agent or agents should comprise the hub. However, we can establish full characterizations when an agent’s probability of direct detection is independent of his action— $\gamma = 1$ or agent-based detection—and when an agent cannot be detected directly if not cooperating— $\gamma = 0$ or cooperation-based detection. These complete characterizations allow us to establish the external authority’s optimal strategy in these cases.

4 Agent-Based Detection

We begin by taking the allocation of detection probabilities chosen by the external agent $\{\beta_1, \dots, \beta_N\}$ as given. Since under agent-based detection $\gamma = 1$ then the probability of direct detection is identical to the extent $\alpha_i = \beta_i$ regardless of the agent’s behaviour. We first identify the most efficient information structure that the members of the organization can form. Given this characterization, in Section 4.1, we step back and study the external agent’s optimal behaviour.

Note that, the optimal structure when all N agents are linked is characterized in Proposition 2. It remains to characterize the optimal structure for $n < N$.

Lemma 2 *If $\gamma = 1$, the optimal information structure with $n < N$ linked agents is a hierarchy with the agent with the lowest probability of detection at the top of the hierarchy and the n agents with the highest probabilities of detection linked to him (i.e., $N, N -$*

$1, \dots, N - n + 1 \rightarrow 1$).²⁴

If the number of linked agents is less than N , the optimal structure is simply an individual-dominated hierarchy, in which the information hub is the member with lowest probability of detection and the $n < N$ linked agents are those with the n highest probability of detection. The proof is very simple. Suppose first that $n = 1$. We need to find the way to generate the “cheapest” possible link in terms of information leakage costs. The only event in which this link becomes costly is the case in which agent i is independently detected and agent j is not. This event has probability $\beta_i(1 - \beta_j)$. The cost of the link is minimized when β_i is as small as possible and β_j is as large as possible. It follows that the “cheapest” possible link is the one that requires agent N to disclose his information to agent 1 (the link $N \rightarrow 1$). If $n = 2$, the second cheapest link one can generate after $N \rightarrow 1$ is $N - 1 \rightarrow 1$, and so on. Notice that Lemma 2 implies that the information leakage cost under an optimal structure in which there are $n < N$ links is simply $b\beta_1 \sum_{i=1}^n (1 - \beta_{N-i+1}) + b \sum_{i=1}^N \beta_i$.

Lemma 2 and Proposition 2 allow us to allow us to define the information leakage cost function $C : \{0, \dots, N\} \rightarrow R$ under agent-based detection as follows

$$C(n) = \begin{cases} b \sum_{i=1}^N \beta_i & n = 0 \\ b \sum_{i=1}^N \beta_i + b\beta_1 \sum_{j=N-n+1}^N (1 - \beta_j) & n = 1, \dots, N - 1 \\ b \sum_{i=1}^N \beta_i + b(\beta_1 + \beta_2 - \beta_1\beta_2) \sum_{i=i^*+1}^N (1 - \beta_i) \\ \quad + b \sum_{i=1}^{\frac{i^*}{2}} [(1 - \beta_{2i-1})\beta_{2i} + (1 - \beta_{2i})\beta_{2i-1}] & n = N \end{cases}$$

On the benefit side, note that for any δ there is a minimal number of agents \underline{m} that must be linked to someone else in order for a link to induce cooperation and generate a fixed benefit $N\lambda - c$.²⁵ Thus, we can characterize the benefit function $B : \{0, \dots, N\} \rightarrow R$ as follows.

²⁴This characterization applies more generally. In particular, it applies when $\gamma > \frac{1}{1+\beta_1}$.

²⁵ \underline{m} is the minimal number of links sufficient to induce cooperation from the agents who reveal their information to someone else. More specifically, it is easy to see that \underline{m} is the smallest integer m for which $\delta \geq \frac{c-\lambda}{\lambda(m-1)+k}$.

$$B(n) = \begin{cases} 0 & 1 \leq n < \underline{m} \\ n(N\lambda - c) & n \geq \underline{m} \end{cases} \quad (4)$$

With costs and benefits established, we can find the number of links n^* that maximizes the value of the organization, or equivalently maximizes $B(n) - C(n)$. The following proposition characterizes the optimal information structure for $\gamma = 1$.

Proposition 3 *If $\gamma = 1$, the optimal information structure is as follows: (i) If $n^* = 0$ the optimal information structure is an anarchy (ii) If $0 < n^* < N$ the optimal structure is an individual-dominated hierarchy where the hub is agent 1 and the subordinates are agents $N, \dots, N - n + 1$. (iii) Finally if $n^* = N$, the optimal structure is as described in Proposition 2.*

Note that when $\beta_i = \beta$ for all i , the additional cost of each link is constant and equal to $b\beta(1 - \beta)$. It follows that in this symmetric case, the optimal structure either consists of no links or all agents are in of binary cells.

Corollary 1 *Let $\gamma = 1$ and $\beta_i = \beta$ for all i . If $\lambda N - c > b\beta(1 - \beta)$ then the optimal information structure is a binary cell structure. Otherwise, the optimal information structure has no links.*

This concludes the characterization of the optimal information structure for a given scrutiny distribution $\{\beta_1, \dots, \beta_N\}$. Next, we endogenize scrutiny and discuss the strategy of the external agent.

4.1 Agent-Based Detection: The External Authority

As discussed in section 2.4.2, we assume that the external authority's objective is to minimize the number of agents who cooperate—that is, the organization's production level.

Recall that the problem of the external agent is to allocate $B \in [0, N]$ to determine the scrutiny $\beta_i \in [0, 1]$ of each agent i such that $\sum_{i=1}^N \beta_i \leq B$. The external authority acts first and chooses these scrutiny levels before the organization forms.

In the next result, we characterize the (weakly) optimal strategy for the external authority to determine how to allocate its resources.²⁶ Note that, if the authority allocates

²⁶This strategy is weakly optimal because there are some value for δ such that the strategy of the external agent is irrelevant, and there may be other strategies that achieve the same result.

the same budget β to each agent, the cost of each link becomes $b\beta(1 - \beta)$. Since this cost is maximized at $\beta = \frac{1}{2}$, it is never optimal to set $\beta > \frac{1}{2}$ in a symmetric allocation. Let then $\widehat{\beta} = \min\{\frac{B}{N}, \frac{1}{2}\}$ be the optimal symmetric allocation.

Proposition 4 *Let $\gamma = 1$. A weakly optimal strategy for the external agent is to set scrutiny symmetrically if $b\widehat{\beta}(1 - \widehat{\beta}) > N\lambda - c$ and to not investigate one agent and detect all others symmetrically (set $\beta_1 = 0$ and $\beta_2 = \dots = \beta_N = \min\{\frac{B}{N-1}, 1\}$) otherwise.*

A symmetric allocation can prevent the formation of any link if the cost of each link $b\widehat{\beta}(1 - \widehat{\beta})$ is greater than the potential benefit of individual cooperation. This is the case when $b\widehat{\beta}(1 - \widehat{\beta}) > N\lambda - c$, and, in these circumstances, a symmetric allocation is optimal as it deters any cooperation.

However, if $b\widehat{\beta}(1 - \widehat{\beta}) < N\lambda - c$, by Lemma 1, a symmetric allocation would yield the formation of a binary cell structure that reaches full efficiency. The question is whether, in these situations, the external agent can do something else to prevent full efficiency. Proposition 4 addresses this question and suggests that, in this case, an allocation in which one agent remains undetected and the budget is equally divided into the other $N - 1$ agents is optimal. Under this allocation, sometimes the organization still reaches full efficiency (in this case, we can conclude that the external agent cannot prevent full efficiency to occur), but in some cases, a hierarchy with $N - 1$ links arises. Since the hierarchy is strictly less efficient than a binary cell structure, this allocation strictly dominates the symmetric one.

If $b\widehat{\beta}(1 - \widehat{\beta}) > N\lambda - c$, we show that there is no other allocation that strictly dominates $\beta_1 = 0$ and $\beta_2 = \dots = \beta_N = \min\{\frac{B}{N-1}, 1\}$.²⁷ The intuition for this part of Proposition 4 is the following. First of all, notice that if two agents remain undetected ($\beta_1 = \beta_2 = 0$), the organization can form N links without incurring any additional information leakage costs with respect to the cost they would incur with no links (this is because the two agents can reveal information to each other at no cost and costlessly act as a hub for the $N - 2$ agents). So, to deter full efficiency, the external agent can leave at most one agent undetected. Suppose now that some cooperation is deterred by an allocation in which all agents are detected with some probability ($\beta_1 > 0$). Then, the agent with the lowest allocation will act as a hub in a hierarchy, as described in Proposition 2. In the Appendix, we prove that under our assumption, there are exactly $N - 1$ links in such a

²⁷Note that considering $\min\{\frac{B}{N-1}, 1\}$ guarantess that the resulting allocation on each agent is in the interval $[0, 1]$.

hierarchy. Then, moving all the resources from the hub to the other agents, as suggested in Proposition 4, is equivalent to the original allocation.

Let us turn to comment the cooperation outcomes in equilibrium. Proposition 4 states that in some circumstances (i.e., if $b\widehat{\beta}(1 - \widehat{\beta}) \geq N\lambda - c$), the external authority can prevent any cooperation in the organization by allocating its budget symmetrically. Note that these circumstances are more likely to occur when b is higher, or the benefit from cooperation $N\lambda - c$ is lower. On the other hand, a higher budget B is beneficial as long as it is below the threshold $\frac{N}{2}$. Increases in B beyond that threshold would not affect the cooperation level in the organization any longer.

On the other hand, when $b\widehat{\beta}(1 - \widehat{\beta}) < N\lambda - c$, the optimal detection strategy for the authority has the outcome to deter *at most* the cooperation from one agent in the organization.

5 Cooperation-based Detection

In this section, we turn to the case of cooperation-based detection. As before, we begin by taking the allocation of detection probabilities chosen by the external agent $\{\beta_1, \dots, \beta_N\}$ as given. Under cooperation-based detection $\gamma = 0$, an agent i is detected with probability β_i if he cooperates and with zero probability if he does not cooperate. This case is appropriate to describe situations in which the organization's daily activity consists of an illegal activity, such as drug trading or gambling, which the external authority is able to detect directly.

In comparison to the agent-based detection case, this modification has two effects. First, as cooperation increases the probability of detection, the costs and benefits of creating links in the organizations cannot be studied separately. Indeed, since cooperation increases the risk of information leakage, full cooperation from all the agents that are linked to someone is not necessarily optimal.

Second, in this case centralization is more desirable. This is because concentrating all the information in the hands of one agent who does not cooperate makes any increase in cooperation of the other agents less costly from an information leakage point of view.

Following the analysis in Section 4 above, we now characterize the optimal information structure given a fixed number n of linked and cooperating agents. While for the $n = N$ case Proposition 2 still applies, in the next result (proved in the Appendix), we characterize of the optimal structure for $n < N$ linked agents.

Lemma 3 *If $\gamma = 0$, a weakly optimal information structure with $n < N$ linked and cooperating agents links the n agents least likely to be detected to agent N . In this organization, the agents $1, \dots, n$ cooperate and all other agents do not.*

Note that for $n < N$ linked agents, the structure described is weakly optimal: linking the n cooperating agents to any of the non-cooperating $N - n$ agents is equivalent since none of the non-cooperating agents is ever detected.

The result of Lemma 3 for $n < N$ links is in stark contrast to the agent-based detection model. In the agent-based detection model if $n < N$ agents are linked, the agents with the highest probabilities of detection are linked to the agent with the lowest probability of detection. In contrast here, it is those with the lowest probability of detection if cooperating who are linked to an agent under higher scrutiny.

The reason is that in the agent-based detection case, the cost of a link $i \rightarrow j$ is solely associated with making agent i vulnerable to agent j 's detection—this cost is minimized if agent i is already under considerable scrutiny and the scrutiny of agent j is relatively low. In contrast, in the cooperation-based detection model, since the link $i \rightarrow j$ makes agent i cooperate, such link exposes agent i to direct detection. The cost associated with direct detection is high if agent i is under considerable scrutiny, that is if β_i is high. On the other hand, when $n < N$, there are agents who are not linked to anybody else, and they are detected with probability zero. Thus, if j is among those not linked to anybody, the link $i \rightarrow j$ does not increase the probabilities of indirect detection for i .

Given the characterization in Lemma 3 and Proposition 2, we can proceed to discuss the optimal number of links. Similarly to how we proceeded in Section 4, we can denote the information leakage costs by $\tilde{C}(\cdot)$. Note that, given the characterization in Proposition 3, we have:

$$\tilde{C}(n) = \begin{cases} 0 & n = 0 \\ b \sum_{i=1}^n \beta_i & n = 1, \dots, N - 1 \\ b \sum_{i=1}^N \beta_i + b(\beta_1 + \beta_2 - \beta_1\beta_2) \sum_{i=i^*+1}^N (1 - \alpha_i) & n = N \\ + b \sum_{i=1}^{\frac{i^*}{2}} [(1 - \beta_{2i-1})\beta_{2i} + (1 - \beta_{2i})\beta_{2i-1}] & \end{cases}$$

We define the benefit function $\tilde{B} : \{0, \dots, N\} \rightarrow \mathbb{R}$ similarly to how we defined it in

the agent-based detection case.²⁸ Define \tilde{n}^* as the optimal number of links: that is the value that maximizes $\tilde{B}(n) - \tilde{C}(n)$. The following proposition characterizes the optimal organization.

Proposition 5 *If $\gamma = 0$, it is optimal for the organization to form the following structure (i) an anarchy if $\tilde{n}^* = 0$ (ii) If $1 \leq \tilde{n}^* < N$, a hierarchical structure in which agents the \tilde{n}^* agents under least scrutiny ($1, \dots, \tilde{n}^*$) are linked to some agent under great scrutiny ($\tilde{n}^* + 1, \dots, N$) (2) if $\tilde{n}^* = N$, a structure as described in Proposition 2.*

In particular, and in contrast with the case of agent-based detection, even with all agents *ex-ante* symmetric, an asymmetric information structure can be optimal. We show this in the next corollary, which we prove in the Appendix.

Corollary 2 *If $\beta_i = \beta$ for all i , then the the optimal information structure is either (i) an anarchy, or (ii) a structure in which $N - 1$ agents are linked to the remaining agent (ii) a binary cell structure.*

5.1 Cooperation-based detection: The External Authority

We turn to examine the strategy of the external authority who sets $\{\beta_1, \dots, \beta_N\}$ before the information exchange and aims to minimize the level of cooperation within the organization.

First note that, as for the agent-based detection case, there are circumstances under which a symmetric allocation of the detection budget prevents all agents in the organization from cooperating. However, whereas in agent-based detection, under the symmetric allocation, verifying the cost-benefit trade-off of a single link is sufficient to determine whether the organization would link all agents, that is no longer true in the case of cooperation-based detection. Indeed, as we proved in Corollary 2, a symmetric allocation of the detection budget could induce asymmetric information structures. Also, while under the agent-based detection model a link $i \rightarrow j$ could be costless for any β_i if $\beta_j = 0$, in the cooperation-based detection model the link $i \rightarrow j$ increases the probability of i 's direct detection and thus is never lower than $b\beta_i$. The external authority could benefit from this fact, as outlined in Lemma 4.

²⁸In particular, there will be threshold values which depend on the number of agents induced to cooperate, on the level of scrutiny and on the other parameters of the model, that will ensure some agents cooperate. From each of those cooperating agents the benefit is $N\lambda - c$. Since the structure is fully characterised, one can determine which agents will be linked for a given number of links n and derive an aggregate benefit function $\tilde{B}(n)$.

Lemma 4 *If $\gamma = 0$, by allocating $\beta_i = \frac{N\lambda - c}{b}$, the external authority can always prevent agent i from cooperating.*

Using this result, we can move on to the characterization of the optimal strategy for the external authority.

Proposition 6 *If $\gamma = 0$, then an optimal strategy for the external agent is to allocate the budget $\beta_i = \frac{N\lambda - c}{b}$ among $\left\lfloor B \frac{b}{N\lambda - c} \right\rfloor$ agents, to allocate the residual budget to one of them, and to allocate nothing to detecting the other agents.*

Suppose $B \geq 2\frac{N\lambda - c}{b}$. In this case, it is possible to allocate $\beta_i, \beta_j = \frac{N\lambda - c}{b}$ to at least two agents i and j and preventing them from cooperating (as Lemma 4 guarantees). With some agent not cooperating, if there is some agent k with $\beta_k < \frac{N\lambda - c}{b}$ then agent can be linked to the non-cooperating agent and be induced to cooperate in equilibrium. It follows that allocating at least $\frac{N\lambda - c}{b}$ to as many agents as possible is the optimal strategy.²⁹

Let us turn to comment the cooperation outcomes in equilibrium. Proposition 6 states that the optimal strategy of the external agent deters cooperation from $\left\lfloor B \frac{b}{N\lambda - c} \right\rfloor$ agents. Thus, more cooperation is deterred the higher are B and b and the lower is $N\lambda - c$. Note that, differently from the agent-based detection case, a high enough increase in the budget B always decreases equilibrium cooperation.

6 Comparing Agent- and Cooperation-based Detection

The agent-based and cooperation-based detection models explored in Sections 4 and 5 are special cases of our more general model. These two models are approximations to different criminal environments, which might vary with the nature of the crime or jurisdiction and the extent to which the law or public opinion takes a stand in defining what kind of detection and enforcement is acceptable.

Under cooperation-based detection, an agent cannot be detected directly if not cooperating with the organization. In other words, under the cooperation-based detection model, the authorities cannot detain a suspect and induce him to give them information about the rest of the organization unless this suspect has been observed engaging directly in criminal activities. Constitutional laws, if present, typically restrict the authorities to abide to such limitation.

²⁹ See the Appendix for the $B < 2\frac{N\lambda - c}{b}$ case.

When limited to cooperation-based detection, authorities can still decide whether to monitor everybody in the same way, or to submit individuals to different levels of scrutiny. As an example, some criminal organizations, such as Mafias, carry out illegal day-to-day activities such as drug dealing, gambling, etc. In these environments, it is typically less costly and controversial for the law enforcement to screen the territory for illegal activities, which is a cooperation-based detection strategy in which every individual is under the same level of scrutiny (i.e., $\gamma = 0$ and $\beta_1 = \dots = \beta_N$). On the other hand, it is possible to have situation in which individuals are under different levels of scrutiny but none of them can be detained unless proven to be directly involved in some crime (i.e., $\gamma = 0$, asymmetric $\{\beta_i\}$ distribution).

On the other hand, the agent-based cooperation model is a situation in which the probability with which an agent is detected directly is independent of the agent's behaviour and depends only on the external authority's level of scrutiny. These are situations in which, because of absence of constitutional boundaries, and because of the nature of the activity at hand, the authorities can detain an agent and force (or convince) him to reveal information without necessarily proving him to be involved in criminal activities.

As an example, many terror organizations carry out ostensibly legal day-to-day activities, for instance, in preparation of an illegal plan (e.g., flying lessons, phone conversations, and meetings in preparation for terrorist attacks) and also operate in areas of extreme conflict (similar considerations apply to insurgency or revolutionary groups). In these situations, because of the different legal standards that are applied, the agent-based detection model may be more appropriate.

Optimal Information Structure The analysis we carried out above allows us the following comparison between the optimal information structure we should observe under the two alternative detection environments.

Remark 7 (i) *The binary cell structure can arise as optimal structure under both agent-based detection and cooperation-based detection.* (ii) *When all agents are treated symmetrically by the authorities ($\beta_1 = \dots = \beta_N$), either a binary-cell structure or an anarchy arise in the agent-based detection model, while in the cooperation-based detection model, anarchy, binary cell structure and hierarchy can arise.*

There are two messages to be learned from Remark 7. First of all, it highlights *the robustness of the binary cell structure as an optimal organization*, as it can arise in equilib-

rium in both models of detection. Second, in an organization that is subject to *symmetric* agent-based detection (such as a terrorist organization), this is *the only alternative to no links*. However, in an organization that is subject to cooperation-based detection (such as the Mafia and traditional organized crime), *a hierarchy can be optimal as well*. This suggests that in the cooperation-based model there is a tendency toward a centralized information structure.

Optimal detection Recall that, under agent-based detection, a link $i \rightarrow j$ does not change the probability that i is detected directly, but makes agent i vulnerable to agent j 's detection. Here, the cost of a link $i \rightarrow j$ depends on the level of scrutiny on agent j , regardless of the whether agent j cooperates or not. In contrast, under cooperation-based detection, the link $i \rightarrow j$, by making agent i cooperate, always has the cost of making agent i more vulnerable to direct detection (while the risk of indirect detection increases *only if* agent j cooperates as well). This cost depends on the level of scrutiny on agent i .

In turn, the optimal strategy for the external authority under agent-based detection is designed to make it unattractive to link *to* agents by allocating the budget symmetrically among all (or among all but one) agents. However, in the cooperation-based detection the optimal strategy is designed to make it unattractive to link *an* agent (by scrutinizing as many agents as possible at a level high enough to discourage them from cooperating).

We think that this is a useful normative insight. In situations in which constitutional limits prevent the detection of non-cooperating agents, the authority's focus should be to directly discourage people from cooperating by maintaining a *sufficient* level of scrutiny on as many agent as possible. On the other hand, when these limitations do not exist, or the daily activities of the criminal organization are legal, spreading the budget symmetrically is likely to be a better approach.

7 Discussion

In the model presented above, we have made a number of strong assumptions, in part for analytical tractability and in part to highlight some effects more clearly. We conclude the paper with a fairly wide-ranging discussion of the assumptions and other potential extensions of this work.

7.1 Motivation of the assumptions

7.1.1 Timing

Throughout, we have assumed that the external authority decides on the level of scrutiny and its strategy is realized before the formation of any links. Such an assumption can be justified on the grounds that law enforcement policies and investigating budgets are broadly laid out and are hard to fine-tune once a certain policy is in place.

In some circumstances, however, it may be more plausible to suppose that the external authority can modify its policy after the information structure has formed (making the strong assumption that the authority can fully observe the information structure μ). In this case, when looking for an equilibrium of the game, we have to worry whether the policy set ex-ante is credible—that is, it is also *ex-post optimal*. Recall that in the agent-based detection model, the optimal allocation policies we characterized in Proposition 4 are either a symmetric allocation (i.e., $\beta_1 = \dots = \beta_N = B/N$) or an allocation in which one agent remains undetected and the others are monitored symmetrically (i.e., $\beta_1 = 0$ and $\beta_2 = \dots = \beta_N = \frac{B}{N-1}$). While the first allocation, when optimal, induces the organization to form no links, the second allocation, when optimal, can induce a hierarchy in which agent 1 holds all the information. It is easy to see that the only allocation optimal ex-post is the symmetric one. This is because if the allocation is asymmetric and a hierarchy emerges, the authority has the incentive to reshuffle all its resources to the information hub of the organization (that is, the agent who was left initially undetected).

Another possibility is that the authority might choose its strategy first, the organization then forms, and the authority’s strategy is then realized. If the organization’s chooses a pure strategy, this timing assumption leads to equivalent results to those considered in this paper. However, if the authority chooses mixed strategies then this timing assumption is substantive and can lead to more effective strategies for the external authority. For example, suppose there are only two agents in the organization. Randomizing between choosing scrutinies of $(0, 1)$ and $(1, 0)$ guarantees that both members of a cell are detected and so is more likely to deter the creation of a cell than a symmetric allocation of $(\frac{1}{2}, \frac{1}{2})$.

7.1.2 Observation Structures

In our model we assumed that each agent can perfectly observe all the other agent’s past actions. Note that this assumption fits a context in which the goal of the organization

is to carry out plans whose success or failure is observable by all the members of the organization. Indeed, just the observation of the outcome of the plan would indicate whether some members deviated from cooperation, and did not perform the task they were assigned to. This could lead to the breakdown of trust within the organization and so result in all its members stopping to cooperate. Of course, a plan's success or failure is influenced by many external noisy variables as well, but as long as these variables are commonly observed, the qualitative analysis of this paper would not change.

However, an interesting alternative modelling choice could be that *the information structure*, rather than allowing for increased punishments, *affects agent's ability to monitor each other's actions*. More precisely, consider a situation in which agents can observe only the actions of the agents they hold information about, and $k = 0$ (there are no additional punishments). Suppose also that the agents cannot deduce the cooperation of the other agents from their own payoffs (for instance, because the payoffs are distributed at a later period in time). Finally, assume that there is *no communication* among agents, that is, if agent j observes the actions of agent i ($i \rightarrow j$), agent j cannot report agent i 's behavior to the other agents.³⁰

Observe that, in this alternative specification of the model, *the analysis of the costs of introducing links in organization (which is the focus of our attention) is unchanged with respect to our model*. On the benefit side, our model has the advantage of making the benefit of introducing a link to a large extent independent of the information structure. This allowed us to state Lemma 1, which guarantees that the cost-minimizing structure is the optimal structure for a given number of links n . In this alternative specification the benefit of a link could depend on the information structure as well. Thus, to determine the optimal information structure for a given number of links n , one should go over all the possible information structures and determine which is the one that maximizes the value of the organization. This makes the analysis significantly more complicated.

However, some considerations allow us to conjecture that the optimality of some of the information structures we characterized above holds in this alternative model as well.

³⁰Suppose instead that we allow for communication to take place, that is, each agent observing other agents' actions is allowed to report what he has seen at everybody else at the end of each period. With this modification, the benefit part of our model is connected to Ben-Porath and Kahneman (1996). They show that having two agents observing each agent is enough to guarantee the existence of efficient equilibria when δ tends to one. Thus, for high δ , the intuition of our results holds in this framework as well, since the introduction of links yields more cooperation in the organizations. The fact that each agent must be connected to two agents in order to cooperate yields consideration similar to the ones we make in the context of Section 7.1.3.

First, note that the available punishments in the repeated game are now rescaled to be weaker than before (Nash-reversion to non-cooperation from all the other agents in case of a deviation from one agent is *always* a possible punishment in our original model, while it is only available *if all agents are directly or indirectly connected to all the other agents* in this specification). Thus, in what follows we focus on high δ , for which the analysis is comparable to the one we carried out above.³¹

Observe that linking one agent i to another one, say j , that is not linked to anybody else does not create an incentive for i to cooperate since j is never going to cooperate in the repeated game.

Thus, introducing only one link yields no benefit to the organization. If two links are introduced, the optimal way to do that is to create a binary cell. Indeed, if agents i and j are linked in a binary cell, and if δ is high enough ($\delta \geq \frac{c-\lambda}{\lambda}$), these agents have an incentive to cooperate (note that a necessary condition for this to happen is $c \leq 2\lambda$).

If one wants to introduce a third link, one way to do that is to link a third agent, say h , to an agent in the $\{i, j\}$ cell.³² Moving on to the fourth link, two effective ways to introduce it are either organizing 4 agents in two binary cells, or link two agents in a hierarchy dominated by the cell $\{i, j\}$, and so on.

All this suggests that the characterization of the optimal organization in Proposition 2 holds in this alternative environment as well. In particular, there are conditions under which a binary cell structure is the optimal organization structure and conditions under which a mixed information structure is optimal. The only information structures that we find optimal in the analysis above that will never arise in this environment are the ones in which a recipient of a link is not linked to another agent himself, as in an individual-dominated hierarchy. Indeed, if this is the case, cooperation from the hub cannot be enforced, and cooperation will unravel throughout the organization. Note however that moving to a hub-dominated hierarchy always solves this problem. This suggests that in the cases in which individual-dominated hierarchies were found to be optimal in our model (i.e., in all cases in which $n^* < N$), either a smaller mixed hierarchy or a partial binary cell structure will be optimal in this context.

³¹Note that for lower δ for cooperation to arise the same agent has to reveal his information to more than one agent. This could lead to larger cells, or multi-headed hierarchies similar to the structures we described in Section 7.1.3.

³²Suppose $i \rightarrow j \rightarrow k$ and k observes that j is not cooperating. This could be caused by a deviation by j or by a punishment that j is inflicting on i . In either cases, if the equilibrium prescribes for k to stop cooperating as well, this generates an incentive for i to cooperate for even lower δ .

7.1.3 Harsher Punishments

Our analysis assumes that if many agents have information about one agent and decide to punish him, then the agent suffers as if only one agent had decided to punish him (that is, he pays only k). There are circumstances in which, if an agent becomes vulnerable to more than one other agent, he can be punished in a harsher way. Such punishment technologies lead to trade-offs and optimal structures that are very similar to the ones we study in this paper, with the exception that the cells would include a larger set of agents that exchange information about each other to sustain cooperation within the cell. Moreover, hierarchies would be characterized by a set of agents that all reveal their information to the same multiple agents, leading to a multi-headed hierarchy. For more details, we refer to the discussion on general link benefits we carry out in Section 7.2.1 for which this extension is a particular case.

7.1.4 Disruption of the Organization

In our model we assume that every time the external agent detects an agent (directly or indirectly), the harm imposed to the organization is represented by a cost $b > 0$ paid by that agent. An easy and natural interpretation is to view b as representing the NPV of the future burdens caused by detection. This formulation is convenient and simplifies the analysis. In particular, it ensures that there is no history-dependence along the equilibrium path of the repeated game. However, this assumption implies that (i) the payment of b does not prevent the agents from present and future cooperation and (ii) the burden of the payment is imposed on the detected agent alone. Nevertheless, we argue that there are a number of ways to relax and moving away from this approach to modelling the harm from detection does not affect the qualitative properties of our results. We address some of these modifications in order of increasing distance from our model.

Cost of Detection Shared by the Organization The simplest modification of the model is to assume that b is a tax paid by the whole organization for every agent detected directly or indirectly and shared between the agent and the rest of the organization in some way (for instance, suppose that the agent detected pays some fraction λb and every other agent in the organization pays $(1 - \lambda)b/(N - 1)$ with $\lambda \in [0, 1]$). This modification would address point (ii) above. Note that, as we characterized the optimal organization from the organization's prospective, this modification would have little impact on our qualitative

results. Indeed, it has no consequences at all in the analysis of the agent-based detection model. It would only quantitatively modify the conditions that describe the incentives to cooperate (and, subsequently, the ranges of δ for which the different structures apply) in the cooperation-based detection model.

Detected Agents Removed and Replaced A further modification that addresses both points (i) and (ii) above is a situation in which all the detected agents leave the organization (for instance, because they are arrested) but each of them can be immediately replaced by a substitute brought in the same position at some cost paid by the whole organization (in a fashion similar to the one discussed in the previous case and perhaps in addition to any penalties, as in the previous case). For example, one could think of the N members of our organization as the leadership in an organization. When a member of this leadership is captured, a rank and file member is promoted and undergoes some initiation, which might include passing on and possibly receiving the kind of information that would form a link. This modification entails a similar qualitative analysis and results to our model. In particular, the possibility of disappearing from the organization can be seen as lowering the effective discount factors that agents face (and making them heterogeneous since they would now depend on each agent's probability of detection). Also, as in the case discussed in the previous paragraph, there would be the relevant parameter ranges of our characterizations in the cooperation-based detection model would change.

Detected Agents Removed Without Replacement Finally, one could address both points (i) and (ii) above by assuming that every detected agent disappears from the organization without the possibility of being replaced.³³ Analytically, this is a far from straightforward extension. The infinitely repeated cooperation game loses its recursive structure as the number of members in the organization varies with the detection and removal of agents and there is history-dependence in the information structure. This is because once an agent is linked to another (that is has given some incriminating information to another) it is hard to imagine that the link could be removed.

In particular, it is not necessarily the case that the structures that emerge from an N -member optimal organization after some members disappear would necessarily be the ones that would have been created if the organization could re-optimize for the smaller number of agents remaining. For example, consider an optimal information structure for an N -agent

³³Similar concerns arise if it takes a number of periods to replace members.

organization in which n agents are linked. For this to be optimal, we must be in a case in which the a linked agent's cooperation is sustained in equilibrium. Recall that this depends not only on the threat of punishment but also the cooperation of $n - 1$ other agents, which ceases following a defection. Suppose now that the external agent detects (either directly or indirectly) k out of the n linked agents. If the number of detected agents k is high enough, it could halt cooperation among the remaining linked agents (see the discussion in Section 4). However, adding more links to try to reach a critical level of cooperation could be ineffective in inducing cooperation if there are few remaining agents or it may be too costly from an information leakage point of view (since presumably previously unlinked agents were the most costly to link for information leakage considerations). In this situation, the information structure consisting of all the remaining agents and the remaining links is not ex-post optimal. This is because, since each link brings no benefit and increases the risk of information leakage, an anarchy would dominate this structure. and the organization would prefer these links to be deleted.

An analysis of the optimal information structure in a forward-looking organization would have to resolve the tension between ex-ante and ex-post optimality and consider the possibility that agents are detected in the future and irreversible links become suboptimal. This suggests that, in general, forward-looking optimal structures will tend to have fewer links than those characterized in the paper (because of the possibility that some of these links will become a burden to the organization later on). However, the effects that we describe in the paper will not substantially change or disappear.

7.1.5 Alternative objectives for the external authority

In our analysis of the external authority's strategy, we assume that the external authority chooses how to allocate its budget with the objective of minimizing the extent of cooperation. There are other objectives that an external authority might pursue: for example maximizing the probability of detecting an agent or maximizing the expected number of agents detected, and perhaps thereby disrupting the organization. Of course, the authority might be interested in all of these objectives to some extent.

In the case of agent-based detection, it is clear that the external authority maximizes the probability of detecting an agent $1 - \prod_{i=1}^N (1 - \beta_i)$ by assigning its budget to scrutinizing all agents symmetrically. In the case of cooperation-based detection, an interesting aspect arises. In contrast to the agent-based detection case, the symmetric allocation of the

budget need not be optimal: for example if $B > N \frac{N\lambda - c}{b}$ this strategy ensures that no agents cooperate and so no agent would ever be detected. Instead, the external authority would be better off allocating its budget in such a way as to ensure that agents cooperate so that the authority has some chance of detecting them.³⁴ Similarly, if the external authority aims to maximize the expected number of agents detected, encouraging links (in both the agent-based detection case or in the cooperation-based detection case) could help in this objective.

In the model, we have considered an external authority who allocates budget in scrutiny of different agents. However, an external authority may have other policy tools and other trade-offs to consider. For example, one might imagine that punishment (such as the building and operations of prisons) is costly for the external authority. Then the authority may face a trade-off between higher punishment on capture (that is increasing b) and the resources devoted to detection. Other policies that might be available to the authority including leniency or even rewards for whistle-blowing. An interesting extension would be to consider, how such policies if available might be used and might effect organization structure.³⁵

7.1.6 Organization formation

We have assumed that the information structure that forms maximizes the sum of the members' payoffs and we have taken a black-box approach to the process of formation, as highlighted in footnote 20. This assumption could be justified, for example by assuming a benevolent outsider who helps to organize the information structure and disappears, as has been suggested for various terrorist organizations and discussed in footnote 10. Consider a formation game in which every agent can simultaneously link to one or more other players before starting to play the repeated stage game. Note that the efficient information structure can always be sustained as an equilibrium of the formation game, by punishing the player who does not comply with the efficient structure design by reverting to non-cooperation in the repeated game.

It would be interesting, though something of a departure from the current model, to

³⁴See Ichino and Muehlheusser (forthcoming) for a related idea in the context of employers screening employees.

³⁵See Spagnolo (2007) and Harrington (forthcoming) and references therein for a currently active research program that addresses leniency in the context of cartels (though typically assuming that all members have the same information).

consider an agent or subset of agents who act as principals with the aim of maximizing their own payoffs while ensuring that others in the group meet some reservation utility. It might also be of interest to consider the organization evolving over time and allowing it to both grow and decline, in which case the issues highlighted in the last paragraph of Section 7.1.4 would apply. In such potential extensions, the interesting question of optimal organizational size might also be addressed.

7.1.7 Decaying Detection

In our model, we have assumed that if an agent is detected, then any other agent who had disclosed his information to this agent is also detected with probability 1. One possible way to relax this assumption is to assume that the information decays—that is if an agent i discloses his information to agent j , then if agent j is detected, agent i is detected with probability < 1 . This implies that if indirect links are formed, the probability of apprehension decreases with the distance between agents in the network. Notice that there is only one instance among all the optimal structures characterized above, in which an indirect link emerges. Specifically, the pair-dominated hierarchy is the only instance of a structure in which a chain might arise. Take for example, the case where N links are formed in the organization, as in Proposition 2 that prescribes a pair-dominated hierarchy with $1 \leftrightarrow 2$ and $i^* + 1, \dots, N \rightarrow 1$, that is agents $i^* + 1, \dots, N$ are indirectly linked to agent 2. It is easy to realize that a decay in the probability of detection will just cause i^* to decrease, leading to the hierarchical part of the organization to become larger. Similarly, it is easy to realize that, besides this adjustment, all our qualitative characterizations of optimal information structures are robust to this extension of the model.

7.2 Extensions

7.2.1 General Link Benefits

In this paper, we focus on the optimal information structure of self-enforcing organizations, as we believe that trust plays a key role in criminal organizations, and repeated game techniques are natural for exploring such considerations. The repeated game we model delivers a very simple benefit structure for linking agents in the organization—typically, a benefit of either 0 or $\lambda N - c$ for any linked agent. This benefit structure has the advantage of simplifying the analysis greatly and allows us to focus on the cost of a link as increasing

the organization’s vulnerability to an external threat. However, to extend this analysis to the optimal information structure of different kinds of organizations (for instance, firms, R&D departments, etc.), it would be interesting to depart from the repeated game we consider in this paper and move on to a reduced form for the links’ benefits that captures more general information technologies and could depend on the information structure itself.

In particular, one interesting direction is the following. Suppose that the benefit of the link $i \rightarrow j$ is defined as $b(n, m) \geq 0$, where n and m are the number of links *departing from* i and *arriving to* j respectively. The analysis in the paper corresponds to the case where $b(1, m) \in \{\lambda N - c, 0\}$ for any m , and $b(n, m) = 0$ for any $n > 1$ and m . However, an agent could be forced to exert more cooperation in the organization if he reveals his information to several individuals (see Sections 7.1.3). If the returns to scale from punishment are decreasing, the function b is decreasing in n . Also, a function b increasing in m could indicate benefits from coordination. We conjecture that the qualitative nature of our results and the basic trade-offs we described above hold in this setting as well.

In particular, if $b(n, m)$ is decreasing in n and constant in m , we can follow the procedure illustrated in Sections 4 and 5. It is straightforward to see that the optimal structure could be either a hierarchy with multiple “heads” independent of each other if the optimal number of linked agent is below some threshold, and a mixed structure otherwise. Note that this follows closely the intuition of Lemma 2 and Proposition 2. However, in this case the mixed structure includes (or in some cases it is entirely constituted by) *cells with several agents* linked to each other. An interesting aspect of this analysis would be to characterize the optimal cell size as a function of the detection strategy.

On the other hand, if b is increasing in m , the model would display an *additional tendency towards centralization* caused by the returns from coordination. However, the trade-offs and effects we analyze above will still hold. The resulting optimal structures will tend to be more hierarchical than the ones we describe in the paper. In this sense, one of the contributions of our results is to show that, *even in the absence of returns from coordination*, there are cases in which hierarchies are the optimal information structures.

7.2.2 Allocation of Organizational Resources for Protection

It is reasonable to suppose that the organization may be able to devote resources to protecting particular agents from detection. For example this might be interpreted literally and reflected in the use of bodyguards and other physical protection, or one could consider

hiring expensive lawyers as a mean of protection. Another means of exerting effort to protect agents from detection is by altering behaviour of agents (for example, there are ways of committing the same crimes more or less covertly and such different means vary in their costs). This last consideration is quite closely related in spirit to the discussion on cooperation-based detection and indeed the discussion in Section 5 well informs our brief discussion here.

Rather than considering the optimal level of protection given its costs or other general considerations, our interest here, as in the rest of the paper, is the effect on information structure. In particular, it is clear that even if agents start out with identical probabilities of detection, the organization may benefit from spending protection resources asymmetrically and in particular it may be beneficial to move towards a more centralized structure, that is either an individual- or pair-dominated hierarchy. Further notice, that if some agents start with a natural advantage (that is they are relatively unlikely to be detected) then the organization may disproportionately spend protection resources on precisely these agents when inducing a centralized structure.

References

- [1] Abreu D. (1988), “On the Theory of Infinitely Repeated Games with Discounting,” *Econometrica*, 56(2), March 1988, 383-396.
- [2] Anderson J. E. and O. Bandiera (2006), “Traders, cops and robbers,” *Journal of International Economics*, Volume 70, Issue 1, September 2006, Pages 197-215.
- [3] Anselmi T. (2003), “La testimonianza di Tina Anselmi sulla Resistenza raccolta da Alessandra Chiappano,” <http://www.novecento.org>.
- [4] Athey S. and K. Bagwell (2001), “Optimal Collusion with Private Information,” *The RAND Journal of Economics*, Vol. 32 (3), pp 428-465.
- [5] Aubert C., Kovacic W. and Rey P. (2006), “The impact of leniency and whistle-blowing programs on cartels,” *International Journal of Industrial Organization*, Volume 24, Issue 6, November 2006, Pages 1241-1266 .
- [6] Bandiera O. (2003), “Land Reform, the Market for Protection and the Origin of Sicilian Mafia,” *Journal of Law, Economics and Organizations*, Vol.19.

- [7] Ballester, C., Calvó-Armengol A., and Zenou, Y. (2006) “Who’s Who in Networks. Wanted: The Key Player,” *Econometrica*, Vol.74, pp.1403-1418.
- [8] Ben Porath, E. and M. Kahneman (1996), “Communication in Repeated Games with Private Monitoring,” *Journal of Economic Theory*, Vol.70, 281-297.
- [9] Benmelech, E. and C. Berrebi (2006). “Attack Assignments in Terror Organizations and the Productivity of Suicide Bombers,” working paper.
- [10] Berman E. (2003), “ Hamas, Taliban and the Jewish Underground: An Economist’s View of Radical Religious Militias,” Working Paper.
- [11] Berman E. and D. D. Laitin (2005), “Hard Targets: Theory and Evidence on Suicide Attacks,” Working Paper.
- [12] Bueno de Mesquita, E. and C. Hafer (2007), “Public Protection or Private Extortion?,” *Economics and Politics*, forthcoming
- [13] Charlie M. (2002) “Into the Abyss: A Personal Journey Into the World of Street Gangs”, Manuscript.
- [14] Dixit, A. (2004), *Lawlessness and Economics: Alternative Modes of Governance*, Princeton University Press, Princeton NJ
- [15] Falcone G. (1991), *Cose di Cosa Nostra*, Rizzoli, Milan.
- [16] Falcone G. and G.Turone (1982), “Tecniche di Indagine in Materia di Mafia,” Report to the consiglio Superiore della Magistratura.
- [17] Farley, J. D. (2003), “Breaking Al Qaeda Cells: A Mathematical Analysis of Counterterrorism Operations (A Guide for Risk Assessment and Decision Making),” *Studies in Conflict and Terrorism*, Vol. 26, 399-411.
- [18] Farley, J. D. (2006), “Building the perfect terrorist cell,” conference presentation.
- [19] Gambetta D. (1993), *The Sicilian Mafia*, Harvard University Press, Cambridge, Mass.
- [20] Garoupa, N. (2007), “Optimal law enforcement and criminal organization,” *Journal of Economic Behavior & Organization*, 63(3), pp. 461-474.

- [21] Garreau, J. (2001), "Disconnect the Dots," *Washington Post*, September 16, 2001.
- [22] Glaeser E. L., B. Sacerdote and J.A. Scheinkman (1996), "Crime and Social Interactions," *Quarterly Journal of Economics*, 111(2), pp.507-548.
- [23] Glaeser E. L., D.I. Laibson, J.A. Scheinkman and C. L. Soutter (2000), "Measuring Trust", *Quarterly Journal of Economics*, 115(3) pp.811-846.
- [24] Harrington J. (2003), "Some Implications of Antitrust Laws for Cartel Pricing," *Economics Letters*, Vol. 79, June 2003, 377-383.
- [25] Harrington J. (2004), "Cartel Pricing Dynamics in the Presence of an Antitrust Authority," *RAND Journal of Economics*, Vol. 35, Winter 2004, 651-673.
- [26] Harrington J. (2005), "Optimal Cartel Pricing in the Presence of an Antitrust Authority," *International Economic Review*, Vol. 46, February 2005, 145-169.
- [27] Harrington, J. (forthcoming): "Optimal Corporate Leniency Programs," *Journal of Industrial Economics*, *forthcoming*.
- [28] Ichino, A. and G. Muehlheusser (forthcoming): "How often should you open the door? Optimal monitoring to screen heterogeneous agents" *Journal of Economic Behavior and Organization*
- [29] Maskin E., Y. Qian, C. Xu (2000), "Incentives, Information, and Organizational Form," *The Review of Economic Studies*, Vol. 67, No. 2, pp. 359-378.
- [30] Pearl M., (2003), "A Mighty Heart: The Inside Story of the Al Qaeda Kidnapping of Danny Pearl", *Scribner Publishing*.
- [31] Piccione M. and A. Rubinstein (2004), "The Curse of Wealth and Power," *Journal of Economic Theory*, Vol. 117, 119-123.
- [32] Radner R. (1992), "Hierarchy: The Economics of Managing," *Journal of Economic Literature*, Vol. 30, No. 3, pp. 1382-1415
- [33] Radner R. (1993), "The Organization of Decentralized Information Processing," *Econometrica*, Vol. 61, No. 5, pp. 1109-1146

- [34] Rajan R. and L. Zingales (2002), “Power in a Theory of the Firm,” *Quarterly Journal of Economics*, Vol. 113.
- [35] Ricart-I-Costa J. (1988), “Managerial Task Assignment and Promotions,” *Econometrica*, Vol. 56 (2), 449-466.
- [36] Schulhofer S. J. (2005), *Rethinking the Patriot Act: Keeping America Safe and Free*, Century Foundation Press.
- [37] Smith A. and F. Varese (2001,) “Payment, Protection and Punishment,” *Rationality and Society*, Vol. 13(3), pp. 349-393.
- [38] Spagnolo, G. (2003), “Optimal Leniency Programs,” *Mimeo*
- [39] Spagnolo, G. (2007): “Leniency and Whistleblowers in Antitrust,” forthcoming as Ch. 12 of P. Buccirosi (Ed.), *Handbook of Antitrust Economics*, 2007, M.I.T. Press.
- [40] Stille A. (1995), *Excellent Cadavers*, Pantheon Books, New York, NY.
- [41] Thompson, T. (2005), *Gangs: A Journey into the Heart of the British Underworld*, Hodder & Stoughton.
- [42] Tucker, C. (2007): “Formal and Informal Influence in the Adoption of Technologies with Network Effects” working paper
- [43] Van Zandt T. (1998), “Organizations that Process Information with an Endogenous Number of Agents,” in *Organizations with Incomplete Information*, edited by Mukul Majumdar, Cambridge: Cambridge University Press. Chapter 7, pages 239-305.
- [44] Van Zandt T. (1999), “Decentralized Information Processing in the Theory of Organizations,” *Contemporary Economic Issues*, Vol. 4: *Economic Design and Behavior*, edited by Murat Sertel, London: MacMillan Press Ltd. Chapter 7, pages 125-160.
- [45] Waldman M. (1984), “Job Assignments, Signalling and Efficiency,” *The RAND Journal of Economics*, Vol.15 (2), 255-267.
- [46] Wasserman and Faust (1994), *Social Network Analysis: Methods and Applications*, Cambridge University Press.

- [47] Williamson O. E. (1983), "Credible Commitments: Using Hostages to Support Exchange," *American Economic Review*, vol.73(4).
- [48] Zbojnik J. (2002), "A Theory of Trade Secrets in Firms," *International Economic Review*, vol.43.

Appendix

Proof of Lemma 1 (i) Consider any information structure μ in which there is at least an agent i linked to at least two agents, j and h . Note that by deleting one of these links, say $i \rightarrow h$ we obtain another information structure μ' . by the assumption in Section 2.3.2, $-k$, which is the minmax payoff for agent i that can be achieved by the other players, is not altered. Thus, agent i 's incentives to cooperate are the same for the new information structure μ' . This implies that the information structure μ' can achieve the same $\sum_i \lambda n(s(h^t)) - c1_{s(h^t)}^A(i) - k1_{s(h^t)}^B(i)$ that can be achieved by μ . However, deleting the link $i \rightarrow h$ may save in information leakage cost. Indeed, in the event in which h is detected and i is not detected (directly or indirectly through the other links that depart from i), the organization saves the cost b . This implies that the information structure μ' is weakly more efficient than the information structure μ . Since the same reasoning can be followed for any link that connects any agent to more than another agent, the claim is proved.

(ii) Take any information structure μ in which the number of agents linked to at least another agent is exactly $1 \leq n \leq N$. Take a link, say $i \rightarrow j$ and obtain a new information structure that is similar to μ except for substituting $i \rightarrow j$ with a link $i \rightarrow h$ for some $h \neq i, j$. Consider any agent's per-period payoff as described in (1). Let us consider the part of the payoff that depend on the agents' strategies $(\lambda n(s(h^t)) - c1_{s(h^t)}^A(i) - k1_{s(h^t)}^B(i))$. It is easy to see that the set of equilibrium payoffs in the repeated game obtained by μ' are the same obtained by μ . Consider now any information structure μ with $1 \leq n < N$ links. Consider an agent j which is not linked to anybody else in μ . (such agent exists since $n < N$). Consider the information structure μ' identical to μ except for the fact that a link $i \rightarrow h$ is substituted by $j \rightarrow l$, for some $l \neq j$. Again, if we focus on the part of the payoff that depend on the agents' strategies $(\lambda n(s(h^t)) - c1_{s(h^t)}^A(i) - k1_{s(h^t)}^B(i))$, it is easy to see that the set of equilibrium payoffs in the repeated game obtained by μ' are the same obtained by μ . It follows then that the most efficient information structure is one that minimizes

$$\sum_i b \left[1 - \prod_{j \in V_i} (1 - \alpha_j) \right]. \blacksquare$$

Proof of Proposition 1

Note that in this environment, following Abreu (1988), the most efficient equilibria can be replicated by equilibria sustained by the most severe equilibrium punishment, which in anarchy entails no cooperation by any of the agents.

Let us now assume that $\delta > \max_i \left\{ \frac{c - \lambda + (1 - \gamma)\beta_i b}{\lambda(N-1)} \right\}_{i=1}^N \equiv \bar{\Delta}$, and consider an anarchy and the candidate equilibrium in which everyone always cooperates except following any deviation (by anybody) from full cooperation. Then if an agent i deviates from the equilibrium strategy he gets $\lambda(N-1) - \frac{\gamma\beta_i b}{1-\delta}$ as he gains $\lambda(N-1)$ in the current period but earns nothing in all future periods, whereas cooperation yields $\frac{\lambda N - c}{1-\delta} - \frac{\beta_i b}{1-\delta}$. Therefore, this equilibrium is sustainable if and only if $\frac{\lambda N - c}{1-\delta} - \frac{\beta_i b}{1-\delta} > \lambda(N-1) - \frac{\gamma\beta_i b}{1-\delta}$, or, equivalently, if and only if $\delta > \frac{c - \lambda + (1 - \gamma)\beta_i b}{\lambda(N-1)}$, which is satisfied by assumption. This implies that, if $\delta > \bar{\Delta}$, anarchy is the optimal information structure, and all agents are as in case (i).

Let us now assume that $\delta < \min_i \left\{ \frac{c - \lambda + (1 - \gamma)\beta_i b}{\lambda(N-1) + k} \right\}_{i=1}^N \equiv \underline{\Delta}$. In this case, it is easy to see that cooperation cannot be achieved even by the threat of the additional punishment. This implies that exchanging information does not have any benefits and, if the probability of detection of the

information receiver is positive, it has the cost of increasing the probability of detection. Thus, if $\delta < \underline{\Delta}$, an anarchy achieves again the highest efficiency, and all agents are as in case (i).

The remainder of the proof appears in the text. ■

Proof of Proposition 2

First step. Recall that $\rho(j, i)$ is decreasing in both α_j and α_i . This follows trivially from the fact that if $x \geq y \geq 0$ then $\frac{xz}{x+z} \leq \frac{yz}{y+z}$ for all $z \geq 0$.

Second step Let us prove that among all possible binary cell information structures that pair N agents to each other $\{\mu \in I \text{ s.t. if } \mu_{ij} = 1 \text{ for some } i \neq j \text{ then } \mu_{ji} = 1 \text{ and } \mu_{ik} = 0 \forall k \neq j\}$ the one which minimizes information leakage costs is $1 \longleftrightarrow 2, 3 \longleftrightarrow 4, \dots, N-1 \longleftrightarrow N$. To see this, let us first show that this result holds for $N = 4$. The claim is true if $1 \longleftrightarrow 2, 3 \longleftrightarrow 4$ is better than either of the alternatives $1 \longleftrightarrow 4, 2 \longleftrightarrow 3$ and $1 \longleftrightarrow 3, 2 \longleftrightarrow 4$. This requires that:

$$\begin{aligned} 2b[1 - (1 - \alpha_1)(1 - \alpha_2)] + 2b[1 - (1 - \alpha_3)(1 - \alpha_4)] &\leq \\ 2b[1 - (1 - \alpha_1)(1 - \alpha_4)] + 2b[1 - (1 - \alpha_3)(1 - \alpha_2)] &\end{aligned} \quad (5)$$

and,

$$\begin{aligned} 2b[1 - (1 - \alpha_1)(1 - \alpha_2)] + 2b[1 - (1 - \alpha_3)(1 - \alpha_4)] &\leq \\ 2b[1 - (1 - \alpha_1)(1 - \alpha_3)] + 2b[1 - (1 - \alpha_2)(1 - \alpha_4)] &\end{aligned} \quad (6)$$

Inequality (5) holds if $\alpha_1\alpha_2 + \alpha_3\alpha_4 \geq \alpha_1\alpha_4 + \alpha_2\alpha_3$ or if $(\alpha_4 - \alpha_2)(\alpha_3 - \alpha_1) \geq 0$, which is always the case. Inequality (6) also always holds.

Now, suppose that for a general even N the claim is not true. Then, there is an optimal structure in which it is possible to find 2 pairs $\{i_1, i_2\}, \{i_3, i_4\}$ such that $\alpha_{i_1} \leq \alpha_{i_2} \leq \alpha_{i_3} \leq \alpha_{i_4}$ is violated. Then, since that is the optimal structure, rearranging the agents in these pairs leaving all other pairs unchanged cannot reduce information leakage costs. However, this contradicts the result for $N = 4$.

Third step. It is clear that the best way to link agents 1 and 2 is to link them to each other since they are the two lowest-probability agents. Now, for any couple $\{N-1, N\}, \dots, \{3, 4\}$ let us compare whether it is better from an information leakage point of view to link the pair to each other and independently from the others, or to have them linked to agent 1 (and 2) instead. If the agents N and $N-1$ are linked to each other, the cost of information leakage corresponding to the couple is $2b[1 - (1 - \alpha_N)(1 - \alpha_{N-1})]$. If they are linked to agents 1 and 2, the cost of information leakage is $b[1 - (1 - \alpha_1)(1 - \alpha_2)(1 - \alpha_N)] + b[1 - (1 - \alpha_1)(1 - \alpha_2)(1 - \alpha_{N-1})]$. Then, the couple $\{N-1, N\}$ should be linked to agent 1 (and then, since we have $1 \leftrightarrow 2$, to the couple $\{1, 2\}$) if and only if

$$\rho(N-1, N) < (1 - \alpha_1)(1 - \alpha_2) \quad (7)$$

If condition (7) fails, by the first step of this proof we know that the condition will fail for any subsequent couple. Then, the optimal way to link the N agents to each other is to create a pairwise structure, and by the second step of this proof we know that the optimal way to do this is to set $1 \leftrightarrow 2, 3 \longleftrightarrow 4, \dots$ and $N \longleftrightarrow N-1$. If condition (7) is satisfied, we can link agents N and $N-1$ to the couple $\{1, 2\}$, and we can repeat this check for the couple $\{N-2, N-3\}$. We repeat this process until we find a couple $\{i-1, i\}$ for which the condition

$$\rho(i-1, i) < (1 - \alpha_1)(1 - \alpha_2)$$

fails. If we find such a couple, by the first step of this proof we know that the condition will fail for any subsequent couple, and, by the second step of the proof, we can arrange any subsequent couple in a pairwise fashion. ■

Proof of Proposition 4

Note that, if the authority allocates the same budget β to each agent, the cost of each links becomes $b\beta(1 - \beta)$. Since this cost is maximized at $\beta = \frac{1}{2}$, it is never optimal to set $\beta > \frac{1}{2}$ in a symmetric allocation.

Next, in order to prove this result, let us assume that $\frac{B}{N-1} \leq 1$ and let us prove the following Lemma first. Let $\tilde{\beta} \equiv \left\{0, \frac{B}{N-1}, \dots, \frac{B}{N-1}\right\}$.

Lemma The allocation $\tilde{\beta}$ increases the information leakage cost of the Nth link (linking agent 1 to agent 2) compared to any other allocation β which generates exactly N-1 links.

Proof: Consider any allocation β that generates exactly $N - 1$ links. Since $\beta_1 \leq \beta_2 \leq \dots \leq \beta_N$ and $\beta_1 \geq 0$, it follows that $\beta_2 \leq \frac{B}{N-1}$. We can compare the additional information leakage costs from the N -th link, $c(N) = C(N) - C(N-1)$ and $\tilde{c}(N) = \tilde{C}(N-1) - \tilde{C}(N-1)$ associated with each agent i under allocations β and $\tilde{\beta}$. In order to do that, let us consider the allocation $\hat{\beta} \equiv \{0, \beta_2, \dots, \beta_N\}$ and first compare β with $\hat{\beta}$. Under the optimal information structures with N links described in Proposition 2, given allocation β , either (a) agent i remains linked to agent 1 or (b) agent i is in a binary cell with some other agent j in the organization (which will be $i + 1$ or $i - 1$ depending on whether i is even or odd). In case (a) the incremental leakage cost for agent i is $b(1 - \beta_i)(1 - \beta_1)\beta_2$, while under allocation $\hat{\beta}$ is going to be $b(1 - \beta_i)\beta_2$. Trivially, $b(1 - \beta_i)(1 - \beta_1)\beta_2 < b(1 - \beta_i)\beta_2$. In case (b), since the incremental information leakage cost for agents i and $i + 1$ of the $N - th$ link under allocation β is $b(1 - \beta_i)\beta_{i+1} + b(1 - \beta_{i+1})\beta_i - b(1 - \beta_1)\beta_i - b(1 - \beta_1)\beta_{i+1}$ where the first positive terms denotes the new information leakage costs associated with these agents and the negative terms the old information leakage costs when they were subordinates in the $N - 1$ hierarchy. Since the cell is preferred to making i and $i + 1$ subordinates to agents 1 and 2, it follows that

$$\begin{aligned} & b(1 - \beta_i)\beta_{i+1} + b(1 - \beta_{i+1})\beta_i - b(1 - \beta_1)\beta_i - b(1 - \beta_1)\beta_{i+1} \\ < & b(\beta_1 + \beta_2 + \beta_1\beta_2)\beta_{i+1} + b(\beta_1 + \beta_2 + \beta_1\beta_2)\beta_i - b(1 - \beta_1)\beta_i - b(1 - \beta_1)\beta_{i+1} \\ = & b(1 - \beta_i)(1 - \beta_1)\beta_2 + b(1 - \beta_{i+1})(1 - \beta_1)\beta_2 \\ < & b(1 - \beta_i)\beta_2 + b(1 - \beta_{i+1})\beta_2 \end{aligned}$$

The last expression is the information leakage cost associated with the allocation $\hat{\beta}$ (that is the information leakage costs beyond those incurred in anarchy).

Next, we show that the allocation $\tilde{\beta}$ has a higher information leakage cost for the $N - th$ link $\tilde{c}(N)$ than the allocation $\hat{\alpha}$, that is $\tilde{c}(N) \geq \hat{c}(N)$. These two costs can be written down trivially:

$$\tilde{c}(N) = b \sum_{i=3}^N \frac{B}{N-1} \left(1 - \frac{B}{N-1}\right) = b(N-2) \frac{B}{N-1} \left(1 - \frac{B}{N-1}\right)$$

and

$$\hat{c}(N) = b \sum_{i=3}^N \alpha_2(1 - \alpha_i) = b(N-2)\alpha_2 - b\alpha_2 \sum_{i=3}^N \alpha_i$$

Since $\sum_{i=3}^N \beta_i < B < N - 2$, it follows that information leakage costs under $\widehat{\beta}$ are increasing in β_2 , whose highest value is $\frac{B}{N-1}$ and when it takes this value the information leakage costs are equal to those under $\widetilde{\beta}$. Thus $\widetilde{c}(N) \geq \widehat{c}(N) \geq c(N)$.³⁶

This concludes the proof of Lemma 7.2.2 ■

Let us now proceed to the proof of Proposition 4

First step. First of all, note that, under some circumstances, the external authority's strategy will be irrelevant. For example, if the benefits are much larger than the punishment from the authority.

Second step. Suppose now that $N\lambda - c < b\frac{B}{N}(1 - \frac{B}{N})$. By Lemma 1, in this case, the symmetric allocation deters the organization from establishing any link, so this will be the optimal strategy for the external agent. *In the rest of the proof we will then assume that $N\lambda - c > b\frac{B}{N}(1 - \frac{B}{N})$.*

Third step. Assume $N\lambda - c > b\frac{B}{N}(1 - \frac{B}{N})$. In points (1)-(3), we go over all the possible budget allocation and show that the allocation $\widetilde{\alpha} = \left\{0, \frac{B}{N-1}, \dots, \frac{B}{N-1}\right\}$ is optimal.³⁷

(1) Consider any allocation such that $\beta_1 = \beta_2 = 0$. Then, the organization can reach full efficiency with zero additional information leakage cost with respect to anarchy. To see this, suppose that $\beta_1 = \beta_2 = 0$; then, an organization with the links $\mu_{1i} = 1$ for all $i \in \{2, \dots, N\}$, $\mu_{21} = 1$ and $\mu_{ij} = 0$ otherwise delivers full efficiency for any $\delta > \Delta(N - 1, k)$. Thus, it must be the case that, in order to prevent links between agent and deter efficiency, *at most one agent can be left with zero probability of detection.*

(2) Consider any allocation such that $\beta_1 > 0$ —that is, *all* the probabilities of detections are set to be positive. Since we are under the assumption that $N\lambda - c > b\frac{B}{N}(1 - \frac{B}{N})$, if these probabilities are symmetric, full cooperation will ensue, and the allocation $\widetilde{\beta}$ cannot do worse than that. Suppose, then, that the allocation is asymmetric that is $\beta_1 < \frac{B}{N}$. Following the characterization in Proposition 2 and Lemma 2, the agents will then form an optimal organization.

First, suppose the parameters are such that the organization has N links. Then, the allocation we are considering reaches full efficiency, and the allocation $\widetilde{\beta}$ cannot do worse than that.

Suppose, instead, that the optimal organization given the allocation β we are considering generates $N - 1$ links. Then by the Lemma 7.2.2, allocation $\widetilde{\beta}$ performs at least as well.

Finally, suppose that under the allocation α the linked agents are $n < N - 1$. We argue that such a structure is impossible. In such organizations, according to Lemma 2, there are three types of agents to consider: the top of the hierarchy agent 1, the $N - n - 1$ independent agents $2, \dots, N - n$, and the n agents who reveal their information to agent 1—that is $N - n + 1, \dots, N$. Without loss of generality, we will restrict our attention to the allocations that give the same probability of detection to each agent in the same category (if the probability is not the same, it is easy to see that it is possible to substitute such probabilities with the average in each category and still obtain the same structure of organization). Let's name such probabilities β_1, β_2 and β_N respectively. The probability allocations we are restricting our attention to have to satisfy the following constraints:

- (i) $0 < \beta_1 \leq \beta_2 \leq \beta_N \leq 1$ (by feasibility and by Lemma 2);

³⁶Trivially, if the mixed strategy is realized before formation, then it can never be strictly optimal for the external agent to choose a mixed strategy.

³⁷We can interpret an allocation $\alpha = \{\alpha_1, \dots, \alpha_N\}$ as a mixed strategy that generates an average probability of detection for agent i of α_i as in the proof of Lemma 7.2.2 and as in that proof. The arguments below effectively assume independent probabilities of detection for agents. A mixed strategy by the external authority introduces negative correlation which can only make the introduction of a link more attractive.

(ii) $b\beta_1(1-\beta_2) \geq N\lambda - c$ (it is not optimal for the organization to link the $N - n - 1$ independent to agent 1);

(iii) $N\lambda - c \geq b\beta_1(1 - \beta_N)$ (it is optimal for the organization to link the n agents to agent 1);

(iv) $\alpha_1 + (N - n - 1)\beta_2 + n\beta_N \leq B$ (the resource constraint).

Note that $b\beta_1(1-\beta_2) \leq b\beta_2(1-\beta_2) \leq b\frac{B}{N}(1-\frac{B}{N})$ since $\beta_2 \leq \frac{B}{N} < \frac{1}{2}$ (otherwise either the (iv) or is violated or it cannot be that $\beta_1 \leq \beta_2 \leq \beta_N$) but then (ii) cannot hold since $N\lambda - c > b\frac{B}{N}(1-\frac{B}{N})$. It follows that such a structure is impossible.

(3) In points (1)-(2) we showed that if $N\lambda - c > b\frac{B}{N}(1 - \frac{B}{N})$, all the allocations such that $\beta_1 = \beta_2 = 0$ or $\beta_1 > 0$ are (weakly) dominated by allocation $\tilde{\beta}$. Finally, let us consider an allocation such that $\beta_1 = 0$ and $\beta_2 > 0$. Under this allocation, it is clear that an organization with $N - 1$ linked agents can arise costlessly. Thus, the best the external agent can do is to try to prevent the $N - th$ link from arising. Observe that, if $\beta_1 = 0$, the characterization in Proposition 2 yields, for each $i \in \{4, \dots, N\}$, to $\frac{2(1-\beta_{i-1})(1-\beta_i)}{2-\beta_{i-1}-\beta_i} \leq 1 - \beta_2$ (easy to check since $\beta_2 \leq \beta_j$ for all $j \in \{3, \dots, N\}$). Then, in the optimal organization, all the agents are linked to agent 1, without binary cells (besides the cell $\{1, 2\}$). Then, the cost of the $N - th$ link for the organization is $b\beta_2 \sum_{i=3}^N (1 - \beta_i)$, and it is maximized (under the constraints $\beta_2 \leq \beta_i$ for all i and $\sum_{i=2}^N \beta_i = B$) by $\beta_i = \frac{B}{N-1}$ for all $i \in \{2, \dots, N\}$, which is allocation $\tilde{\beta}$. ■

Proof of Lemma 3 First consider linking a single agent i to some other agent. Such link yields to the organization the benefit $N\lambda - c$. Regardless of which agent j he is linked to agent j is not cooperating, and so the cost of linking agent i is simply $b\beta_i$. This is minimized when β_i is as small as possible. Thus, the cheapest link connects agent 1 to some other agent. A similar argument applies for all $n < N$. ■

Proof of Corollary 2 Note that in this case $\tilde{C}(n) = \begin{cases} 0 & n = 0 \\ bn\beta & n = 1, \dots, N - 1 \\ bN\beta(2 - \beta) & n = N \end{cases}$. It is

easy to see that if it is worth linking one agent (that is, if $N\lambda - c > b\beta$) then it is worth linking at least $N - 1$ agents. However the cost of the N^{th} link $\tilde{C}(N) - \tilde{C}(N - 1) = b\beta + Nb\beta(1 - \beta)$ is higher than the cost of all the previous links. Thus, if $b\beta < N\lambda - c < b\beta + Nb\beta(1 - \beta)$, a structure with $N - 1$ links is optimal even though all agents are symmetric. Clearly, if $N\lambda - c$ is sufficiently high then a binary cell structure is optimal and if it is sufficiently low, then an anarchy is optimal. ■

Proof of Lemma 4 For the organization perspective, the per-period benefit of a link is at most $N\lambda - c$. The information leakage cost for linking agent i is bounded below by $\beta_i b$ (this would be exactly equal to the cost of the link if the agent were linked to an agent who is not cooperating). Thus, an agent will not cooperate if $\beta_i b$ is higher than $N\lambda - c$. Thus, by allocating a budget of $\frac{N\lambda - c}{b}$ to scrutinizing an agent the external authority can prevent this agent from cooperating. ■

Proof of Proposition 6 First suppose $B \geq 2\frac{N\lambda - c}{b}$. In this case, then it is possible to allocate $\beta_i, \beta_j \geq \frac{N\lambda - c}{b}$ to at least two agents i and j and preventing them from cooperating (as Lemma 4 guarantees). Given that agent i is not cooperating, the threshold value of β_j at which j stops cooperating is exactly $\frac{N\lambda - c}{b}$ (this is now not just the lower bound for the threshold but the exact value for it). Similarly, since j is not cooperating, $\frac{N\lambda - c}{b}$ is the exact threshold for i . Therefore, with some agent not cooperating, if there is some agent k with $\beta_k < \frac{N\lambda - c}{b}$, then agent k can be linked to a non-cooperating agent and be induced to cooperate in equilibrium. It follows that allocating at least $\frac{N\lambda - c}{b}$ to as many agents as possible is the optimal strategy.

Suppose now that $B < 2\frac{N\lambda - c}{b}$. We consider two cases: (i) $B < \frac{N\lambda - c}{b}$. In this case, the external

authority's optimal strategy cannot deter any agents. Thus, the external authority's strategy is irrelevant since all budget allocations are optimal. (ii) $\frac{N\lambda-c}{b} \leq B \leq 2\frac{N\lambda-c}{b}$. In this case, the external authority can deter directly exactly one agent, say agent i . The other agents cannot be deterred directly from cooperating. However, consider the cooperation of the other agents. The organization will not prefer agent j to be linked (and cooperating) as long as the benefit from the cooperation $N\lambda - c$ is lower than the lowest possible information leakage cost $\min_h \beta_j b + (1 - \beta_j)\beta_h b$. It follows that allocating all resources to scrutinizing one agent, and not scrutinizing the other agents at all makes the condition that the organization prefers that all agents cooperate most difficult to satisfy. Note that

$$N\lambda - c \geq \max_i \min_j \beta_i b + (1 - \beta_i)\beta_j b \quad (8)$$

is the appropriate condition. ■