# How to Organize Crime*

Mariagiovanna Baccara and Heski Bar-Isaac†

October 2006

**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 facilitating cooperation). 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. Finally, we discuss a number of applications and extensions.

## 1 Introduction

In this paper, we study the interplay between cooperation 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 important factor in building internal trust and cohesion, but it can leave the organization

*Previous drafts have circulated under the titles "Power and organizational structure under an external threat" and "Crime, Punishment and Organizational Structure." We thank Pierpaolo Battigalli, Jean-Pierre Benoît, 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 INSEAD, Pennsylvania State, NYU, Tel Aviv University, USC, Bocconi University, University of Pennsylvania, University of Venice, IDEI Toulouse, Paris-Jourdan Sciences Economiques, the 2005 IIOC, the 2005 EEA, the 2006 ESSET, the Third Conference on Mathematical Methods in Counterterrorism and the Second Conference on "Conflict and Cooperation" at Northwestern University for useful conversations and comments. All errors are ours.

†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.

vulnerable to an external threat. Understanding how this trade-off affects the information structure of an organization and its productivity, allows us to assess detection policies designed to destabilize such organizations.

Since 9/11, a $400 billion annual budget has been passed 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 in the last fifty years. 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 considerable insight into answering this question.

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, such as the anarchist and revolutionary organizations in the late 19th century in Europe and the communist organization in the early 20th century.[2]

[1] Another 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).

[2] Among 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

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 self-enforcing organizations. In particular, 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. Note that in practice this information structure can coexist and may interact with organization structures (such as communication and decision-making structures). We abstract from these interactions to highlight the information hierarchy in the organization.

We consider an organization of $N$ agents and characterize its optimal information structure. When we talk of information structure, we have in mind characterizing which members of the organization know the real name (rather than nickname) of some other member of the organization, or who within the organization holds some incriminating evidence about him or other detailed information that would harm him if it came to light. In the example in Section 2 we simply assume a reduced form for the benefits of agents in the organization holding such information, while in the rest of the paper we provide a model which rationalizes the benefit of information links—essentially arguing that it leads to greater trust which can be 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 changes in the external agents policies and resources.

We consider two alternative models of detection available to the external authority. In the first model (*agent-based detection*), the authority allocates a budget to detect each agent independently of his cooperation in the organization. For example, regardless of the activities he is currently engaged in, authorities are actively seeking Osama Bin Laden and his lieutenants. In the second model (*cooperation-based detection*), each agent's probability of detection is a function of the level of cooperation within the organization. For instance, if the organization members are drug dealers, a possible policy for the authority is to look

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...".

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

We characterize the optimal information structure within the organization in the two models and compare them. In the agent-based detection model, we find that if the probabilities of detection are sufficiently similar, the optimal structure consists of either an anarchy (where no agent is linked to any other) or an organization constituted by "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 fact, by doing so, the agents will be induced to form a hierarchy *strictly less efficient than a binary cell structure.*

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. Despite the fact that in this model all agents are ex-ante symmetric, we show that the optimal information structure can be a hierarchy in which the information hub does not cooperate at all, and thus remain undetected. If each individual agent's contribution to the organization is sufficiently high, the optimal organization can also be a binary cell structure.

We compare the two detection strategies and, in particular, highlight two considerations. First, there are situations in which an agent-based detection model is the only feasible strategy for law enforcement. For instance, the authority is forced to use an agent-based detection model when the day-to-day activity of the organization consists of tasks that are not illegal or outside the norm, such as meetings, phone conversations, or flying lessons. In these circumstances, the authority's detection strategy must be individually targeted to specific agents, and a crucial decision for the authority is how spread out the detection should be. Then, we can verify the positive results of the agent-based detection model using the available evidence. Our results predict that, when detection is agent-based, a symmetric detection strategy leads to a binary cell structure, which is similar to the way terrorists seem to be organized. Moreover, our results on the optimal budget allocation take a step towards the normative side of this application.

There are, however, situations in which an authority can choose between the two detection models.[3] In fact, many criminal organization carry out daily illegal activity, as has been historically the case with Mafias, organizations involved traditionally in gambling and liquor trading and, more recently, in drug and gun dealing. In these cases, our results provide a direct comparison between the two alternative detection models. Indeed, our results suggest than when the authority chooses a symmetric agent-based detection, the equilibrium information structure is either an anarchy or a binary cell structure, but never a hierarchy. If the authority chooses a cooperation-based detection strategy, the equilibrium information structure is either a hierarchy or a binary cell structure. Traditionally, Mafias have been investigated mainly through cooperation-based detection models, and these results seem to match with the evidence we have on these organizations.

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.[4] Thus, our results can shed some light on the optimal information sharing protocols of these organizations.

The remainder of the paper is organized as follows. After a summary of the related literature, in Section 2, we present examples to illustrate the trade-offs and preview some of the results of the following sections. We formally introduce the model in Section 3. In Sections 4 and 5, we study the agent-based detection model. In particular, in Section 4, we take the behavior of the external agent as given and characterize the optimal information structure, and in Section 5, we endogenize the choice of the external agent. In Section 6, we study the cooperation-based detection model. In Section 7, we compare the two models and discuss the robustness of the results and, in Section 8, we present a number of extensions of the model and we conclude. All proofs are in the Appendix.

[3] A cooperation-based detection model is often easier to justify politically than the agent-based detection model. In fact, the agent-based detection model, in which agents are monitored independently on their illegal activity, is subject to reasonable (and controversial) legal and social constraints. An interesting debate related to this issue is the one surrounding the Patriot Act (see, for example, Schulhofer (2005)).

[4] For instance, consider secrecy issues in patent races and R&D departments.

### 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.[5] However, there are several strands of the literature that have elements in common with our work.

The only paper of which we are aware that considers how information structure affects behavior in a repeated game is Ben Porath and Kahneman (1996), which focuses on the structure of how agents observe the actions of other agents. In our model, in contrast, actions are observed and we characterize the information structure as discussed above.

Another strand of the literature related to this paper is on social networks. Notably, Glaeser, Sacerdote and Scheinkman (1996) focuses on the link between crime and social networks. The authors consider a fixed social network (for instance, given by the urban structure of a city) and focus their analysis on how the agents' decisions about crime are a function of their neighbors' decisions. Also, Ballester et al. (2006), under the assumptions that the network structure is exogenously given and observed, characterizes 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.

There is a wide literature on information 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 information structure that is most internally efficient. 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 behavior within the organization, but renders agents more vulnerable to an external threat.

[5] 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 hierarachy.

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 papers that tackle different notions of power include Rajan and Zingales (2002) and Piccione and Rubinstein (2004).[6]

Recent contributions to the literature on cartels deal with the impact of an external authority on the cartel's behavior. In particular, Harrington (2003(a),(b) and (c)) looks at the impact of an external authority on the cartel's optimal pricing behavior, and Spagnolo (2003) and Aubert *et al.* (2003) study the effect of leniency programs on cartels' stability. Also in the cartel literature, the papers by Athey and Bagwell (2001) and Green and Porter (1984) share with this paper the notion that communication (in their case, on costs and demand, respectively) enable a more efficient form of collusion.

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 (2002), Bandiera (2003), Bueno de Mesquita and Hafer (2005), and Dixit (2004).[7] Other attempts to use rational models to understand the behavior of terroristic groups include Berman (2003), Berman and Laitin (2005) and Benmelech and Berrebi (2006).

## 2 Illustrative examples

As a preview of our results, we provide some examples that the reader may find convenient to refer to later on in the paper.

Consider four agents, $A$, $B$, $C$ and $D$, who can form an information structure by creating directional links among themselves. A link is created when an agent reveals some information about himself to another agent. We assume that learning information about agent $i$ allows agent $j$ to foster greater cooperation from $i$. In the rest of the paper,

[6] Also, Zabojnik (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.

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

the benefit of a link endogenously arises from sustaining cooperation in a self-enforcing organization. However, in this example, for simplicity, we take a reduced-form approach and assume that if an agent reveals his information to another agent, *the benefit to the organization is fixed and equal to one*. Moreover, we assume that once $i$ is linked to some agent, there is no incremental benefit from linking $i$ to some other agent.

Linking agents has a cost for the organization as well. In particular, these costs arise from the possibility of external detection. Having an agent detected costs the organization 2. *If an agent is detected, all the ones who revealed their information to that agent are detected as well.*

We consider two different models of detection–*agent-based detection* and *cooperation-based detection*–and we find the most efficient information structure for each model

**Agent-based detection.** Suppose that agents $A'$s and $B$'s probability of detection is $\frac{1}{4}$, while $C$ and $D$ have a probability of detection $\beta \in \left[\frac{1}{4}, 1\right]$. Suppose that one wants to introduce one link in this organization. In order to do this at the minimal cost, an agent relatively likely to be detected should be linked to one unlikely to be detected. Thus, agent $D$ (or $C$) becomes linked to $A$ (or $B$); the probability with which $D$ is detected increases by $\frac{1}{4}(1-\beta)$, and the cost of this link is $2\frac{1}{4}(1-\beta)$. Similarly, if we want to introduce two links, then both $C$ and $D$ should be linked to $A$ (or $B$) at a total cost of $4\frac{1}{4}(1-\beta)$. Next, consider an organization with three links. The cheapest way to introduce a third link is to link $B$ to $A$, and the cost of this last link to the organization is $2\frac{1}{4}(1-\frac{1}{4}) = \frac{3}{8}$. Suppose, finally, that one wants to link all four agents to someone else. If, starting from the hierarchy we have generated so far, we link agent $A$ to agent $B$, as in Figure 2, the cost of this link is relatively high because if agent $B$ is detected, agents $A$, $C$ and $D$ will all be detected as well. It is easy to see that the total information leakage cost of the organization is $2\times 2\times \frac{1}{4}(1-\frac{1}{4})+2\times 2\,(1-\beta)\left(\frac{1}{4}+\frac{1}{4}-\frac{1}{16}\right)$. On the other hand, if we generate a *binary cell* structure, as in Figure 1, in which the couple $A$ and $B$ and the couple $C$ and $D$ are linked to each other, the total cost of information leakage is $2\times 2\times \frac{1}{4}(1-\frac{1}{4})+2\times 2\times \beta(1-\beta)$. The optimal information structure with four linked agents depends on the comparison between the information leakage costs of these two organizations.[8]

Then, if $\beta \leq \frac{7}{16}$ (that is, if the probabilities of detection of the agents are relatively symmetric), the optimal organization with four linked agents is the binary cell structure, as in Figure 1.

[8]It is easy to show that generating 4 links in any alternative pattern would just lead to higher costs.

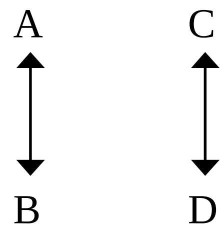


Figure 1: Binary Cells

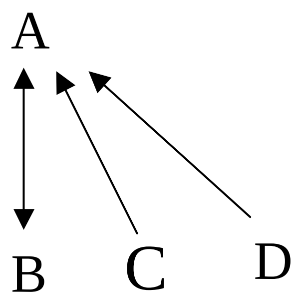


Figure 2: Centralised hierarchy with A and B as the information hub

However, if $\beta > \frac{7}{16}$ (that is, if the probabilities of detection of the agents are relatively asymmetric), the optimal organization with four linked agents is the more *centralized* one described in Figure 2. Now that we have characterized the optimal structure for any number of links, it is easy to see that the additional costs each link generates is always offset by the additional benefit (which we assumed to be equal to 1). Then, the structures described in Figures 1 and 2, indeed, represent the optimal information structures in this example.

In Section 4, in addition to endogenizing the benefits of the links, we show that the intuition of this example translates to more general detection probability distribution; and in Section 5, we address the problem of an external authority that has to set a budget to determine these probabilities of detection.

**Cooperation-based detection.** Suppose, now, that each agent's probability of getting independently detected depends on whether an agent cooperates.[9] Specifically, we assume that the probability of an agent's independent detection is $\frac{1}{4}p$ where $p \in [0,1]$ is the probability with which he cooperates. An agent cooperates only if he is linked to at least another agent, and, again, suppose that the cost to the organization of having an agent detected is 2, while the benefit of having one more agent cooperating is 1. In this circumstance, the optimal structure is a *centralized hierarchy*, as in Figure 3. In such a structure, beyond the direct costs induced by cooperating, the links are costless since A, the hub of the hierarchy, is not cooperating and, thus, never detected. Thus, the additional value of this information structure as compared to having no links is 3 (as the hierarchy ensures that three agents cooperate but has no cost in terms of incremental vulnerability to detection).

To show that this is the optimal structure, let us compare it to the binary cell structure in Figure 1. Consider a binary cell structure in which each agent cooperates with probability $p$. It is easy to see that, in this example, the efficiency of this structure is maximized for $p^* = 1$. Then, this organization would yield a net benefit of only $\frac{1}{2}$: a benefit 4 from ensuring that all four agents cooperate, but incurring incremental vulnerability costs of $\frac{7}{2}$ (for each of the four agents there is an incremental probability of detection $\frac{1}{4}\left(2 - \frac{1}{4}\right) = \frac{7}{16}$, yielding an expected cost of $\frac{7}{2}$).

Finally, consider linking agents $A$ and $B$ to each other while keeping agents $C$ and $D$ linked to $A$, as in Figure 2. In this organization, cooperation of agents $A$ and $B$ is costly from the point of view of the organization because, besides increasing the exposure of agents $A$ and $B$ to detection, it also increases the exposure of agents $C$ and $D$. Indeed, it is easy to show that the optimal cooperation level of agents $A$ and $B$ in this example is $p^* = 0$. This implies that the structure in Figure 2 is dominated by the hierarchy in Figure 3 (indeed, in the hierarchy, we have three agents cooperating at zero cost, while in this structure only agents $C$ and $D$ cooperate).

These simple examples demonstrate some prominent organizational forms–binary cells and centralized hierarchies–that emerge as optimal more generally and illustrate the trade-off between fostering cooperation and vulnerability to detection. Below, we endogenize the links' benefits, consider more general environments, and allow for a more active role for the external threat.

[9] For instance, cooperation involves an illegal activity and the authority's detection policy is based on monitoring this activity.

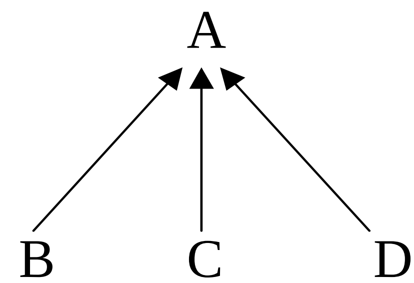


Figure 3: Centralized Hierarchy

## 3 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]

In the first of the two models of detection we study, the authority moves first and sets a given detection strategy as specified in Section 3.1.[11] Then, the $N$ players have the possibility of forming an information structure by exchanging information among themselves as specified below in Section 3.2.[12] After forming an information structure, the $N$ agents start playing an infinitely repeated stage game as specified in Section 3.3.

In Section 4, we assume that the $N$ agents take the choice of the external agent as given and we focus on the choice of the information structure. Then, we turn to the strategies available to the external agent and its optimal decision in Section 5. In Section 6 we study an alternative model of detection in which the game starts directly with the $N$ players forming an organization.

[10] Allowing $N$ to be an odd number presents no conceptual difficulties, but adds to the number of cases that need be considered with regard to how to treat the last odd agent, with no real gain in insight. Details of characterization of optimal organizational structures with an odd number of agents are available from the authors upon request.

[11] The authority moves first in the "agent-based" detection model. In the second model we will consider (the "cooperation-based" detection model) the authority's behavior is fixed and, as it will become clear later, the authority has no strategic role.

[12] We 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 22 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 a cell.

### 3.1 External Authority

At each period of the repeated stage game, each agent could be detected by the external authority. We assume that each time an agent is detected he has to pay an amount $b > 0$. This payment may represent a punishment such as a period in prison, reductions in consumptions or productivity, etc.

There are two ways for an agent to be detected, a *direct* way and an *indirect* one. First, an agent $i$ can be detected directly by the authority according some probability $\alpha_i$. We consider two alternative models through which $\alpha_i$ is determined, the *agent-based detection* model and the *cooperation-based detection* model. While in the first model, analyzed in Sections 4 and 5, $\alpha_i$ is *determined by the external authority at the beginning of the game*, in the second model, studied in Section 6, $\alpha_i$ is determined by the cooperation level of agent $i$ at each period of the game.

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 immediately detects these agents as well with probability one. Thus the external authority's ability to detect agents indirectly depends on the information structure.

#### 3.1.1 Agent-Based Detection

The external agent allocates a budget $B \in \left(0, \frac{N}{2}\right)$ to detect the other $N$ agents. In particular, it devotes $\alpha_i \in [0,1]$ to detecting member $i$ where $\sum_{i=1}^{N} \alpha_i \leq B$, and without loss of generality $\alpha_1 \leq \alpha_2 \leq ... \leq \alpha_N$. Once the budget is allocated, each agent $i$ is directly detected *independently* with probability $\alpha_i$ at every period of the subsequent repeated game.

#### 3.1.2 Cooperation-Based Detection

In the cooperation-based detection model we analyze in Section 6 we assume that, rather then being determined by the external agent, the probability of independent detection of an agent at a given period increases with his level of cooperation in the organization in the same period.

Let $p_i$ be the probability of agent $i$ cooperating at a certain period. Then, she is independently detected with probability $\alpha(p_i)$ with $\alpha : [0,1] \to [\underline{\alpha}, \overline{\alpha}]$ an increasing function.

For simplicity, we assume that $\alpha(p_i) = \alpha p_i$ with $\alpha \in (0,1)$.[13]

### 3.2 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.[14] We formalize the fact that player $j$ discloses his information to player $i$ by an indicator variable $\mu_{ij}$, 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 $j \to i$ to represent $\mu_{ij} = 1$ (and, similarly, for instance $i, j, k \to l$ to represent $\mu_{li} = \mu_{lj} = \mu_{lk} = 1$).[15] 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 $\mu$ 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 $\{h_1, .., h_n\}$ such that $i \to h_1$, $h_1 \to h_2$, .., $h_n \to j$. Thus, given an information structure $\mu$, 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 this set as $V_i$, $i$'s *vulnerability set*, and to $V \equiv \{V_1, ..., V_N\}$ as a *vulnerability structure*. Note that the vulnerability structure is induced by the choice of the information structure.

**Definition 1** *An "anarchy" is an information structure with no links, that is such that* $\mu_{ij} = 0$ *for all* $i \neq j$*, or equivalently* $V_i = \{i\}$ *for all* $i$*.*

The information structure affects the agents' probabilities of detection by the external authority. Specifically, if $i$ has information about another player $j$ and if $i$ is detected (either directly or indirectly), player $j$ is detected as well.[16] Given an information structure $\mu$ and

[13] Notice that this assumption implies that there is no additional information leakage cost in sharing information with someone who is not going to cooperate. All our qualitative results still hold for a more general function $\alpha(p_i)$ with $\alpha : [0,1] \to [\underline{\alpha}, \overline{\alpha}]$ as long as $\underline{\alpha}$ is small enough.

[14] In this model, as will become apparent in Section 3.3, relinquishing this information is a means for allowing another agent to have power over herself and reflects that in many organizations "good" behavior can be encouraged through such information exchange. Thompson (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.

[15] Note that it is always the case that $\mu_{ii} = 1$.

[16] Allowing for "detection decay"–that is, supposing that if agent $i$ has information 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.

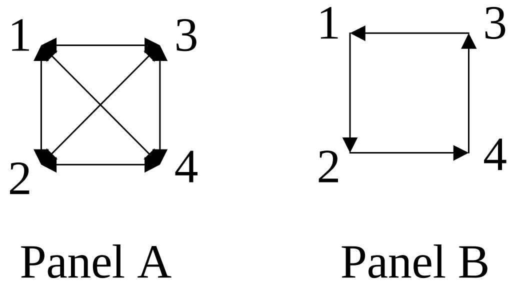


Figure 4: Examples of equivalent organizations

independent detection probabilities $\{\alpha_1, ..., \alpha_N\}$, agent $i$ is detected in one period if and only if at least one agent in $V_i$ is detected.[17]

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

Note that different information structures $\mu$ might lead to an identical vulnerability structure $V$. 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 4, 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 4. In fact, $V_i = \{1, 2, 3, 4\}$ for all $i$, and the probability of detection is $\left(1 - \prod_{j=1}^{4} (1 - \alpha_j)\right)$ for each player in both cases.

### 3.3 The Stage Game

After exchanging information about each other, the agents play an infinitely repeated stage game. 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 $j$ incur a punishment. The cooperation choice and the punishment choice are made simultaneously.

[17] Recall that $\{\alpha_1, ..., \alpha_N\}$ may or may not depend on the cooperation level of the agents, depending on whether the detection is cooperation-based or agent-based.

### 3.3.1 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 $i = 1, .., N$. Cooperation is a productive action that increases the sum of the total amount of 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 (this is because if $c > \lambda$,then $\lambda n - c < \lambda(n-1)$ for all $n$), and that $\lambda N > c$ (which implies that full cooperation is the most efficient outcome of the stage game).

### 3.3.2 Punishment technology

Suppose that player $i$ has revealed his information to player $j$ (or $\mu_{ji} = 1$). 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 ($NP$). Then, 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.

## 3.4 The Game

### 3.4.1 Timing

The timing of the game is the following:

(1) *In the agent-detection model only*, the external agent chooses the allocation $\{\alpha_1, ..\alpha_N\}$. 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 $\mu$ and implied vulnerability structure $V$, the stage game described above is played an infinite number of times at periods $t = 1, 2...$ 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, every agent is directly detected in accordance with the probabilities of independent detection (i.e., the allocation $\{\alpha_1, ..\alpha_N\}$ chosen above in the agent-based detection model and according to the cooperation level in the cooperation-based detection model) and, if detected (either directly or indirectly), he has to pay a cost $b$.

In Sections 4 and 5, we analyze the agent-based detection model. In particular, in Section 4, we analyze the subgame that starts after the external agent's decision has been made, while in Section 5, we analyze the whole game, endogenizing the external agent's choice. In Section 6, we analyze the cooperation-based detection model.

#### 3.4.2 Payoffs

**The Agents and the Organization** Let $h^t$ denote a period $t \geq 1$ history in the repeated game.[18] Let $\mathcal{H}$ denote the set of histories. Then, player $i$'s (pure) strategy is denoted as $s_i : \mathcal{H} \rightarrow A_i \times A_i'$.

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

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

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.[19]

The per-period payoff of agent $i$, $\pi_i^t(s(h^t))$, can be decomposed into $\lambda n(s(h^t)) - c1_{s(h^t)}^A(i) - k1_{s(h^t)}^B(i)$,which is the payoff coming from the interaction among the $N$ agents in the stage game, and $-b\left[1 - \prod_{j \in V_i}(1-\alpha_j)\right]$,which we refer to as the per-period "*information leakage*

[18] $h^t$ contains information on the allocation $\{\alpha_1, ..\alpha_N\}$ (in the agent-based detection model only), on the organization structure $\mu$, and on the previous decisions to cooperate or not and to punish or not by all players.

[19] Note that (1) contains a slight abuse of notation since in a cooperation-based detection model, the probability $\alpha_j$ is a function of the level of cooperation of agent $j$ at history $h^t$. Instead, in an agent-based detection model, $\alpha_j$ is determined by the external agent at the beginning of the game. Therefore, $\alpha_j$ is constant with respect to $s(h^t)$.

*cost*" for agent $i$ associated with the information structure $\mu$ (recall that $\mu$ determines the set $V_i$).

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 $\Pi(s) = \sum_{i=1}^{N} \pi_i(s)$.

Note that the *first best* for the $N$ agents in the repeated game is full cooperation and no information to be exchanged (because exchanging information causes a higher information leakage cost); however, full cooperation may not be sustainable as a Subgame Perfect Nash Equilibrium in the repeated game. Finally, notice that no cooperation and always punishing can always be sustained as an equilibrium of the repeated game for any information structure $\mu$; it represents the minmax for each player $i$ and, given any information structure $\mu$, this is clearly the equilibrium that minimizes $\Pi(s)$.

**The External Agent** The external agent appears only in the agent-detection model. 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. This is because often these payments are costly for the external authority, as they may consist of detention in prison facilities, for example.

### 3.4.3 Efficient Information Structure

For each information structure $\mu$ and for any $\delta$, 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 $\mu$ and for any $\delta$, *we identify the most efficient SPNE achievable under* $\mu$ (the SPNE that maximizes $\Pi(s)$) *when the discount factor is equal to* $\delta$*.* Let us refer to such an equilibrium as $s^*(\mu, \delta)$.

For a given $\delta$, we say that one information structure $\mu$ is *strictly more efficient* than another information structure $\mu'$ if we have $\Pi(s^*(\mu,\delta)) > \Pi(s^*(\mu',\delta))$. Then, we assume that, once the external agent chooses the allocation $\{\alpha_1, ..., \alpha_N\}$, the organization $\mu$ that will be formed is the most efficient one—that is, one that achieves the highest $\Pi(s^*(\mu, \delta))$. In other words, we assume that the $N$ agents select $\mu^* \in \arg\max_\mu \Pi(s^*(\mu, \delta))$.[20]

# 4 Agent-Based Detection: Information Structure

In this section, we start analyzing the agent-based detection model. We take the allocation of detection probabilities chosen by the external agent $\{\alpha_1, ..., \alpha_N\}$ as given, and we identify the most efficient information structure that the other $N$ agents can form. Given this characterization, in Section 5, we step back and study the external agent's optimal behavior.[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 agents' vulnerability sets. In this section, we study how this trade-off affects the organization's optimal structure, and we fully characterize the optimal information structure for any allocation $\{\alpha_1, ..., \alpha_N\}$.

First, in Section 4.1, we focus on one side of the trade-off: the information leakage costs. Then, in Section 4.2, we compare such costs with the efficiency gains that information exchange generates, and we finish by characterizing the optimal information structure.

[20] We 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 behavior 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.

Notice, also, that it is always possible to sustain this behavior 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.

[21] Note that even though we assume that the external authority determines these probabilities of detection, as we assume in the agent-based detection model, the probability of detection 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 analysis of this section can be seen as self-contained.

## 4.1 Information Leakage Costs

We 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 3.3.2), 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.

**Lemma 1** *If the number of linked agents is fixed, an efficient organization minimizes the information leakage costs.*

By Lemma 1, from the organization's point of view there is no gain in an agent disclosing his information to more than one other agent. Indeed, by doing so, the information leakage costs of the organization may increase (since the vulnerability set of the agent expands), while the maximum cooperation level obtainable from that agent remains the same.[22] This observation yields the next Lemma.

**Lemma 2** *Any efficient information structure is equivalent to another organization in which each agent reveals his information to at most one other agent.*

Lemma 2 suggests that, for any given number of linked agents, we have to understand, first, which agents should reveal their information and, second, to whom they should reveal it.

We begin by characterizing the optimal information structure when the number of linked agents $n$ is strictly less than $N$ in the next Proposition.

**Proposition 1** *The optimal structure to link $n < N$ 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., $N, N-1, ..., N-n+1 \longrightarrow 1$).*

If the number of linked agents is less than $N$, the optimal structure is simply a hierarchy, in which the top of the hierarchy (the agent who receives the information from the others) is

[22] Sometimes, additional links may come at no additional information leakage cost, as Figure 4 illustrates.

the member with lowest probability of detection and the $n < N$ "linked" agents are those with the $n$ highest probability of detection. The proof of Proposition 1 is very simple. Suppose that you have $N$ agents, and you want to generate a structure with $n < N$ agents linked to someone else. Recall that, without loss of generality, we have $\alpha_1 \leq \alpha_2 \leq ... \leq \alpha_N$. Suppose first that $n = 1$, so 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 $\alpha_i(1 - \alpha_j)$. Then, the cost of the link is minimized when $\alpha_i$ is as small as possible and $\alpha_j$ is as large as possible. If 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 Proposition 1 implies that the information leakage cost under an optimal structure in which there are $n < N$ links is simply $b\alpha_1 \sum_{i=1}^{n} (1 - \alpha_{N-i+1}) + b \sum_{i=1}^{N} \alpha_i$.

The next step is to characterize the optimal structure and the information costs under that structure when there are $N$ linked agents in the organization. Before proceeding with the characterization, we introduce the following definitions.

**Definition 2** *Two agents $\{i, j\}$ constitute a "binary cell" if they are linked to each other ($i \longleftrightarrow j$) and neither of them is linked to anybody else ($V_i = V_j = \{i, j\}$).*

**Definition 3** *Consider a cell $\{i, j\}$. Let its "independence value ratio" $\rho(i, j)$ be defined as $\rho(i, j) \equiv \frac{2(1-\alpha_i)(1-\alpha_j)}{2-\alpha_i-\alpha_j}$.*

To understand the intuition of the "independence value ratio," observe that 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* $\beta$, agent $i$ will not pay $b$ with probability $\beta(1 - \alpha_i)$, and agent $j$ will not pay $b$ with probability $\beta(1 - \alpha_j)$. 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

$$2(1 - \alpha_i)(1 - \alpha_j) > \beta(1 - \alpha_i) + \beta(1 - \alpha_j), \tag{2}$$

or, equivalently,

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

Thus, for any couple of agents, the higher is the independence value ratio the greater is the advantage of forming a cell rather than being linked to a third agent. *Notice that the independence value ratio is decreasing in both* $\alpha_i$ *and* $\alpha_j$–that is, the higher the probability of detection of an agent, the lower the independence value ratio of each cell he is part of.

We now characterize the optimal information structure with $N$ linked agents in the following Proposition.

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

Proposition 2 suggests 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. $1 \leftrightarrow 2$, $3 \longleftrightarrow 4,..,i^*-1 \longleftrightarrow i^*$). All the agents belonging to the second group reveal their information to agent 1 ($i^*+1,..,N \to 1$).[23] Figure 5, Panel A illustrates an example of the optimal structure described in Proposition 2.

Let us discuss the intuition behind Proposition 2. Suppose that, after following the procedure described in Proposition 1 and having $N-1$ agents linked to agent 1, one wants to link agent 1 to someone else. It is clear that the best way to do it is to create the link $1 \to 2$. However, because agent 1 already holds the information of all the other agents, this last link is expensive for the organization. Indeed, the rest of the organization will be now detected *if either 1 or 2 is detected*. In particular, agents with relatively low probability of detection could now have an *overall lower probability of detection by being linked to each other, rather than to both* 1 *and* 2. This suggests that, if one wants to link all the

[23] Notice that, because agents 1 and 2 form a cell, agents $i^*+1,..,N$ could equivalently reveal their information to agent 2 or to both agents 1 and 2.

$N$ agents to someone else in the organization, the best way to do it may be to create a *mixed* structure, in which agents with high probabilities of detections are linked to the cell $\{1, 2\}$ in a hierarchical fashion, while agents with relatively low probabilities of detection (agents 3, 4, etc.) remain organized in independent cells. In particular, an easy result to show is that, if we have to organize a set of agents in binary cells, the arrangement that minimizes the information leakage cost is the one in which the agents are linked to each other sequentially as $1 \leftrightarrow 2$, $3 \leftrightarrow 4, .., i^* - 1 \leftrightarrow i^*$.[24]

The number of agents $i^*$ belonging to the independent cell component depends on how steeply the independence value ratio of each couple with subsequent probabilities of detection grows. If $\alpha_1$ and $\alpha_2$ 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 $i = 4, .., N$. In this case, Proposition 2 requires that an optimizing organization links all the agents $3, .., N$ to agent 1 (who remains linked in a cell with agent 2).[25] On the other hand, if $\alpha_3$ and $\alpha_4$ are close enough to $\alpha_2$, 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 structure described in Proposition 2 is illustrated in Figure 5, Panel A, which shows the optimal structure when there are $N = 6$ links and when $(1 - \alpha_1)(1 - \alpha_2) \in [\rho(7, 8), \rho(5, 6)]$. Note that in terms of information leakage costs and in their ability to sustain cooperation the three structures in Panels A, B, and C of Figure 5 are equivalent.

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.

The following Corollary to Proposition 2 characterizes the optimal information structure if the probabilities of detection are symmetric.

**Corollary 1** *When each agent is equally likely to be detected ($\alpha_i = \alpha_j$ for all $i, j$) then $i^* = N$: the optimal information structure with $N$ links is a binary cell structure.*

The optimal information structure described in Corollary 1 is illustrated in Figure 1 for $N = 4$. Corollary 1 follows from the fact that in the symmetric case, the characterization

[24] We show this claim in the second step of the proof of Proposition 2 in the Appendix.

[25] In particular, if $\alpha_1$ and $\alpha_2$ approach zero, all these link have a arbitrarily small information leakage cost, so the organization information leakage cost is the same as in anarchy.

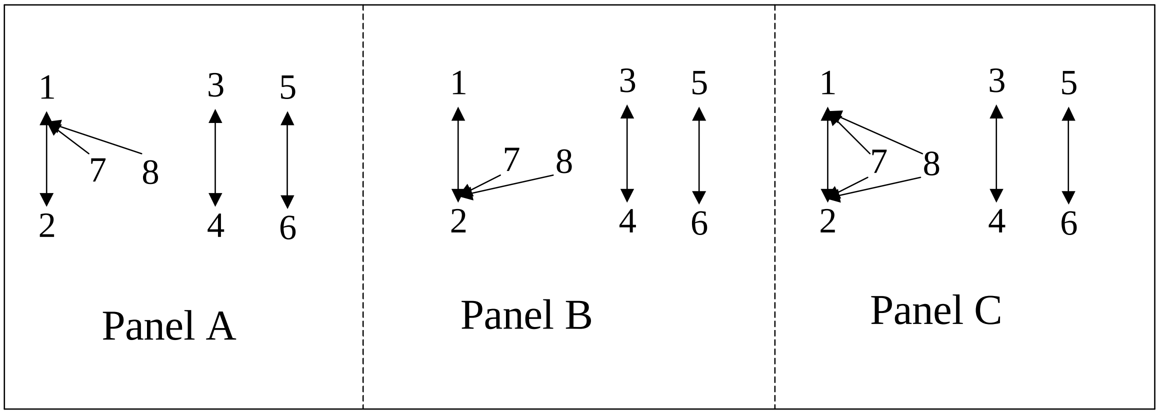


Figure 5: Equivalent optimal structures with $N = n = 8$ and $i^* = 6$.

in Proposition 2 implies $i^* = N$. Indeed, if all the agents have the same probability of detection,a it is never optimal to link one agent to another agent who is already linked to someone else rather than having him forming a cell with another agent.

The characterization of the optimal organization given a number of linked agents $n$ that we carried out in Propositions 1 and 2 allows us to define the information leakage cost function $C : \{0, .., N\} \to R$ as follows

$$C(n) = \begin{cases} b\sum_{i=1}^{N}\alpha_i & n = 0 \\ b\sum_{i=1}^{N}\alpha_i + b\alpha_1 \sum_{j=N-n+1}^{N}(1-\alpha_j) & n = 1, .., N-1 \\ b\sum_{i=1}^{N}\alpha_i + b(\alpha_1 + \alpha_2 - \alpha_1\alpha_2)\sum_{i=i^*+1}^{N}(1-\alpha_i) & n = N \\ +b\sum_{i=1}^{\frac{i^*}{2}}\left[(1-\alpha_{2i-1})\alpha_{2i} + (1-\alpha_{2i})\alpha_{2i-1}\right] & \end{cases}$$

The function $C$ represents the total information leakage cost of the organization. Notice that Corollary 1 implies that when $\alpha_i = \alpha$ for all $i$, then $i^* = N$ and $C(n) = Nb\alpha + nb\alpha(1-\alpha)$ for all $n$.

## 4.2 Optimal Number of Links

In the previous section, we characterized the optimal information structure *given a number of linked agents,* and we derived the cost function $C$ that the organization has to incur to link any number of agents. In this section, we compare this cost with the benefit that information diffusion throughout the organization provides and thereby characterize the optimal information structure.

First of all, it is useful to characterize the circumstances in which information links induce more cooperation. In order to do so, for $m \geq 1$ and $k \geq 0$, let $\Delta(m,k)$ be defined as $\Delta(m,k) \equiv \frac{c-\lambda}{\lambda m+k}$. Observe that $\Delta(m,k)$ is decreasing in both $m$ and $k$.

**Proposition 3** *When the discount factor is sufficiently high or sufficiently low ($\delta \geq \Delta(N-1,0)$ or $\delta < \Delta(N-1,k)$), the most efficient structure is anarchy.*

Recall, by Definition 1, that an anarchy is a structure with no links. Proposition 3 (whose proof is in the Appendix) states that if $\delta$ is either high enough or low enough, the most efficient information structure is anarchy. This is because if $\delta$ is high enough (i.e., higher than $\Delta(N-1,0)$), full cooperation can be reached in anarchy. Then, since we are comparing organizations by looking at the most cooperative equilibrium that can be reached under them, adding links in the organization would induce additional information leakage costs but no benefits. On the other hand, if $\delta$ is low enough (i.e., lower than $\Delta(N-1,k)$), even the threat of the additional punishment that the information exchange yields (i.e., the payment $k$) is insufficient to induce cooperation. Then, in this case also, linking agents to each other induces a positive information leakage cost but no benefits. As we show in the following results (whose proof is in the Appendix), in the range $\delta \in [\Delta(N-1,k), \Delta(N-1,0))$, an organization in which information is exchanged can achieve a strictly better outcome than an anarchy.

**Lemma 3** *If $m$ other agents cooperate, then (1) in equilibrium it is possible to sustain cooperation from an agent who is not linked to other agents if and only if $\delta \geq \Delta(m,0)$. (2) If in addition the agent is linked to another agent, then cooperation can be sustained in equilibrium if and only if $\delta \geq \Delta(m,k)$.*

Since $\Delta(\cdot,\cdot)$ is decreasing the second argument, Lemma 3 guarantees that information exchange is beneficial because it increases the range of $\delta$ under which any agent can be induced to cooperate.

Given Lemma 3, we are ready to characterize the optimal information structure. On the side of the links' benefits, suppose we have $m$ linked agents and one wants to add another linked agent. Then, for all $\delta$ in the interval $[\Delta(m,k), \Delta(N-1,0))$ we get one more agent cooperating and an increase in production of $N\lambda - c$. Notice that as the number of links increases, the critical $\delta$ for which an agent can be induced to cooperate decreases, so an increase in production is easier to achieve.

To start studying the trade-off between links' benefits and costs for any $\delta$, we can define the set of *effective* organizations.

**Definition 4** *An organization is effective if it has a number of links sufficient to induce cooperation from the agents who reveal their information to someone else. In particular, let $m(\delta)$ be the smallest integer $m$ for which $\delta \geq \Delta(m-1,k)$. Then, an organization is effective if it has at least $m(\delta)$ linked agents.*

The definition of an effective organization captures the fact that, in our model, for each $\delta$ there is a critical minimal level of cooperation that induces a linked agent to cooperate. If the number of agents who cooperate (because they are linked to someone else) is below that critical level, the threat of the punishment $k$ is not sufficient to induce cooperation. Thus, any link created in an organization that is not effective (i.e., in which the linked agents are less than $m(\delta)$) has additional information leakage cost and no benefits. This implies that an organization that is not effective is never optimal since it is always dominated by anarchy. Thus, if $\delta < \Delta(N-1,0)$, the production level of an organization as the number of linked agent $n$ increases, can be represented by the function $W : \{0,..,N\} \rightarrow R$ defined as follows:[26]

$$W(n) = \begin{cases} 0 & n \in \{0,..,m(\delta)-1\} \\ n(N\lambda - c) & n \in \{m(\delta),..,N\} \end{cases} .$$

Now, among the number of links that make the organization effective, let us identify as $n^*(\delta)$ the one that maximizes the value of the organization.

**Definition 5** *Let $n^*(\delta)$ be the integer in $\{0,..,N\}$ that maximizes $W(n) - C(n)$.*

[26] Notice that if $\delta \geq \Delta(N-1,0)$, by Proposition 3 we achieve full cooperation in anarchy, so, as we are restricting our attention to the most efficient equilibria of the repeated game, the benefit of the links is zero.

At this point we are ready to characterize the optimal information structure. Recall that by Proposition 3, if $\delta \geq \Delta(N-1,0)$, the optimal structure is anarchy. In Proposition 4 we focus on the case $\delta < \Delta(N-1,0)$.

**Proposition 4** *If the discount factor is below the level where co-operation is self-sustaining in the absence of information links ($\delta < \Delta(N-1,0)$), the optimal information structure is described as follows: (1) It is a hierarchy with $n^*(\delta)$ subordinates as described in Proposition 1 if $n^*(\delta) \in \{1,..,N-1\}$ (2) It is an organization with $N$ links as described in Proposition 2 if $n^*(\delta) = N$ (3) It is anarchy if $n^*(\delta) = 0$.*

Proposition 4 fully describes the optimal information structure given a detection probabilities distribution $\{\alpha_1,..,\alpha_N\}$. To identify the optimal information structure, one should start by looking at all the possible hierarchies as described in Proposition 1. As more and more agents are linked to agent 1, the additional cost of information leakage also increases. If among these hierarchies there are effective ones, the most efficient one among them is the one before the additional benefit of one more link $(N\lambda - c)$ crosses its additional cost.

Finally, if there are no effective organizations among the organizations considered in Propositions 1 and 2, the optimal organization is of course an anarchy.

Corollary 1 implies that when $\alpha_i = \alpha$ for all $i$, the additional cost of each link is constant and equal to $b\alpha(1-\alpha)$. Thus, the optimal structure is either an anarchy or a set of binary cells.

**Corollary 2** *Suppose $\alpha_i = \alpha$ for all $i$. If $\lambda N - c > b\alpha(1-\alpha)$ and $\delta \in [\Delta(N-1,k), \Delta(N-1,0))$ then the optimal structure is a binary cell structure. Otherwise, the optimal structure is an anarchy.*

This concludes the characterization of the optimal information structure for a given detection probability distribution $\{\alpha_1,..,\alpha_N\}$. In the next section, we endogenize such probabilities and discuss the strategic issues regarding the external agent.

# 5 Agent-Based Detection: The External Authority

In the previous section, we took the external authority's behavior as exogenously given, and we characterized the optimal information structure given an agent-based detection probability distribution $\{\alpha_1,..,\alpha_N\}$. In this section, we focus on the external agent's strategic choices and derive some normative results on optimal detection strategies of secret organizations.

As we discussed in section 3.4.2, in what follows, 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 external agent has a budget $B \in \left(0, \frac{N}{2}\right)$ available for direct detection.[27] The problem of the external agent is to allocate $B$ to determine the probability $\alpha_i$ of detection of each agent $i$ such that $\sum_{i=1}^{N} \alpha_i \leq B$. The external authority acts first and chooses these probabilities of detection before the organization forms. Though we discuss (and relax) the timing assumption of the game in Section 8, 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.

In the next result, we characterize a (weakly) optimal strategy for the external authority to determine how to allocate its resources. This strategy is weakly optimal because there are some value for $\delta$ such that the strategy of the external agent is irrelevant, and there may be other strategies that achieve the same result.

**Proposition 5** *The strategic agent's optimal strategy is to set detection probabilities symmetrically ($\alpha_1 = \alpha_2 = ... = \alpha_N = \frac{B}{N}$) if $b\frac{B}{N}(1 - \frac{B}{N}) > N\lambda - c$ and, otherwise, it is to not investigate one agent and detect all others symmetrically (set $\alpha_1 = 0$ and $\alpha_2 = ... = \alpha_N = \frac{B}{N-1}$).*[28]

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

However, if $b\frac{B}{N}(1 - \frac{B}{N}) < N\lambda - c$, by Corollary 2, a symmetric allocation would yield for all $\delta \in [\Delta(N-1, k), \Delta(N-1, 0))$ 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 5 addresses this question and suggests that, in

[27] By assuming that $B < N/2$, we ensure that the authority prefers to spend all its budget. Otherwise, we may have situations in which the authority prefers to set $\alpha_1 = ... = \alpha_N = 1/2 < B/N$ and dispose of part of the budget. This adds cases to consider, and brings little additional economic insight.

[28] Note that this allocation is feasible as $\frac{B}{N-1} < 1$ because $B < N/2$ and $N > 2$.

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.

We show that there is no other allocation that strictly dominates $\alpha_1 = 0$ and $\alpha_2 = ... = \alpha_N = \frac{B}{N-1}$ if $b\frac{B}{N}(1-\frac{B}{N}) > N\lambda - c$. The intuition for this part of Proposition 5 is the following. First of all, notice that if two agents remain undetected ($\alpha_1 = \alpha_2 = 0$), following the characterization of Proposition 4, the organization can form $N$ links without incurring any additional information leakage costs with respect to the cost they would incur in anarchy (this is because one of the two agents will act as a hub for the other $N-1$ and he can reveal his information to the second agent without any additional information leakage cost). 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 ($\alpha_1 > 0$). Then, the agent with the lowest allocation will act as a hub in a hierarchy, as described in Proposition 1. In the Appendix, we prove that if this is the case, there are exactly $N-1$ links in such a hierarchy. Then, moving all the resources from the hub to the other agents, as suggested in Proposition 5, is equivalent to the original allocation.

Proposition 5 implies that in the case in which a symmetric allocation is unable to prevent full efficiency (i.e., if $b\frac{B}{N}(1-\frac{B}{N}) < N\lambda - c$), the best strategy available to the external agent is to allow $N-1$ links to form at no cost and try to make the last link as costly as possible by leaving one agent undetected. If that is the case, the undetected agent will become the hub of the organization. Since he is not subject to punishment from the other members of the organization, he will cooperate less often and the organization will be less efficient than a binary cell structure. Notice that there are instances (if $m(\delta) = N$) in which full cooperation is necessary for the organization to sustain any cooperation. Then, if one agent does not cooperate, the entire organization becomes non effective and cooperation falls apart. Otherwise, Proposition 5 implies that the external agent cannot prevent $N-1$ agents in the organization from cooperating.

In Section 6, we abandon the agent-based detection model we have studied so far, and we discuss the cooperation-based detection model–that is, a model in which the probability of detecting the agents increases with their level of cooperation.

## 6 Cooperation-Based Detection

In this section, we assume that, rather then being determined by the external agent, the probability of detecting an agent is an increasing function of his cooperation level in the organization. This alternative detection model is available in situations in which the organization's daily activity consists of an illegal activity, such as drug trading, gambling, etc. which the external authority is able to detect directly.

This modification to the model has two effects. First, as cooperation increases the probability of detection, it changes the incentives to cooperate. Second, we will see that it makes centralization more desirable. This is because concentrating all the information in the hands of one agent who cooperate less makes any increase in cooperation of the other agents less costly from an information leakage point of view. Moreover, notice that, with respect to the agent-based detection model, cooperation from the $N-th$ agent (the hub of the hierarchy) is less likely to occur as it would involve an increase in his probability of detection.

Recall that, as specified in Section 3.1.2, in this detection model, if $p_i$ is the probability of agent $i$ cooperating at a certain period, then agent $i$ is detected with probability $\alpha(p_i) = \alpha p_i$ with $\alpha \in (0,1)$ in that period. We also assume that cooperation is the most efficient action in anarchy, which requires $b\alpha < \lambda N - c$. In the spirit of Lemma 3, let $\widetilde{\Delta}(N-1,k) \equiv \frac{c-\lambda+\alpha b}{\lambda(N-1)+k}$.

**Lemma 4** *(1) If $m-1$ other agents cooperate, an agent who did not reveal his information has an incentive to cooperate if $\delta \geq \widetilde{\Delta}(m-1,0)$. If the agent revealed his information to someone else, he can be induced to cooperate if $\delta \geq \widetilde{\Delta}(m-1,k)$ (2) If $\delta \geq \widetilde{\Delta}(N-1,0)$ anarchy is the most efficient information structure.*

Note that, in anarchy, cooperation is harder to sustain than in the previous model because the incentive to deviate increases; this is true since the agent not only saves the direct cost of cooperating, but also reduces the probability of detection. Also, note that, in contrast to the previous model, if no agents can be induced to cooperate (as is the case when $\delta < \widetilde{\Delta}(N-1,k)$), all organizations are equivalent as information exchange has no information leakage cost. *Because of this consideration and point (2) of Lemma 4, in the rest of this section and the next, we focus on $\delta \in \left[\widetilde{\Delta}(N-1,0), \widetilde{\Delta}(N-1,k)\right)$.*

In this model, information exchange is costly only if the agent who receives the information is going to cooperate with positive probability. This is because as long as an agent does not cooperate, he cannot be detected independently by the external agent.

## 6.1 Optimal information structure

Following the analysis of Section 4, let us now characterize the optimal information structure given a number of linked agents. Note that the analysis is made more complicated than the one in Section 4 by the fact that, in this specification of the model, the level of cooperation affects the information leakage cost of any organization, so the costs and benefits of creating links in the organizations cannot be studied separately. This implies that, for any information structure, we have to check whether the organization can reach a higher efficiency by imposing a lower level of cooperation by any of its members.

Let us now move on to the characterization of the optimal structure given a number of linked agents $n \leq N$ links. Note that besides specifying the link structure, this characterization also has to specify the cooperation level of each agent in the organization.[29]

**Lemma 5** *(1) The optimal information structure with $n < N$ linked agents is a hierarchy $N-n+1,..,N \to 1$. The agents $2,..,N$ fully cooperate and agent 1 does not. (2) The optimal information structure with $N$ links is a binary cell structure in which all agents cooperate if $N\lambda - c \geq \max\left[b\alpha\left(2-\alpha\right), b\alpha\left(\frac{N+2}{2} - \frac{N}{2}\alpha\right)\right]$, or a hierarchy with $N-2$ agents cooperating and linked to the cell $\{1,2\}$ in which neither agent 1 nor 2 cooperate otherwise.*

For any number of linked agents $n < N$, if agent 1 is the agent with the lowest probability of cooperating $p_1$, and $N-n,..,N$ are the agents with the highest probability of cooperating, a hierarchy with $N-n,..,N \to 1$ is optimal. In particular, if $p_1 = 0$ and $p_{N-n+1},..,p_N = 1$, such hierarchy does not impose any information leakage cost for the organization (besides $nb\alpha$, which are the ones imposed by the cooperation of the $n$ agents and which we know is optimal given the assumption $b\alpha < \lambda N - c$). The total payoff of such hierarchy is $n(N\lambda - c - b\alpha)$.

Suppose, now, that we want to link $N$ agents. Suppose that $N\lambda - c \geq b\alpha\left(2-\alpha\right)$. If this is the case, it is efficient to fully cooperate in a binary cell structure. In Lemma 5, we show that in this case the optimal structure with $N$ links is either a binary cell structure or a hierarchy in which agents 1 and 2 hold the information of all others and do not cooperate, while the other $N-2$ agents cooperate with probability 1. Note that

[29] Since if agents do not cooperate, the links are costless, and the specified information structures would be optimal anyway, we can specify these links and efficient cooperation levels without worrying about $\delta$ being high enough to induce cooperation from the agents.

this structure is as efficient as a hierarchy with $N-2$ links and is strictly dominated by a hierarchy with $N-1$ links, which implies that it will never arise as optimal organization in the next result, when we endogenize the number of links. On the other hand, the binary cell structure is the optimal organization with $N$ links if the benefit of inducing 2 more agents to cooperate outweighs the additional information leakage costs of having agents linked to each other rather than to 2 agents who never cooperate–that is, if $2(N\lambda - c) \geq Nb\alpha\,(2-\alpha) - (N-2)\,b\alpha$, or $N\lambda - c \geq b\alpha\left(\frac{N+2}{2} - \frac{N}{2}\alpha\right)$.[30]

Given the characterization in Proposition 5, we can proceed to discuss the optimal number of links. Following how we proceeded in Section 4, we can denote by $\widetilde{C}(\cdot)$ the total information leakage cost function for a given number of linked agents. Note that, given the characterization in Proposition 5, we have

$$\widetilde{C}(n) = \begin{cases} n\alpha b & n < N \\ Nb\alpha(2-\alpha) & n = N,\ N\lambda - c \geq \max\left[b\alpha\,(2-\alpha)\,, b\alpha\left(\frac{N+2}{2} - \frac{N}{2}\alpha\right)\right] \\ (N-2)\alpha b & n = N, \text{ otherwise} \end{cases} .$$

Similarly, we can denote by $\widetilde{m}(\delta)$ the minimal number of links to make the organization effective, and by $\widetilde{W}(\cdot)$ the total benefit function given a number of links. We have

$$\widetilde{W}(n) = \begin{cases} 0 & n \in \{0, .., \widetilde{m}(\delta) - 1\} \\ n\,(N\lambda - c) & n \in \{\widetilde{m}(\delta), .., N\} \end{cases} .$$

The optimal number of links $\widetilde{n}$ maximizes $\widetilde{W}(n) - \widetilde{C}(n)$. The following proposition easily follows from the previous considerations and characterizes the optimal organization.

**Proposition 6** *If the benefits of cooperation are high enough ($N\lambda - c \geq \max\left[b\alpha(2-\alpha), Nb\alpha\,(1-\alpha)\right]$) the optimal organization is a binary cell structure in which everybody cooperates. Otherwise, the optimal organization is a hierarchy with $N-1$ links ($2, .., N \to 1$), agent 1 does not cooperate and the other $N-1$ agents fully cooperate if $\widetilde{m}(\delta) \leq N$ (while they do not cooperate if $\widetilde{m}(\delta) = N$).*

The intuition of Proposition 6 comes from the fact that, as we discussed before, in a hierarchy in which the top does not cooperate, the first $N-1$ links do not have additional

[30] If $N\lambda - c < b\alpha\,(2-\alpha)$, the optimal cooperation in a binary cell structure is zero, and the optimal organizational structure with $N$ links is again the hierarchy with agents 1 and 2 linked to each other and not cooperating and all the other $N-2$ agents linked to them and cooperating.

information leakage costs besides $b\alpha$, induced by cooperation. Then, it is never optimal to link fewer than $N-1$ agents. A possible alternative with respect to this hierarchy is to link the agents in a binary cell structure. Since, now, each agent is linked to a cooperating agent, this organization implies additional information leakage cost of $Nb\alpha(1-\alpha)$. If this cost is lower than the additional benefit of $N\lambda - c$ (generated by one additional cooperating agent), this organization is optimal.

# 7 Applications and Discussion

In this section, we first compare the two models of detection we have studied so far, and then we discuss several assumptions we made in the model.

## 7.1 Model Comparison and Applications

Proposition 6 allows us to compare the two models we have analyzed so far. Recall that in Sections 4 and 5, we studied a model in which an external agent determines ex-ante each agent's probability of detection (*agent-based detection model*). In Section 6, we analyzed a model in which the probability of detection is an increasing function of the cooperation level of the agents (*cooperation-based detection model*).

Comparing the results of the different models is interesting for two reasons. First, they correspond to alternative detection policies that may be available to external authorities. One possibility for an external authority is to invest resources in detecting the illegal activity carried out in a society. In this way, the agents that cooperate the most with the organization are the most likely to be apprehended. On the other hand, an external authority can decide ex-ante to invest resources in monitoring agents independently of how much illegal activity they carry out (for instance, taping phone calls, monitoring movements and relationships, etc.). Since our results yield predictions on how the organizations react optimally to the different policies, this comparison highlights the consequences of the authority's choice when selecting a detection strategy.

Second, there are criminal environments that fit one of the models better than the other, or situations where the law takes a stand in defining what is constitutionally acceptable detection and enforcement. In these cases, we observe an authority typically using one strategy rather than the other.

For instance, some criminal organizations, such as Mafias, carry out illegal day-to-day activities such as drug dealing, gambling, etc. In these environments, since cooperation-based detection is typically less costly and controversial, it has been the most common one used by law enforcement agencies.

On the other hand, there are organization who 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). In these situations, because of the ordinary nature of these activities, it is difficult for an external authority to apprehend agents on the basis of their cooperation with the organization. However, it is still possible for an authority to invest resources in targeting and investigating agents independently of their activities, as suggested in the agent-based detection model.

Associating the two models with the different applications brings out the positive aspects of our results and allows us to tie the normative implications of the results to some more applied setting.

In particular, we are able to state the following remark.

**Remark 7** *When all agents are similar and are treated symmetrically by the authorities, in the range of $\delta$ that allows for information links to sustain cooperation, either anarchy or a binary-cell structure arise in equilibrium in the agent-detection model, whereas in the cooperation-detection model either a hierarchy or a binary-cell structure arises.*

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 equilibrium in both models of detection. Second, in an organization that is subject to *symmetric* agent-detection (such as a terrorist organization), this is *the only alternative to anarchy.* 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.* Indeed, since a hierarchical organization offers the possibility to store all the information in the hands of one agent who, as he does not cooperate, is never apprehended and so incurs no information leakage costs, a cooperation-based detection strategy always generates some degree of cooperation in the range where a link can make a difference to cooperation.

As we discussed in the Introduction, this prediction is consistent with new evidence about Mafia organizations that suggests the presence of the so-called "third level," formed by agents who collect a lot of information about the structure and the members of the organization, but who are never detected as they never carry out any illegal activity.[31]

[31] Note that in a modification of the model in which the contribution of each agent of the organization

Our results suggest two more applied considerations that are worth highlighting.

First, Proposition 5 suggests that in an agent-based detection environment, such as terrorist detection, when a symmetric detection allocation fails to prevent the (fully efficient) binary cell structure, sometimes an asymmetric allocation of the detection resources in which one agent is left undetected and the other are equally detected can lead to a strictly less efficient criminal organization.

Second, the development of new technologies has made the agent-based detection strategy increasingly available to the authorities. This implies that long-lived organizations such as Mafias, which have been investigated mainly through the cooperation-detection model in the past, have been increasingly investigated through the agent-based model. By Remark 7, this change should imply a transition of the internal information structure of mafia organizations from the centralized structure prevalent before the 1990s to more decentralized structures. Although it is hard to verify this prediction for the lack of evidence, the ability–with the apprehension or whistle-blowing of a key agent–to detect a substantial part of the organization is evidence consistent with a hierarchical information structure. Instead, if the organization is structured in cells, the apprehensions may be more frequent but less substantial from an information point of view.[32]

## 7.2 Discussion of the Assumptions

### 7.2.1 Timing

Throughout the agent-based detection model, we have assumed that the external authority moves and chooses the agents to target 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. On the other hand, a criminal organization has fewer constraints to satisfy and is more flexible when an adjustment of strategy is needed.

In other circumstances, however, it may be more plausible to suppose that the external authority can modify its policy after the information structure is formed (making the

$(N\lambda - c)$ is constant rather then increasing in $N$, Proposition 6 implies that a hierarchy becomes *always optimal* in a cooperation-detection model for a large enough organization.

[32] It is worth noting that in the 1980s it was possible for Sicilian judges to organize the so-called "Maxi-processo," in which 324 inter-connected individuals were found guilty at the same time. These kinds of inter-connected large trials are less frequent now, and it would be interesting to understand whether this is due to a deliberate choice by the judges or a less interconnected structure of the organization.

strong assumption that the authority can fully observe the information structure $\mu$). In particular, it can adjust the detection probabilities allocation with the goal of detecting as many agents as possible (notice that, ex-post, the external agent can no longer influence the level of cooperation in the organization). If this is the 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 the optimal allocation policies we characterized in Proposition 3.1 are either a symmetric allocation (i.e., $\alpha_1 = .. = \alpha_N = B/N$) or an allocation in which one agent remains undetected and the others are monitored symmetrically (i.e., $\alpha_1 = 0$ and $\alpha_2 = ... = \alpha_N = \frac{B}{N-1}$). While the first allocation, when optimal, induces anarchy, 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).

#### 7.2.2 Harsher Punishments

Our analysis has assumed 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. We conjecture that such punishment technologies might lead to information structures similar to the ones described in this paper, with the exception that the cells, instead of being binary, like the ones we found in our results, would include a small set of agents that exchange information about each other to sustain cooperation within the cell. Then, a possible interesting direction for further research would be to characterize the optimal cell size and, more generally, optimal structure.

#### 7.2.3 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, that if an agent, say $i$, discloses his information to agent $j$, if agent $j$ is detected, agent $i$ is detected with probability $\gamma < 1$. 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 in Section 4, in which an indirect link emerges. Specifically, if $N$ links are formed in the organization, Proposition 4 prescribes that $1 \leftrightarrow 2$ and $i^* + 1, .., N \rightarrow 1$, that is agents $i^* + 1, .., N$ are indirectly liked 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. It is easy to realize that, besides this adjustment, all our characterizations of an optimal information structure are robust to this extension of the model.

# 8 Extensions and Conclusion

We now discuss some results related to several extensions of our model that we consider interesting directions for further research, and we conclude.

## 8.1 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 has the advantage of delivering a very simple benefit structure for linking agents in the organization–that is, as discussed in Section 4.2, no benefits for fewer than $m(\delta)$ links, and a constant benefit of $\lambda N - c$ for any link generated after that. We also discussed 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 considered in this paper and move on to a reduced form for both the links' benefits and costs that captures more general information technologies. For instance, depending on the application, one can think of either increasing or decreasing returns of scale in generating links, local complementarities, and so on.

## 8.2 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 6 well informs our brief discussion here.

Rather then 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. Following Proposition 4, a centralized structure entails either that the top of the hierarchy does not cooperate but enforces cooperation from all the other members of the organization, or otherwise (if there are sufficient resources available for protection) of two well-protected members of the organization who enforce cooperation from each other and from all other members of the organization. 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.

### 8.3 Prison or police?

Proposition 5 describes how the strategic detection authority should allocate a fixed budget in order to minimize the efficiency of the organization. However, there are further normative implications to be learned from the result that could lead to interesting further research. In particular, we saw that the external agent is able to prevent cooperation that would otherwise occur if $b\frac{B}{N}(1-\frac{B}{N}) \geq N\lambda - c$. Suppose, now, that detection and punishment are costly activities for the external agent. For example, assume that the cost function $L(b, B)$ captures the cost of imposing a punishment $b$ to all detected agents (for instance, the cost of building and maintaining prison facilities) and to allocate a budget $B$ to detect them. Then, Proposition 5 suggests the conjecture that, in order to minimize collusion, the external agent should minimize this cost under the constraint that $b\frac{B}{N}(1-\frac{B}{N}) = N\lambda - c$.

Second, a possible interpretation for the punishment an agent can inflict on another agent, once he knows his information, is the ability to (anonymously) disclose incriminating evidence to the authorities. If this is the case, $k$ will be correlated with $b$. Suppose, for simplicity, that $k = b$. Then, the external authority will face a trade-off in setting $b$. In particular, increasing $b$ prevents collusion since it makes more likely that $b\frac{B}{N}(1 - \frac{B}{N}) \geq N\lambda - c$. However, if the external authority is not able to raise $b$ to this threshold level, then raising $b$ actually helps rather than harms the effectiveness of the organization because it increases the cooperation within the organization. This is because, as $\Delta(N-1, k) = \frac{c-\lambda}{\lambda(N-1)+k}$, raising $b$ (and thereby raising $k$) increases the range of discount factors for which information exchange increases cooperation in the organization.

### 8.4 Conclusions

This paper presents a simple model highlighting the trade-off between concerns to increase internal efficiency (sustaining cooperation) against the threat of greater vulnerability to an external threat (increasing the probability of indirect detection). We consider two alternative detection models, and we highlight how, in anticipating the reaction of the organization, the external authority should allocate its resources for detection.

In presenting a fairly simple model, we are able to fully characterize strategies for the organization and the external authority. The results we obtain are consistent with anecdotal evidence on the structure of criminal organizations such as the Mafia and terror networks and, in particular, do not rely on interactions with the production structure. We hope that this model could be explored further and extended to derive further normative implications for the role of the authorities in detecting these organization and to study their dynamics as they grow.

## 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 (2002), "Private Enforcement and Efficiency," *Mimeo*, London School of Economics.

[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. (2003), "The Impact of Leniency Programs on Cartels," Mimeo.

[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-Armengolz 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 (2005), "Contracting without Commitment: Economic Transactions in the Political Economy of States and Mafias," *Mimeo*

[13] Dixit, A. (2004), *Lawlessness and Economics: Alternative Modes of Governance*, Princeton University Press, Princeton NJ

[14] Falcone G. (1991), *Cose di Cosa Nostra*, Rizzoli, Milan.

[15] Falcone G. and G.Turone (1982), "Tecniche di Indagine in Materia di Mafia," Report to the consiglio Superiore della Magistratura.

[16] 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.

[17] Farley, J. D. (2006), "Building the perfect terrorist cell," conference presentation.

[18] Gambetta D. (1993), *The Sicilian Mafia*, Harvard University Press, Cambridge, Mass.

[19] Garreau, J. (2001), "Disconnect the Dots," *Washington Post,* September 16, 2001.

[20] Glaeser E. L., B. Sacerdote and J.A. Scheinkman (1996), "Crime and Social Interactions," *Quarterly Journal of Economics,* 111(2), pp.507-548.

[21] 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.

[22] Harrington J. (2003a), "Optimal Cartel Pricing in the Presence of an Antitrust Authority," *International Economic Review,* forthcoming.

[23] Harrington J. (2003b), "Some Implications of Antitrust Laws for Cartel Pricing," *Economics Letters.*

[24] Harrington J. (2003c), "Cartel Pricing Dynamics in the Presence of an Antitrust Authority," *RAND Journal of Economics,* forthcoming.

[25] Maskin E., Y. Qian, C. Xu (2000), "Incentives, Information, and Organizational Form," *The Review of Economic Studies,* Vol. 67, No. 2, pp. 359-378.

[26] Piccione M. and A. Rubinstein (2004), "The Curse of Wealth and Power," *Journal of Economic Theory*, Vol. 117, 119-123.

[27] Radner R. (1992), "Hierarchy: The Economics of Managing," *Journal of Economic Literature*, Vol. 30, No. 3, pp. 1382-1415

[28] Radner R. (1993), "The Organization of Decentralized Information Processing," *Econometrica*, Vol. 61, No. 5, pp. 1109-1146

[29] Rajan R. and L. Zingales (2002), "Power in a Theory of the Firm," *Quarterly Journal of Economics,* Vol. 113.

[30] Ricart-I-Costa J. (1988), "Managerial Task Assignment and Promotions," *Econometrica,* Vol. 56 (2), 449-466.

[31] Schulhofer S. J. (2005), *Rethinking the Patriot Act: Keeping America Safe and Free*, Century Foundation Press.

[32] Smith A. and F. Varese (2001,) "Payment, Protection and Punishment," *Rationality and Society, Vol. 13(3)*, pp. 349-393.

[33] Spagnolo, G. (2003), "Optimal Leniency Programs," *Mimeo*

[34] Stille A. (1995), *Excellent Cadavers*, Pantheon Books, New York, NY.

[35] Thompson, T. (2005), *Gangs: A Journey into the Heart of the British Underworld*, Hodder & Stoughton.

[36] 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.

[37] 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.

[38] Waldman M. (1984), "Job Assignments, Signalling and Efficiency," *The RAND Journal of Economics,* Vol.15 (2), 255-267.

[39] Wasserman and Faust (1994), *Social Network Analysis: Methods and Applications*, Cambridge University Press.

[40] Williamson O. E. (1983), "Credible Committments: Using Hostages to Support Exchange," *American Economic Review*, vol.73(4).

[41] Zabojnik J. (2002), "A Theory of Trade Secrets in Firms," *International Economic Review,* vol.43.

# Appendix

**Proof of Proposition 2** *First step.* Recall that $\rho(j,i)$ is decreasing in both $\alpha_j$ and $\alpha_i$. 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 $\{\mu \in I$ s.t. if $\mu_{ij} = 1$ for some $i \neq j$ then $\mu_{ji} = 1$ and $\mu_{ik} = 0 \; \forall k \neq j\}$ the one which minimizes information leakage costs is $1 \longleftrightarrow 2$, $3 \longleftrightarrow 4$,..., $N-1 \longleftrightarrow N$. 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:

$$\begin{aligned} &2b\left[1-(1-\alpha_1)(1-\alpha_2)\right]+2b\left[1-(1-\alpha_3)(1-\alpha_4)\right] \leq \\ &2b\left[1-(1-\alpha_1)(1-\alpha_4)\right]+2b\left[1-(1-\alpha_3)(1-\alpha_2)\right] \end{aligned} \tag{4}$$

and,

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

Inequality (4) 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 (5) 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 $\{N-1, N\}, ..., \{3,4\}$ 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 $2b\left[1-(1-\alpha_N)(1-\alpha_{N-1})\right]$. If they are linked to agents 1 and 2, the cost of information leakage is $b\left[1-(1-\alpha_1)(1-\alpha_2)(1-\alpha_N)\right]+ b\left[1-(1-\alpha_1)(1-\alpha_2)(1-\alpha_{N-1})\right]$. 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

$$\rho(N-1, N) < (1-\alpha_1)(1-\alpha_2) \tag{6}$$

.

If condition (6) 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 $1 \leftrightarrow 2$, $3 \longleftrightarrow 4, ..$ and $N \longleftrightarrow N-1$. If condition (6) 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 3** 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 (as additional punishments are not possible).

Consider an anarchy and the candidate equilibrium in which everyone always cooperates except following any deviation (by anybody) from full cooperation. Then a deviation from the equilibrium strategy will yield an agent $\lambda(N-1)$ as she gains $\lambda(N-1)$ in the current period but earns nothing in all future periods, whereas cooperation yields $\frac{\lambda N - c}{1-\delta}$. Therefore, this equilibrium is sustainable if and only if $\frac{\lambda N - c}{1-\delta} > \lambda N - \lambda$, or, equivalently, if and only if $\delta > \frac{c-\lambda}{\lambda(N-1)} = \Delta(N-1, 0)$.

Consider now a situation in which an agent revealed his information to someone else. Thus, this agent can be punished by the additional payment of $k$. However, if $\delta < \frac{c-\lambda}{\lambda(N-1)+k} = \Delta(N-1, k)$, cooperation cannot by achieved even by the threat of the additional punishment. This implies that exchanging information does not have any benefits and, if $\alpha > 0$, it has the cost of increasing the probability of detection. Thus, if $\delta < \Delta(N-1, k)$, an anarchy achieves the highest efficiency.¥

**Proof of Lemma 3** To prove (2), consider a candidate equilibrium in which $m+1$ agents are supposed to cooperate at every period. Focus on an agent $i$ who is supposed to cooperate and has revealed his information to someone else. Suppose that if that agent does not cooperate at some period, all the agents revert to the equilibrium in which nobody cooperates, and all the agents who hold information about someone else (including the agent who holds agent $i's$ information) punish the agents who revealed their information to them. Then, agent $i$ will not deviate from cooperation as long as $\frac{\lambda m-c}{1-\delta} \geq \lambda(m-1)-\frac{\delta k}{1-\delta}$, or $\delta \geq \frac{c-\lambda}{\lambda(m-1)+k} \equiv \Delta(m-1,k)$. If the additional punishment $k$ is not available, he does not deviate from cooperation if $\frac{\lambda m-c}{1-\delta} \geq \lambda(m-1)$, or $\delta \geq \frac{c-\lambda}{\lambda(m-1)} = \Delta(m-1,0)$. This proves (1)¥

**Proof of Proposition 5** In order to prove this result, we prove the following Lemma first. Let $\widetilde{\alpha} \equiv \left\{0, \frac{B}{N-1}, .., \frac{B}{N-1}\right\}$.

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

*Proof:* Consider any allocation $\alpha$ that generates exactly $N-1$ links. Since $\alpha_1 \leq \alpha_2 \leq ... \leq \alpha_N$ and $\alpha_1 \geq 0$, it follows that $\alpha_2 \leq \frac{B}{N-1}$. We can compare the additional information leakage costs from the $N-th$ link, $c(N) = C(N)-C(N-1)$ and $\widetilde{c}(N) = \widetilde{C}(N-1)-\widetilde{C}(N-1)$ associated with each agent $i$ under allocations $\alpha$ and $\widetilde{\alpha}$. In order to do that, let us consider the allocation $\widehat{\alpha} \equiv \{0, \alpha_2, .., \alpha_N\}$ and first compare $\alpha$ with $\widehat{\alpha}$. Under the optimal information structures with $N$ links described in Proposition 2, given allocation $\alpha$, 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 $b(1-\alpha_i)(1-\alpha_1)\alpha_2$, while under allocation $\widehat{\alpha}$ is going to be $b(1-\alpha_i)\alpha_2$. Trivially, $b(1-\alpha_i)(1-\alpha_1)\alpha_2 < b(1-\alpha_i)\alpha_2$. In case (b), since the incremental information leakage cost for agents $i$ and $i+1$ of the $N-th$ link under allocation $\alpha$ is $b(1-\alpha_i)\alpha_{i+1} + b(1-\alpha_{i+1})\alpha_i - b(1-\alpha_1)\alpha_i - b(1-\alpha_1)\alpha_{i+1}$ 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

$$\begin{aligned}
& b(1-\alpha_i)\alpha_{i+1} + b(1-\alpha_{i+1})\alpha_i - b(1-\alpha_1)\alpha_i - b(1-\alpha_1)\alpha_{i+1} \\
< \; & b(\alpha_1+\alpha_2+\alpha_1\alpha_2)\alpha_{i+1} + b(\alpha_1+\alpha_2+\alpha_1\alpha_2)\alpha_i - b(1-\alpha_1)\alpha_i - b(1-\alpha_1)\alpha_{i+1} \\
= \; & b(1-\alpha_i)(1-\alpha_1)\alpha_2 + b(1-\alpha_{i+1})(1-\alpha_1)\alpha_2 \\
< \; & b(1-\alpha_i)\alpha_2 + b(1-\alpha_{i+1})\alpha_2
\end{aligned}$$

The last expression is the information leakage cost associated with the allocation $\widehat{\alpha}$

(that is the information leakage costs beyond those incurred in anarchy).

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

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

and

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

Since $\sum_{i=3}^{N}\alpha_i < B < N-2$, it follows that information leakage costs under $\widehat{\alpha}$ are increasing in $\alpha_2$, whose highest value is $\frac{B}{N-1}$ and when it takes this value the information leakage costs are equal to those under $\widetilde{\alpha}$. Thus $\widetilde{c}(N) \geq \widehat{c}(N) \geq c(N)$. This concludes the proof of Lemma 8.4¥

Let us now proceed to the proof of Proposition **5**.

*First step.* First of all, note that, under some circumstances, the external authority's strategy will be irrelevant. In particular, for $\delta < \Delta(N-1,k)$ and $\delta \geq \Delta(N-1,0)$ the organization will be anarchic and all agents will either not cooperate (if $\delta < \Delta(N-1,k)$) or cooperate (if $\delta \geq \Delta(N-1,0)$), regardless of the external authority's allocation of its investigative budget $B$. *In the rest of the proof we will focus on* $\delta \in [\Delta(N-1,k), \Delta(N-1,0))$.

*Second step.* Suppose now that $N\lambda - c < b\frac{B}{N}(1-\frac{B}{N})$. By Corollary 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 $\widetilde{\alpha} = \left\{0, \frac{B}{N-1}, .., \frac{B}{N-1}\right\}$ is optimal.

(1) Consider any allocation such that $\alpha_1 = \alpha_2 = 0$. Then, the organization can reach full efficiency with zero additional information leakage cost with respect to anarchy. To see this, suppose that $\alpha_1 = \alpha_2 = 0$; then, an organization with the links $\mu_{1i} = 1$ for all $i \in \{2, .., N\}$, $\mu_{21} = 1$ 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 prevents links between agent and deter efficiency, *at most one agent can be left with zero probability of detection.*

(2) Consider any allocation such that $\alpha_1 > 0$–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 $\widetilde{\alpha}$ cannot do worse than that. Suppose, then, that the allocation is asymmetric that is–$\alpha_1 < \frac{B}{N}$. Following the characterization in Proposition 4, 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 $\widetilde{\alpha}$ cannot do worse than that.

Suppose, instead, that the optimal organization given the allocation $\alpha$ we are considering generates $N-1$ links. Then by the Lemma 8.4, allocation $\widetilde{\alpha}$ performs at least as well.

Finally, suppose that under the allocation $\alpha$ the linked agents are $n < N-1$. We argue that such a structure is impossible. In such organizations, according to Proposition 1, there are three types of agents to consider: the top of the hierarchy agent 1, the $N-n-1$ independent agents $2, ..N-n$, and the $n$ agents who reveal their information to agent 1–that is $N-n+1, ..N$. 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 $\alpha_1$, $\alpha_2$ and $\alpha_N$ respectively. The probability allocations we are restricting our attention to have to satisfy the following constraints:

(i) $0 < \alpha_1 \leq \alpha_2 \leq \alpha_N \leq 1$ (by feasibility and by Proposition 1);

(ii) $b\alpha_1(1-\alpha_2) \geq N\lambda - c$ (it is not optimal for the organization to link the $N-n-1$ independent to agent 1);

(iii) $N\lambda - c \geq b\alpha_1(1-\alpha_N)$ (it is optimal for the organization to link the $n$ agents to agent 1);

(iv) $\alpha_1 + (N-n-1)\alpha_2 + n\alpha_N \leq B$ (the resource constraint).

Note that $b\alpha_1(1-\alpha_2) \leq b\alpha_2(1-\alpha_2) \leq b\frac{B}{N}(1-\frac{B}{N})$ since $\alpha_2 \leq \frac{B}{N} < \frac{1}{2}$ (otherwise either the (iv) or is violated or it cannot be that $\alpha_1 \leq \alpha_2 \leq \alpha_N$) but then (ii) cannot hold since $N\lambda - c > b\frac{B}{N}(1-\frac{B}{N})$. If 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 $\alpha_1 = \alpha_2 = 0$ or $\alpha_1 > 0$ are (weakly) dominated by allocation $\widetilde{\alpha}$. Finally, let us consider an allocation such that $\alpha_1 = 0$ and $\alpha_2 > 0$. 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 $\alpha_1 = 0$, the characterization in Proposition 2 yields, for each $i \in \{4,..,N\}$, to $\frac{2(1-\alpha_{i-1})(1-\alpha_i)}{2-\alpha_{i-1}-\alpha_i} \leq 1-\alpha_2$ (easy to check since $\alpha_2 \leq \alpha_j$ for all $j \in \{3,..,N\}$). 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 $b\alpha_2 \sum_{i=3}^{N}(1-\alpha_i)$, and it is maximized (under the constraints $\alpha_2 \leq \alpha_i$ for all $i$ and $\sum_{i=2}^{N}\alpha_1 = B$) by $\alpha_i = \frac{B}{N-1}$ for all $i \in \{2,..,N\}$, which is allocation $\widetilde{\alpha}$.¥

**Proof of Lemma 5** (1) Note that if $N-n+1,..,N \to 1$, agents $n,..,N$ cooperate and agent 1 does not, we have $n$ links at no additional information leakage cost (besides to ones imposed by the cooperation of the $n$ agents, which we know is optimal given the assumption $\lambda N - c > b\alpha$), as the probability of detection of agent 1 is zero. (2) Suppose that one wants to generate a structure with $n < N$ linked agents. First, let us analyze the optimal binary cell allocation. If the agents are all linked in binary cells, we have to find the optimal cooperation levels $p^*$, $q^*$ for each agent in a cell. The most efficient pairwise structure solves

$$\max_{p,q\in[0,1]} p\,(N\lambda - c) + q\,(N\lambda - c) - 2b\left(p\alpha + q\alpha - pq\alpha^2\right) \tag{7}$$

Note that we have a corner solution, in particular $p^* = q^* = 1$ if $N\lambda - c > b\alpha\,(2-\alpha)$, $p^*$ and $q^* \in [0,1]$ if $N\lambda - c = b\alpha\,(2-\alpha)$ and $p^* = q^* = 0$ if $N\lambda - c < b\alpha\,(2-\alpha)$.

(a) Suppose first that $N\lambda - c > b\alpha\,(2-\alpha)$, so in the most efficient binary cell structure there is full cooperation. Let us consider the following different link structure instead. In particular, let us consider a hierarchy in which $N-2$ agents, rather than being arranged in binary cells, are all linked to the cell $\{1,2\}$ (notice that this structure is the optimal one in Proposition 2, in the case in which the probabilities of detection of agents $3,..,N$ are all the same and $\rho(3,4) < (1-\alpha_1)(1-\alpha_2)$). In such an organization, it is again optimal to set $p_3,..,p_N = 1$.

The question is, since now $N-2$ agents are linked to the cell $\{1,2\}$ whether it is optimal to lower the cooperation level of such a cell to lower its probability of detection. In finding the optimal level of cooperation for 1 and 2, we can restrict our attention to a positive level of cooperation for agents 1 and 2 since if they do not cooperate at all, a hierarchy with $N-2$ links would dominate this organization. Moreover, the level of cooperation of agents 1 and 2 must be such that it is more efficient to link the other agents to them, rather than keeping them in binary cells. This is satisfied if and only if, setting $p$ and $q$ be the cooperation levels of agents 1 and 2, we have $\alpha(1-\alpha) \geq (1-\alpha)\left(\alpha p + \alpha q - \alpha^2 pq\right)$ ,or $p + q - \alpha pq \leq 1$. Note that as $\alpha < 1$, it cannot be the case that both $p$ and $q$ are equal to 1.

Overall, the optimal $p^*$ and $q^*$ solve the following problem

$$\begin{array}{cc} \max_{p,q \in (0,1]} & (N\lambda - c)(p+q) - 2b\left(\alpha p + \alpha q - \alpha^2 pq\right) + \\ p + q - \alpha pq \leq 1 & -(N-2)b(1-\alpha)\left(\alpha p + \alpha q - \alpha^2 pq\right) \end{array} \tag{8}$$

It is easy to see that optimality requires that $p^* = q^*$,so problem (8) is equivalent to

$$\begin{array}{cc} \max_{p \in (0,1]} & 2p(N\lambda - c) - \left(2\alpha p - \alpha^2 p^2\right)[2b + b(N-2)(1-\alpha)] \\ 2p - \alpha p^2 - 1 \leq 0 & \end{array} \tag{9}$$

Since the objective function of problem (9) is convex in $p$, the solution is a corner one. Since a higher $p$ tightens the constraint, we have two possible cases: (i) $p^*$ is such that the constraint is binding, that is, $2p^* - \alpha(p^*)^2 - 1 = 0$, which yields $\frac{1-\sqrt{1-\alpha}}{\alpha}$, or (ii) $p^* = 0$.

If case (i) is true, let us compare such an outcome with the binary cell structure we considered before. Notice that (as the constraint in problem (9) is binding at the optimum) the agents $3, .., N$ incur the same information leakage cost in the hierarchical structure we just considered and in the binary structure. Also, they cooperate with probability one in both cases. Then, let us focus on agents 1 and 2. Since these agents constitute a cell in both structures, they behave more efficiently in a binary cell structure, as problem (8) guarantees. Then, a binary cell structure dominates the structure we just considered. If case (ii) is true, we have a structure in which agents 1 and 2 are linked together, they do not cooperate and the other $N-2$ agents do cooperate and are linked to agent 1 (or, equivalently, to agent 2 or both).

(b) Consider now the case in which $N\lambda - c < b\alpha(2-\alpha)$. In this case, it is not efficient to cooperate in a binary cell structure. Let us consider the possibility to linking agents 1 and 2 to each other and all the other agents to agent 1. If this is the case, we know that agents 1 and 2 cannot benefit from cooperating (as the information leakage cost generated by their cooperation is going to be greater than the one they would incur in a binary cell structure, and we are in the case in which their cooperation is zero). Then, agents 1 and 2 should not cooperate. This structure would lead to $N-2$ agents cooperating at no cost.¥