Coordination Against Crime: Can Policing Efforts Influence Network Design

Shikeb Farooqui

PRELIMINARY AND INCOMPLETE*

March 2009

Abstract

We present a model of criminal activity and policing, in which a network planner recruits individuals to be part of a criminal network. Criminal activity is subject to monitoring by a local policing authority. In equilibrium the policing authority assigns an individual criminal a probability of detection that is a weighted average of his Bonacich and Intercentrality. Police monitoring influences the network primitives and higher skilled policing can induce the network planner to design small decentralised networks. We then proceed to show how coordination between different local policing authorities can influence the global network topology and lead to location specific centralisation.

*All comments and suggestions welcome at shikeb.farooqui@upf.edu

1 Introduction

We study a model of strategic interaction between a criminal organisation and an independent local policing authority that monitors the activity of the criminal organisation. The criminal organisation recruits agents willing to partake in the criminal activity and connects them in a network designed by a network planner. The policing authority monitors the activity of the network and assigns a probability of detection to each network participant based on its understanding of the network primitives. We use the model to predict how equilibrium probability profiles depend on network characteristics through the policing authority's skill and resource allocation. We also show that the degree of centralisation in criminal operations is a complex function of the type of actions being undertaken and the resources available to the network planner relative to policing efficiency.

The second part of the paper uses some of the salient features of the crime and punishment framework to build another model on coordinated policing and network design. We consider a set of independent local policing authorities that decide whether to coordinate policing amongst one another. Coordination exhibits strategic complementarity and is costly to induce. The network planner sets up locally independent networks with the aim to maximise global criminal activity. Under a strict assumption that restricts the network planner from designing a global network that induces cooperation between policing authorities we analyse the degree of local centralisation that arises in equilibrium.

Our analysis is built on the assumption that like any other activity, criminal behaviour is a function of the trade-off between global rivalry (substitutability) and local complementarity (information sharing) in actions. We argue that in equilibrium criminal actions will depend on the policing authority's ability to predict the true extent of this trade-off which in turn will impact the optimal design of each local cell. Differences in policing ability and resources will extend to the choice of cooperation and coordination and lead to optimal networks that exhibit location specific centralisation.

In formulating this story we explicitly have the case of an international terrorist organisation in mind, however the normative issues considered and policy recommendations expressed here can easily be extended to criminal activity that extends beyond national boundaries.

Terrorist acts are by no means a new phenomena, but the twenty-first century has seen a noticeable resurgence in such activity. What defines the recent shift in the terrorist paradigm is the blurring of boundaries between the universal and particularistic causes of this activity which make standard carrot and stick procedures for deterrence harder to implement. More importantly terrorist activities and organisation are no longer constrained to national borders inducing a direct need for international cooperation in policing efforts.

As the need for combatting terrorism grows a pertinent question that arises relates to the mode of combat. Should governments treat such acts as criminal activity and use the standard criminal justice model of policing and judiciary or should they resort to stricter measures and military warfare tactics? To a certain extent governments have been forced to adopt a combination of both.

From the policing perspective almost all governments that have been subject to terrorist attacks in the recent past have substantially enhanced policing power, resources dedicated towards reducing further incidents. The effectiveness of these policies cannot be evaluated without first understanding how criminal organise their operations. If we know the organisational structure of an operation then we are well placed to target it optimally.

Criminal networks balance the need for coordinated activity with the requirement to conceal and Baker and Faulkner (1993) postulate that networks based activities requiring high information exchange that are not easy to conceal should group in a decentralised way. More formally, Bar-Isaac and Baccara (2006) shows that the exchange of potentially incriminating information between criminals can act as a coordination device and rationalise both hierarchical structures and organisation in cells.

Given a particular network structure how does one optimally disrupt the network. As Borgatti (2003) argues this is a problem of identifying the key player in the network. The identity of the key player will necessarily be a function of underlying actions taken within the network and the nature of information flow which will determine the order of influence each node has in the network. The idea of nodal influence within the network has been captured by the notion of centrality in the social networks literature. A plethora of centrality measures exist in the literature each with its own ad hoc evaluation criteria.

One particular measure of centrality suggested by Bonacich (1987) has been shown by Ballester et al. (2006) to be extremely useful in the economic context. This measure of centrality will also play a pivotal role in our analysis. Simply put a more central node exerts greater influence in the network and removing it from the network will lead to greater disruption.

Since centrality measures can be very sensitive to network topology this line of reasoning hinges on perfect knowledge of the network structure. Given that in reality one does not perfectly know the network architecture a reasonable criteria to assess the usefulness of different centrality measures may be to gauge how well they perform when network structures are partially sampled.

Other related lines of research have looked which network structures are particularly vulnerable to random targeting of nodes.

Related Literature

Our primary motivation is to write an economic model rationalising an individuals decision to enter a criminal network and the organisational structure of criminal activity. Using this theory we also wish to look at how effective various tools available to policing authorities are, characterise optimal detection strategies and evaluate how strategic interaction between the two organisations impacts their individual organisational structure.

In defining the utility function of the individual criminal we adopt the prescription forwarded by Becker (1968). This is becoming increasingly standard procedure when studying the cost and benefits associated with crime at the individual level (see for example Polinsky and Shavell (1999) and Garoupa (1997)). Ballester (2008) study a model of delinquent networks in which criminals also have the option of entering the labour market. This is an individual choice decision and individuals self-organise into stable criminal networks if labour market opportunities are prohibitive. Individuals factor into their decision process the possibility that the resulting network will be subject to targeting and this can reduce the size of the network and also influence structure. Our work follows Ballester (2008) closely, however, we assign the network design problem to a network planner.

Bar-Isaac and Baccara (2006) also looks at individual sorting into criminal networks. The enforcement device here is the sharing of incriminating information with ones neighbours to build trust. This information can potentially be used for internal punishment but also leaves the organisation vulnerable to external threat. They show that depending on the nature of law enforcement individuals either prefer to sort into binary cells or hierarchical structures which in turn influences a policing authorities optimal strategy. We will also motivate our network structure in terms of information exchange, however the role of this information exchange will be to enhance production.

Farley (2003,2006) studies the robustness of terrorist networks to ordered targeting and finds that a hierarchical organisation is least vulnerable. BCZ study the problem of optimal targeting when effort levels are linked through strategic complementarities embedded in the network structure. The key player problem amounts to finding the individual that has the highest influence on the production decisions of all other network participants.

Garoupa (2007) studies the trade-off between enhancing internal productivity of an illegal activity while and increasing the vulnerability of detection of organisation members. Taking the internal structure as given he focuses on the optimal size of the criminal organisation.

Whereas in the first part of the model we require policing authorities to act independently, in the second part we allow them to coordinate efforts. One interpretation of the superior policing technology they obtain through coordination could be a move to military methods. Poveda analyse a two stage repeated setting game in which countries coordinate military effort in the first stage thereby reducing aggregate terrorist resources and then decide on defensive policing measures in the second stage. They show the existence of a set of countries that coordinate militarily in the first stage that depends on political and economic power and military might. More generally there is newly emerging game theoretic literature on counterterrorism[1].

Our work is also closely related to organisational structure and in particular endogenous coordination in organisations. Papers concerned with internal efficiency and organisational form include Maskin et al. (2000) who study the impact of organisational form on incentives and Radner (1990), Radner (1993), VZ and Garicano and Rossi-Hansberg (2006) that study the role of hierarchical organisations in improving information processing. However, in the illegal context Klerks (2001) argues that we should not directly assume that criminal networks are organised hierarchically just because the organisational literature suggest so.

A number of recent papers analyse coordination in organisations. looks at the trade-off between coordination and the private benefits of independent actions. In decentralised structures managers do not fully internalise the benefits of coordination, whereas under centralised activity the decision maker fails to take into account private benefits. Hart and Moore (2004) looks at the optimal hierarchical structure that arises when some agents specialise thoughts about the coordinated use of some assets, while others develop ways to independently use a particular asset. Dessein and Matouschek (2008) compare centralised and decentralised coordination in the firm context when managers are privately informed and communicate strategically and find that higher coordination improves horizontal communication at the expense of vertical communication causing decentralisation to dominate centralisation. Finally, focusing on cases in which divisions differ in their need for coordination Rantakari (2008) analyses coordination in organisations in which information is dispersed. He argues that it in be optimal design asymmetric organisational structures in which all decision rights are concentrated in one division.

Our paper is also related to the burgeoning literature on peer effects, criminal behaviour and unemployment. Verdier and Zenou (2001), for example look at the link between racial beliefs and crime and Burdett et al. (2003) study the interaction between crime and unemployment[2].

And somewhat related to the literature on organised crime although much of this literature has focused on role organised crime plays in providing governance

[1] See for example Atkinson et al. (1987), Lapan and Sandler (1988), Sandler and Siqueira (2006) and Sandler and Arce (2007).

[2] See also Huang et al. (2004). For an empirical study of social interactions and crime see Glaeser et al. (1995)

and private contract enforcement[3]. Recently Berman and Laitin (2005) propose and test empirically a club framework emphasizing the function of voluntary religious organisations as efficient providers of public goods.

Other papers that rationalise religious extremism and its connection to terrorist activities include Benmelech and Berrebi (2007) and Berman (2003).

2 The Model

The model we present studies the interaction between three sets of agents: a) Individual criminals i, b) a network planner T and c) a set of independent policing authorities L. We wish to focus primarily on the strategic interaction between the network planner T and the set of independent policing authorities L, and we will simplify the strategic interaction between T and individual recruits as much as is feasible to retain qualitative insights.

The network planner links individual criminals in a network ensuring that the participation constraint of each individual recruit is satisfied and that the resulting network is stable. Each criminal recruit is subject to policing by his local authority. His position in the network determines his effort level x_i and his detection profile p_i. In equilibrium these will be linked to his Bonacich Centrality and Intercentrality respectively.

The interaction between all three sets of agents provides the micro-foundation for the extended game in which we are primarily interested. The extended game studies the strategic interdependence of organisational structure between T and L. Local policing authorities are allowed to act independently mimicking decentralised activity or coordinate operations in a centralised fashion.

Does the coordination decision of local policing authorities impact the degree of centralisation of the criminal network? We answer this question by allowing T to choose from centralised and decentralised networks at the local level and look at equilibrium actions.

The remainder of this section states the concept of a network and the notions of centrality and centralisation before introducing the actors in the game and their preferences and actions.

2.1 The Network

A network or graph g consists of a set of n agents (recruits) linked by a set of weighted edges, where $g_{ij} \in [0,1]$ is the weight assigned to the link $(i \to j)$ when i is directly connected to j. We consider unweighted networks with at most one link

[3]Good examples from this strand of literature include Bandiera (2003) and Mesquita and Hafer (2008) amongst others.

between i, j for any $i, j \in g$, where $g_{ij} \in \{0, 1\}$, such that $g_{ij} = 0$ when $(i \nrightarrow j)$ and $g_{ij} = 1$ when $(i \rightarrow j)$.

We restrict attention to undirected networks where $g_{ij} = g_{ji}$ for all $i, j \in g$. We also assume that the network has no self-loops i.e. $g_{ii} = 0$ for all $i \in g$. A recruit $i \in g$ is isolated if $g_{ij} = 0$ for all $j \in g$. An isolated recruit is not considered part of the criminal network and his output is normalised to 0.

A path in g of length k from i to j is a sequence $p = \{i_0 \rightarrow i_1 \rightarrow ... \rightarrow i_k\}$ connecting recruits in k-steps such that $i_0 = i$, $i_k = j$, $i_p \neq i_{p+1}$, and i_p and i_{p+1} are directly linked. Let $g^k_{ij} = 1$ denote the fact that $i \rightarrow j$ indirectly through a path of length k. A connected graph is a network structure in which $g^k_{ij} = 1$ for all pairs of recruits $i, j \in g$ and some finite length k. In particular we are interested in two specific network structures in g. Before defining these structures we will spend a couple of lines discussing the concept of network centralisation.

Network Centralisation

To each node in the network we can assign a centrality measure that captures the importance and/or power of that node relative to all other nodes in the network. The information content of this measure depends on the precise definition of centrality used, however all centrality measures suggested in the literature induce a node ranking over the network structure.

When talking about network centralisation we are moving from a discussion of individual centrality to that of group centrality and heterogeneity within it. In the social networks literature, the standard procedures used to evaluate centralisation either rely on assessing by how much the highest individual centrality differs from the rest, or by also looking at how individual centrality is dispersed.

We are primarily interested in the strategic interdependence between the organisational structure of the criminal outfit and the local policing authorities. In particular we wish to evaluate how this interdependence impacts the centralisation of criminal activity. Below we identify clear examples of decentralised and centralised networks that meet both the centralisation criteria discussed above.

A Decentralised Network

Let d denote a fully connected graph in which all n recruits are directly connected to all other $n - 1$ recruits i.e. $g^1_{ij} = 1$ for all $j \neq i \in g$. In this particular network structure each recruit has exactly the same of connections, $n - 1$ and is connected to every other recruit in the network, although he may not know this. Any information originating from recruit i can potentially flow, and in the model we present will flow directly to all other recruits. From a more technical viewpoint any measure of centrality that we may consider will assign the same value to each node in this structure. As there is no disparity in node centrality between recruits, this structure gives us the ideal candidate for modeling decentralised activity.

A Centralised Network

Let c denote a star graph consisting of one hub h directly connected to all other $n-1$ recruits i.e. $g^1_{hj} = 1$ for all $j \neq h \in g$ and $n-1$ peripheries such that $g^2_{ij} = 1$ for all $i, j \neq h \in g$. In this structure any recruit that is not the hub can only share information directly with the hub. Whereas the activities of the hub are directly visible to the periphery, only the hub can see the activities of individual periphery recruits through his direct links. Although the periphery nodes are as central as one another, for any measure of centrality we may consider their centrality is negligible compared to that of the hub. The social networks literature considers the star graph the most centralised structure in the set of all graphs. We follow this prescription and use the star for modeling centralised activity.

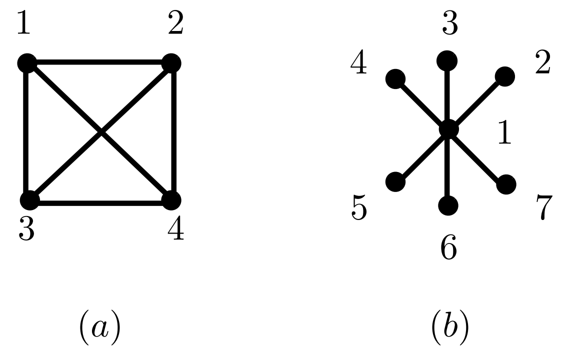


Figure 1: (a) A decentralised network - d, (b) A centralised network - c

Figure 1 depicts an example of d for $n = 4$ and c for $n = 7$. It is evident that both structures use the same number of direct links but altering the distribution the centralised network is able to connect more nodes.

2.2 Network Centrality

The social networks literature offers a plethora of criteria, not always micro-founded, for assessing the centrality of a node. Our preferred measures of centrality are Bonacich and Intercentrality. However, two other notions of centrality that are widely known and used are that of degree and betweenness centrality. We mention them here because they are closely related to our preferred measures.

Degree centrality counts the number of direct links each nodes has. A node with more direct links is considered more influential and hence more central in the network. In Figure 1, each node in d has the same number of direct links, hence there is no dispersion and the network is decentralised. In c node 1 has centrality of 3 and all other nodes have centrality 1, hence c is centralised. As we will see Bonacich Centrality is closely related to degree centrality.

When interested in information flow or influence within a network Betweenness centrality measures how well a node bridges, through the shortest path, two other nodes that are not directly connected. In Figure 1, node 1 in c controls the information flow from $2 \rightarrow 3$ who never directly interact, hence he should be more central than any of the two peripheries. Since 1 lies on the shortest path for three distinct interaction pairs $(2,3),(2,4),(3,4)$ he has a Betweenness centrality of 3. The peripheries do not lie on any shortest path communication directed to someone else hence all three have a Betweenness centrality of 0. According to this criteria c is again a centralised structure. On the other hand none of the nodes in d lie on the shortest path of a communication stream not directly intended for them hence each node has a centrality of 0. Therefore, d is decentralised according to this criteria. As we will see Intercentrality is closely related to the concept of Betweenness.

The model that we propose on criminal behaviour and network design will, in equilibrium, depend on the criminal recruits Bonacich and Intercentrality in the network. These are explained below.

Bonacich Centrality

The nature of direct links in g can be be succinctly summarised in the n-square adjacency matrix G, so that entry $g_{ij} = \{0, 1\}$ in G captures the direct link between (i, j). Because we are dealing with undirected graphs, G is symmetric. Let G^k be the $k-th$ power of G with coefficients g_{ij}^k for some integer k. The number of indirect paths between (i, j) of length k are counted by $g_{ij}^k \geq 0$ and stored in G^k.

Let $\theta \geq 0$, to be defined more precisely in later sections, be a decay factor such that $m_{ij}(g, a\theta) = \Sigma_{k=0}^{+\infty} \theta^k g_{ij}^k$ counts the total number of paths in g starting at i and ending at j, where paths of length k are weighted by some factor θ^k. The Bonacich Centrality of recruit i, $b_i(g, \theta) = \Sigma_{j=1}^{n} m_{ij}(g, \theta)$, counts the total number of paths in g starting from i.

Equivalently we can capture this information in the form of a matrix where $m_{ij}(g, \theta)$ are the coefficients of the following matrix:

$$\mathbf{M}(g, \theta) = [\mathbf{I} - \theta\mathbf{G}]^{-1} = \sum_{k=0}^{+\infty} \theta^k \mathbf{G}^k$$

So that the vector of Bonacich Centralities can be written in the following form:

$$\mathbf{b}(g, \theta) = \mathbf{M}.\mathbf{1} = [\mathbf{I} - \theta\mathbf{G}]^{-1}.\mathbf{1}$$

However, in order for $\mathbf{M}$ to be well defined we require the decay factor θ, that scales down the relative weight of longer paths, to be small enough. In particular we require that θ is less than the inverse of the largest eigenvalue of G, denoted $\rho(g)$[4].

The above restriction on θ is required to ensure that the coefficients $m_{ij}(g, \theta)$ are bounded and do not explode. In fact the relationship between $|\theta - \frac{1}{\rho(g)}|$ will be crucial in equilibrium and for this reason we state the following technical lemma.

Lemma 1. *Let $\rho(g)$ denote the largest eigenvalue of $g \in c, d$. For a given n let $g(n)$ denote a graph of type g with n connected nodes. Then the following is true:*
(1) $\rho(c(n)) = (n-1)^{\frac{1}{2}}$
(2) $\rho(d(n)) = n - 1$

The model presented below considers decentralised and centralised networks of differing sizes. In order to facilitate comparison of aggregate effort levels resulting from the different network types we make the following assumption.

Assumption 1. *For any $n' \geq n$ we require that $n' < (n-1)^2 + 1$.*

In the analysis that follows this assumption will always be satisfied. We state it here explicitly because it helps clarify the underlying forces of the model. As long as this assumption is satisfied the largest eigenvalue of $d(n)$ will always be greater than the largest eigenvalue of the largest $c(n')$ we consider. This means that a decentralised network will always have a more restrictive bound than any centralised network and $|\theta - \frac{1}{\rho(d)}| < |\theta - \frac{1}{\rho(c)}|$ for any (c, d) and $n' \geq n$.

We can re-write b_i in the following manner:

$$b_i(g, \theta) = m_{ii}(g, \theta) + \sum_{j \neq i} m_{ij}(g, \theta)$$

A node's Bonacich Centrality is increasing not only in his direct links, i.e. his degree centrality, that in turn allow him to access a higher number of indirect links but also in θ which allows the information generated at node i to travel farther distances thus compounding the effect of higher direct links.

It is clear that $b_i(g, \theta)$ is increasing in θ. However, the number of paths a nodes influence can traverse in the network depends not only on θ but also on $|\theta - \frac{1}{\rho(g)}|$.

Let $\mathbf{b} \in \mathbb{R}^n$ be a vector that gathers the Bonacich profiles of all the nodes in the network, and let $b = \sum_{i=1}^{n} b_i$ denote the sum of these coordinates. In equilibrium we will be interested in this sum and the above remark applies equally strongly to $b(g, \theta)$. The implications of this remark on aggregate Bonacich centrality will help us build a preference ranking over various types of network structures.

[4]Because we are dealing with a square and symmetric matrix G, the Perron-Frobenius theorem guarantees that the largest eigenvalue of G is real and positive.

Figure 2 makes clear that the gradient of $b(g, \theta)$ is also increasing in θ and only really starts to asymptote as θ approaches its bound. For the two structures depicted in Figure 2 we also see that the aggregate Bonacich curves cross only once. For low θ the centralised network produces a higher aggregate Bonacich count but as $|\theta - \frac{1}{\rho(d)}|$ becomes small, aggregate Bonacich is able to count more paths in d compared to c. This observation anticipates a result that will prove useful in categorising the equilibrium and which we will develop further on.

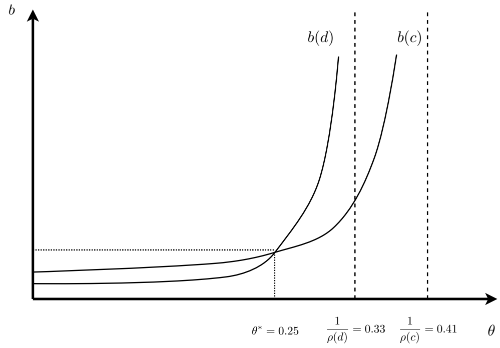


Figure 2: A comparison of total Bonacich in centralised and decentralised networks

One last thing to note is that b_i includes the number of self-loops, $m_{ii}(g, a)$. This is information that originates at i but flows back to i after mixing with information from other nodes $j \neq i$ in path k. By definition $m_{ii}(g, \theta) \geq 1$ therefore $b_i(g, \theta) \geq 1$. Equality is attained when $\theta = 0$[5].

In order to measure the pure influence node i has on the network excluding himself we turn to our other preferred measure of centrality.

Intercentrality[6]

Knowing the Bonacich Centrality and number of self loops of recruit i, we can define the Intercentrality of recruit i as:

$$r_i(g, \theta) = \frac{b_i(g, \theta)^2}{m_{ii}(g, \theta)}$$

Intercentrality counts the total number of paths in g ending at i and therefore measures the amount of new information flowing to i from elsewhere in the network. A node that exhibits high Betweenness centrality should also, on average, exhibit high Intercentrality depending on the magnitude of θ.

[5]We are using a normalised Bonacich Centrality. In the original Katz-Bonacich measure $\mathbf{b} = [\mathbf{I} - \theta\mathbf{G}]^{-1}.\mathbf{G}.\mathbf{1}$, and in this case Bonacich Centrality is exactly the same as degree centrality when $\theta = 0$.

[6]This measure is due to Ballester et al. (2006)

In a quadratic utility setup where the utility of agent i depends, through strategic complementarities, on the effort levels of all other agents in the network Ballester et al. (2006) show that removing the agent with the highest Intercentrality from the network maximally reduces the output of the resulting network.

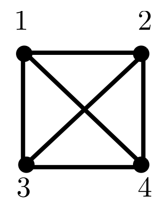
As we will see Intercentrality will also play a role in our model through its impact on the probability of detection.

A Numerical Example

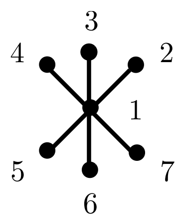
Table 1 presents the Bonacich and Intercentrality for the two network structures in Figure 1 assuming two different values for θ. Also presented are the eigenvalues for each network, the inverse of which exceed θ. Later on we will give a general characterisation of the two centrality measures based on the model primitives. What is clear is that both measures of centrality are increasing in θ and that according to our criteria of centralisation the star network remains centralised for both centrality measures, whereas the fully-connected network is completely decentralised.

$n_d = 4, n_c = 7$

	$\theta = 0.1$		$\theta = 0.2$	
	d	c	d	c
b_1	1.43	1.70	2.5	2.89
$b_i : \{i \neq 1\}$	1.43	1.17	2.5	1.58
r_1	1.96	2.72	5.0	6.37
$r_i : \{i \neq 1\}$	1.96	1.36	5.0	2.37
	$\rho(d)^{-1} = 0.34$		$\rho(c)^{-1} = 0.41$	



d



c

Table 1: A comparison of Bonacich and Intercentrality in centralised and decentralised networks

Also seen from the example is that the ranking of centralities is preserved for the two cases when θ changes. This is a special case of these two network structures and is not true in general. Later on we will see examples where the ranking is not preserved under an increase in θ.

2.3 The Actors

We now move to a description of the actors.

2.3.1 The Network Planner

There is a set of locations $L \in \mathbb{N}$. The criminal organisation T has access to n potential recruits[7] and has to decide how many of these agents he will use when designing a network structure in each location. The choice of network structures is restricted to the set $g \in \{c, d\}$. The criminal organisation T is endowed with a limited number of edges $E \in \mathbb{N}$ to be distributed across locations.

T designs separate network structures for each location denoted g_l. In the network g_l, each recruit $i = 1, 2, ..., n_l$ produces an effort of $x_i(g_l, \theta_l)$ that depends only on the local network structure and, still to be defined, the network primitive θ_l. This is an individual choice that we can interpret as a desire to take part in criminal offences.

Let $\mathbf{x}(g, \theta) = (x_1(g, \theta), ..., x_i(g, \theta), ..., x_n(g, \theta))$ denote the criminal effort profile of recruits within the network g, so that $x(g, \theta)$ is the sum of the coordinates of $\mathbf{x}(g, \theta)$. T is interested in designing g_l in such a way that the sum of all efforts from all locations is maximised. Let $x_l(g_l, \theta_l)$ denote the total effort exerted in location l. T maximises:

$$x = \sum_{l \in L} x_l(g_l, \theta_l) \tag{1}$$

2.3.2 The Criminal Recruit

Individual recruits decide between entering a given network g_l or entering the labour market and obtaining an outside option of $\omega_l \in [0, 1]$.

Given the network g_l all recruits that decide to enter into the network simultaneously commit to exert effort $x_i(g_l, \theta_l)$. Let $u_i(g_l, \theta_l)$ denote the utility recruit i obtains from being in the network g_l. In order to abstract away from any strategic considerations between an individual recruit and the network planner we implicitly assume that the network planner explicitly commits to providing $u_i(g_l, \theta_l) \geq \omega_l$ if he recruits the individual into the network.

Benefits

The gross benefit $y_i(g, \theta)$ recruit i receives from g is not only a function of x_i but of the entire network profile $\mathbf{x}(g, \theta)$. An individual derives benefits from the actions of all other network participants even if he is not directly connected to them. This is a global relationship captured by δ. The global relationship is augmented with local benefits that an individual derives through his direct connections. These are

[7] We impose no restrictions on n in order to focus the analysis on the edge restriction. The edge restriction can then be interpreted as the wealth endowment of T. Relaxing this assumption so that n is also bounded could be an extension to the present analysis.

captured by λg_{ij}.

$$y_i(\mathbf{x}(g,\theta)) = x_i \max\{1 - \delta \sum_{j=1}^{n} x_j + \lambda \sum_{j=1}^{n} g_{ij} x_j, 0\}$$

To give more structure to the problem we make the following assumption.

Assumption 2. $\lambda \in [0,1]$ *and* $\delta \in (0,1]$ *such that* $\lambda \leq \delta$.

Assumption 2 introduces local complementarities and global substitutability into the gross benefit function. It also serves a more technical purpose. Lemma 1 implies that for the set of all connected graphs $\inf\{\rho(g)\} = 1$, which arises when two nodes are connected by one link. Assumption 2 imposes the requirement that as a minimum a binary cell can exist in equilibrium and that at an equilibrium $\mathbf{x}^* \geq 0$, which allows us to write:

$$y_i(\mathbf{x}(g,\theta)) = x_i - \delta \sum_{j=1}^{n} x_i x_j + \lambda \sum_{j=1}^{n} g_{ij} x_i x_j \geq 0$$

Monitoring and Punishment

A local policing authority l tries to monitor all actions undertaken by the criminal network with the aim of punishing a recruit with fine f if successfully captured. A recruits probability of detection, $p_i(g,\theta)$, depends on the entire criminal profile $\mathbf{x}(g,\theta)$ through the triplet $(p_0, \tilde{\lambda}, \tilde{\delta})$.

p_0 can be thought of as the fraction of policing time l devotes to monitoring the type of crime being carried out by the criminal network. We assume that a policing authority can directly influence the probability of apprehension by increasing the amount of resources, p_0, it devotes to policing a certain type of crime.

We depart from BCZ in several ways. Firstly, instead of stipulating a negative relationship between λ and p_i we follow Heski and assume that the probability of apprehension increases with the degree of complementarity embedded in the network. However the probability of apprehension does not directly depend on the complementarities captured by λ but on the policing authorities estimate of these complementarities $\tilde{\lambda}$.

Secondly, adding to the existing debate on criminal activity and policing, we believe that a high degree of global substitutability within the criminal network should make it harder for l to predict the true perpetrator of the end crime. We model this by assuming that an individual's probability of apprehension decreases with $l's$ estimate of the substitutability, $\tilde{\delta}$, embedded within the network.

$$p_i(\mathbf{x}(g_l,\theta_l)) = p_0 x_i \max\{1 - \tilde{\delta} \sum_{j=1}^{n} x_j + \tilde{\lambda} \sum_{j=1}^{n} g_{ij} x_j, 0\} \quad (2)$$

An analogue to Assumption 2 can be made[8], allowing us to write:

$$p_i(\mathbf{x}(g_l,\theta_l)) = p_0(x_i - \tilde{\delta}\sum_{j=1}^{n} x_i x_j + \tilde{\lambda}\sum_{j=1}^{n} g_{ij}x_i x_j) \geq 0$$

Criminal Utility

In line with Becker (1968) and BCZ (2007) a network recruit has to trade off his gross benefits from being in the network with his expected fine. His expected gains are given by:

$$u_i(\mathbf{x}(g_l,\theta_l)) = y_i(\mathbf{x}(g_l,\theta_l)) - p_i(\mathbf{x}(g_l,\theta_l))f$$

Letting $\pi \equiv p_0 f \in [0,1]$ denote the marginal expected punishment cost of an isolated criminal, by direct substitution we have that:

$$u_i(\mathbf{x}(g_l,\theta_l)) = \max\{(1-\pi)x_i - (\delta - \pi\tilde{\delta})\sum_{j=1}^{n} x_i x_j + (\lambda - \pi\tilde{\lambda})\sum_{j=1}^{n} g_{ij}x_i x_j, \omega\} \quad (3)$$

2.3.3 The Local Policing Authority

Each local policing authority inherits a skill level $s_l \in [0,1]$ that enables it to predict the primitives of the local network g_l. The accuracy of its estimates increase with s but also depends on the true value of the primitives in the following form[9]:

$$\tilde{\lambda} = f(\lambda, s)$$

$$\tilde{\delta} = f(\delta, s)$$

We make the following explicit assumption about the monitoring technology $f(a,s)$ and its relation to skill and model primitives where $a = \{\lambda, \delta\}$.

Assumption 3. *For any $a, s \in [0,1]$, $f(0,s) = 0$, $f(1,s) = 1$, $f(a,0) = 0$, $f(a,s) < a$ and $f(a,1) = a$. Also $f_a > 0$ and $f_{aa} > 0$. Finally, $f_s > 0$ and $f_{ss} < 0$.*

Only the highest skilled policing authority is able to perfectly estimate the model primitives. Policing authorities with lower skills underestimate with varying degrees depending on the true primitives. Larger primitives make the criminal network more visible and make it easier for the policing authority to estimate

[8] Assumption 3 below makes this analogue redundant hence we do not explicitly state it here

[9] Our results will be influenced by the fact that both λ and δ are estimated through the same functional form. More generally we could consider the case where policing authorities use different technologies to estimate different primitives and this may lead to qualitatively different results. This may provide an interesting avenue for future research but for now we believe that this a natural starting point.

these primitives with less bias. Increasing skill levels improves estimates but at a diminishing rate.

What we have in mind is the use of a surveillance technology that allows the policing authority to partially monitor the activities of the criminal network. The policing authority knows that a more technologically evolved surveillance system or a better trained police force will allow it to build a sharper understanding of the network primitives, however if the scale of network interactions are very large, as captured by the primitives, then even a relatively bad policing authority should be able to pick this up.

Always underestimating the true model primitives is tantamount to modeling the policing authority as a bounded rational agent that knows his estimate of the network primitives are downward biased, but are the best achievable given his skill level and therefore does not attempt to adjust them by some arbitrary adjustment function.

In the analysis presented below, for tractability, we will consider a specific functional form for $f(a, s)$ that satisfies Assumption 3. In particular:

$$\tilde{a} = a^{\frac{1}{s}}$$

According to its estimates $(\tilde{\delta}, \tilde{\lambda})$ and its choice of π the policing authority is able to assign a probability of detection to each recruit in the network through (3). Let $\mathbf{p} = (p_1, ..., p_i, ...p_n)$ denote the entire detection profile of the criminal recruits. Given its skill level and choice of π we assume that the policing authority only benefits from $p_m \equiv \max\{\mathbf{p}\}$. Because in expectation this is the most likely target for capture it seems a natural candidate from which a policing authority can derive benefit[10].

Policing is costly and inversely related to policing skill in the following form

$$c(\pi, s) = \frac{\pi^2}{s}$$

The net payoff to the policing authority is given by the following expression:

$$p_m - \frac{\pi^2}{s} \tag{4}$$

2.4 On Global and Local Network Effects

So far we have not said much about the nature of local and global interactions. Only that they enter a criminal recruits gross benefit function. We refrained

[10]Of course one may consider more general utility functions for the policing authority. We will return to this point later in the text.

from doing so this far because in equation 3 we have a very general formulation. Depending on the values $\pi, \tilde{\lambda}, \tilde{\delta}$ take relative to λ, δ the precise nature of the local and global interactions change and we can distinguish four independent cases, of which the fourth case is the primary focus of this study.

1) When $\delta - \pi\tilde{\delta} < 0$ and $\lambda - \pi\tilde{\lambda} > 0$ the model exhibits both **local and global complementarities**. In this case it is clear that the effort levels of each criminal recruit within the network are self and cross reinforcing and there is no uniquely stable equilibrium in the system.

2) When $\delta - \pi\tilde{\delta} > 0$ and $\lambda - \pi\tilde{\lambda} < 0$ the model exhibits both **local and global substitutability**. In this framework effort levels of direct and odd path length neighbours act as direct substitutes for the effort levels of recruit i and reinforce the global substitutability of effort inherent in the model. However the effort levels of indirect neighbours of even path length work against global substitutability and act as complements for the effort level of recruit i. A uniquely stable equilibrium for this system exists.

3) When $\delta - \pi\tilde{\delta} < 0$ and $\lambda - \pi\tilde{\lambda} > 0$ the model exhibits **local substitutability and global complementarity**. As long as $|\delta - \pi\tilde{\delta}| \leq |\lambda - \pi\tilde{\lambda}|$ local substitutability will outweigh the global complementarity inherent in the model and a uniquely stable equilibrium will exist.

While each of the above frameworks are worthy of study in their own right, here we wish to focus mainly on the case of **global substitutability and local complementarity**. This arises

4) When $\delta - \pi\tilde{\delta} > 0$ and $\lambda - \pi\tilde{\lambda} > 0$. A sufficient condition for the existence of this particular case is $\pi \leq \min\{\frac{\lambda}{\tilde{\lambda}}, \frac{\delta}{\tilde{\delta}}\}$[11]. However an explicit reason for introducing assumption 3 is to ensure that we are in world of complementarities at the local level and substitutability in effort levels at the global level for any $\pi \in [0, 1]$.

3 Crime and Punishment: A One Shot Game

3.1 The One Shot Game

We want to model a one shot game between T and only one local policing authority l in which the outside option ω, the policing authority's skill level s and its policing technology $f(a, s)$ are common knowledge. The policing authority inherits s and

[11]Hence a model that allows for the policing authority to overshoot one of its estimates, say $\tilde{\lambda} > \lambda$ will have to restrict attention to the range of values $0 \leq \pi < \frac{\lambda}{\tilde{\lambda}} < 1$.

cannot upgrade to $s' > s$. The network planner faces an edge restriction E and decides (λ, δ) known only to himself and the criminal recruits. The network planner chooses n and $g \in \{c, d\}$ and the policing authority simultaneously chooses the expected punishment level π.

We first study the equilibrium characteristics of this one shot game and then offer some comparative statics.

3.2 The Equilibrium

An equilibrium of the one shot game is defined as a situation in which:
(1) an individual criminal is earning as much utility from criminal activity as he would from his outside option
(2) The network structure maximises aggregate effort levels and the design is stable i.e. given the network structure, existing network participants do not want to quit and no potential recruit wants to join.
(3) Given his skill level the policing authority chooses an optimal π to maximise equation 4.

3.2.1 Individual Participation

An individual criminal recruit takes the network structure g and π as given. He is aware of the network primitives (λ, δ) and of the policing authority's estimate of these primitives $(\tilde{\lambda}, \tilde{\delta})$. Denote $\theta = \frac{\lambda - \pi\tilde{\lambda}}{\delta - \pi\tilde{\delta}}$ and let $\rho(g)$ be the spectral radius of g, then the following result holds.

Proposition 1. *If $\theta\rho(g) < 1$ then there exists a Nash equilibrium in which the effort level of recruit i is uniquely determined. In this equilibrium recruit i exerts effort in proportion to his Bonacich Centrality in the network g.*

$$x_i^*(g, \theta) = \frac{1 - \pi}{\delta - \pi\tilde{\delta}} \frac{b_i(g, \theta)}{1 + b(g, \theta)}$$

Once part of the network an individual criminal produces an effort level in direct proportion to his Bonacich centrality relative to the sum of Bonacich. Thus a criminal that is relatively central will also exert high effort. Increasing policing resources devoted to targeting the crime does not necessarily lead to a direct reduction in effort levels. There is a direct reduction effect through $1 - \pi$ but by increasing π a policing authority also reduces $\delta - \pi\tilde{\delta}$ hence the direct reduction in crime is less than $1 - \pi$.

In addition increasing π can lead to an increase in θ which will increase effort levels for any given network structure, through higher b_i.

Now that we have shown how the decision of a recruit to exert effort is shaped by his positioning in a network we inspect under what conditions the recruit is willing to join the network.

Proposition 2. *If $\theta\rho(g) < 1$ in any Nash equilibrium a criminal recruit i participates in the criminal network if and only if:*

$$\frac{b_i(g,\theta)}{1+b(g,\theta)} \geq \frac{\sqrt{\omega(\delta-\pi\tilde{\delta})}}{1-\pi}$$

A network participant derives utility not only from the effort he exerts but also from the global effort profile. In equilibrium we see that a recruits utility is precisely a function of his Bonacich centrality relative to total Bonacich i.e. the more central a recruit is the higher is his utility. For a recruit to stay in the network it must be the case that the utility he receives from being in the network is at least as great as his outside option. His equilibrium outside option relative to network utility is determined not only by ω but also by (π, s, δ).

3.2.2 Network Design and Stability

First we state a general result relating to network stability. We have seen that utility within a network is heterogenously determined as function of network positioning. The stability condition requires that even the recruit with the lowest network utility, prefers to stay in the network. This stability is supported in equilibrium as long as another agent cannot enter the network and derive utility greater than his outside option by doing so.

Proposition 3. *The network g is stable if and only if:*

$$\max_{j\in N\backslash g} \frac{b_j(g\cup\{j\},\theta)}{1+b(g\cup\{j\},\theta)} \leq \frac{\sqrt{\omega(\delta-\pi\tilde{\delta})}}{1-\pi} \leq \min_{i\in g} \frac{b_i(g,\theta)}{1+b(g,\theta)}$$

Proof. Use Proposition 8 from Ballester (2008). □

With the above stability condition in mind the network planner T has a maximum number of edges E which he can use to link at most $n = E - 1$ recruits. Taking θ as given T designs a network $g \in \{c, d\}$ in such a way that equation 1 is maximised.

To fix ideas consider a regular (decentralised) network made up of n nodes. In order to fully connect the n nodes we require $\frac{n(n-1)}{2}$ edges. Equivalently we can link up to $\frac{n(n-1)}{2} + 1$ nodes in a star (centralised) network using $e \leq \frac{n(n-1)}{2}$ edges. Now imagine we have access to an extra n edges. This allows us to move from

$d(n)$ to $d(n+1)$. However, if we only have access to an extra $n-1$ edges, $r(n+1)$ is no longer feasible, but we can still design a star with up to $\frac{n(n-1)}{2}+(n-1)+1=\frac{n(n+1)}{2}$ nodes. We are interested in knowing which of these configurations leads to highest aggregate effort levels and whether our preferred network meets the stability criteria.

The following proposition quantifies aggregate effort levels in two graph types as a function of n and θ.

Lemma 2. *Fix n and let $d(n)$ denote a decentralised (regular) network connecting n nodes. For any n if and only if $\theta\rho(d(n))<1$ then in equilibrium each recruit in d exerts effort level*

$$x^*(d,\theta)=\frac{1-\pi}{\hat{\delta}(n+1-(n-1)\theta)}$$

and has a Bonacich centrality of

$$b^*(d,\theta)=\frac{1}{1-(n-1)\theta}$$

Lemma 3. *Fix n and let $c(n)$ denote a centralised (star) network connecting n nodes. For any n if and only if $\theta\rho(c(n))<1$ then in equilibrium*
(a) The hub of c, denoted h, exerts an effort level

$$x_h(c,\theta)=\frac{(1-\pi)(1+(n-1)\theta)}{\hat{\delta}((n+1)+2(n-1)\theta-(n-1)\theta^2)}$$

and has a Bonacich centrality of

$$b_h(c,\theta)=\frac{1+(n-1)\theta}{1-(n-1)\theta^2}$$

(b) Each periphery node of c, denoted p, exerts an effort level

$$x_p(c,\theta)=\frac{(1-\pi)(1+\theta)}{\hat{\delta}((n+1)+2(n-1)\theta-(n-1)\theta^2)}$$

and has a Bonacich centrality of

$$b_p(c,\theta)=\frac{1+\theta}{1-(n-1)\theta^2}$$

Proposition 4. *Fix n and n' such that $n\leq n'$. Let $\{c(n),d(n)\}$ denote, respectively, a centralised (star) and decentralised (regular) network connecting n nodes.*

For any $n, n' > 2$ the following hold:

(1) For any $\theta\rho(d(n)) < 1$ we have $x(c(n,\theta)) < x(d(n,\theta))$.
(2) Let g denote a particular type of network from $\{c(n), d(n)\}$ with n nodes. And let g' from $\{c(n'), d(n')\}$ denote the same type of network as g, but with $n' > n$ nodes. For any $\theta\rho(g') < 1$ we have $x(g(n,\theta)) < x(g(n',\theta))$.

For $n > 4$ and $n \leq n' \leq \frac{n(n+1)}{2}$

(3) $x(c(n',\theta)) \geq x(d(n,\theta))$ if $\theta \leq \theta^*$ and $x(c(n',\theta)) < x(d(n,\theta))$ if $\theta > \theta^*$. Where

$$0 \leq \theta^* \leq \frac{1}{n-1}$$

and

$$\theta^* = \frac{n'-n}{(n'-1)(n-2)}$$

In its construction Proposition 4 makes direct use of an interesting fact pertaining to Bonacich Centrality first highlighted in Figure 2. Bonacich Centrality is an infinite sum counting the total number of paths in the network that originate from node i. The magnitude of this sum depends on the value of θ relative to its natural bound $\frac{1}{\rho(g)}$. By designing a network that minimises $|\theta - \frac{1}{\rho(g)}|$ a network planner can induce a higher total Bonacich Centrality b than from any other configuration. If this is not feasible and $|\theta - \frac{1}{\rho(g)}|$ is large enough b is more likely to be maximised by concentrating all available links on one node.

This observation leads us to prescribe the following algorithm the network planner should follow when designing a network to maximise criminal activity.

Proposition 5. *Given θ and the edge restriction E find the largest n such that $r(n,\theta)$ is feasible and calculate $\frac{1}{\theta}$.*
(1) If $\frac{1}{\theta} =$ integer and $n \geq \frac{1}{\theta}$ set $n^* = \frac{1}{\theta}$. Design $r(n^*,\theta)$ if stable.
(2) If $\frac{1}{\theta} \neq$ integer and $n \geq \frac{1+\theta}{\theta}$ set $n^* = \frac{1+\theta}{\theta}$. Design $r(n^*,\theta)$ if stable and $s(n',\theta)$ is either not feasible, or feasible but not stable. If $s(n',\theta)$ is both feasible and stable, then find largest $n'^* > n'$ such that $s(n'^*,\theta)$ is both feasible and stable and design. Where:

$$n' = \frac{(n^*-2)\theta + n^*}{(n^*-2)\theta + 1}$$

(3) If (1) or (2) are not satisfied find highest $n'' \leq n$ such that $r(n'',\theta)$ is stable. Design $r(n^*,\theta)$ if stable and $s(n',\theta)$ is either not feasible, or feasible but not stable. If $s(n',\theta)$ is both feasible and stable, then find largest $n'^* > n'$ such that $s(n'^*,\theta)$ is both feasible and stable and design. Where:

$$n' = \frac{(n''-2)\theta + n''}{(n''-2)\theta + 1}$$

3.2.3 Network Targeting

The policing authority inherits a skill level s and decides on the optimal expected punishment π for a successful capture. Given its skill level it assigns a probability profile to each criminal recruit that is increasing π. The optimal π so as to maximise the chances the capturing the recruit with the highest probability profile. The next proposition characterises this profile.

Proposition 6. *If $\theta\rho(g) < 1$, in equilibrium a criminal recruit i, that participates in the criminal network, faces a probability of detection that is a weighted average between his Bonacich and Intercentrality. In particular:*

$$p_i = \frac{\pi(1-\pi)}{(\lambda - \pi\tilde{\lambda})(\delta - \pi\tilde{\delta})} \frac{(\lambda - \tilde{\lambda})b_i + (1-\pi)\tilde{\lambda}(r_i m_{ii})}{(1+b)^2} \tag{5}$$

Depending on the policing authorities skill level s and the true local complementary parameter λ, criminal $i's$ probability of detection falls in the following range:

$$p_i \in \left[\frac{\pi(1-\pi)}{\delta} \frac{b_i}{(1+b)^2}, \frac{\pi}{\delta} \frac{r_i m_{ii}}{(1+b)^2} \right]$$

We characterise the ability of a policing authority to assign detection probabilities based on how well it is able to ascertain the network primitives, in particular λ. Ballester et al. (2006) suggest an optimal policy for targeting the Intercentral player from a network. In an ideal world where the policing authority can perfectly visualise the network it should do so.

However in a world where policing is subject to the imperfections of monitoring technologies and policing skills, probabilities of detection will be a function both of Intercentrality and Bonacich Centrality. High skilled policing authorities will place more weight on the Intercentrality of a criminal relative to total network output, whereas relatively low skilled policing authorities will put more weight on how much a criminal produces relative to others in the network when assigning a probability profile.

For the particular network structures we are considering the ranking of nodes induced by the two measures of centrality coincide. Therefore, both a high and low skilled policing authority will assign the largest detection probability to the hub when facing a centralised network and equal probabilities on criminals when facing decentralised activity. By definition, however, the Intercentrality of a node is always higher than the Bonacich Centrality hence when facing the same network with the same aggregate Bonacich, the higher skilled policing authority will assign a higher detection probability to each criminal in the network.

Proposition 6 allows us to characterise p_m which depends not only on policing skill but also on the expected punishment π. Using this fact we state the following result that determines the optimal π.

Proposition 7. *The optimal π is the solution to fourth-order polynomial:*

$$\alpha\pi^4 + \beta\pi^3 + \eta\pi^2 + \kappa\pi + \tau = 0$$

Where $\alpha, \beta, \eta, \kappa$ are all functions of $(b_m, r_m, s, \lambda, \delta)$ and τ is a function of (b_m, λ, δ).

The characteristic equation in Proposition 7 takes a polynomial form indicating the possibility for multiple equilibria in π. In any of these equilibria π^* will be a function of policing skill, network primitives and more specifically of the network structure. Considering more general objective functions for the policing authority, that do not explicitly rely on probability profiles can potentially alleviate problems of multiplicity of equilibria[12]

Another Numerical Example

Bonacich and Intercentrality always coincide in the simplified class of networks we are considering. In particular cases the two centrality measures can lead to different rankings of node influence. When this happens it is not necessarily true that the node with a higher Intercentrality obtains a higher probability of detection even if policed by a high skilled authority. We proceed to quantify the message of Proposition 6 by way of examples.

We present three networks of varying sizes and centralisation, even though the largest eigenvalue for each graph are very similiar. For each graph we also rank the nodes and report centralisation according degree and betweeness centrality. For all three examples we assume $\theta_o = \{0.1, 0.2\}$, defined as $\theta_o = \frac{\lambda}{\delta}$.

For the special cases of $s = 0$ and $s = 1$ we have that $\theta = \theta_o$. This allows us to focus on how good and bad policing influence probability profiles, holding the output of the network constant.

We also use different values of λ, δ to show how relative probability profiles can differ for the same θ_o.

[12]We note that $\lambda - \pi\tilde{\lambda}$ can be considered an efficiency loss in policing local criminal interactions. Given its policing skill s a policing authority can reduce this efficiency loss by increasing π. Imagine a situation where a policing authority cares about minimising the product of efficiency losses both at the local and global level. Its objective function takes the form:

$$-(\lambda - \pi\tilde{\lambda})(\delta - \pi\tilde{\delta}) - \frac{\pi^2}{s}$$

The F.O.C of this maximisation problem is linear in π and yields the following interior solution for $s > 0$:

$$\pi = \frac{\tilde{\lambda}\delta + \lambda\tilde{\delta}}{2(s + \tilde{\lambda}\tilde{\delta})}$$

For the specific case of $\tilde{a} = a^{\frac{1}{s}}$ we have that optimal π is increasing with skill for small enough (λ, δ) but eventually starts to decline for large (λ, δ).

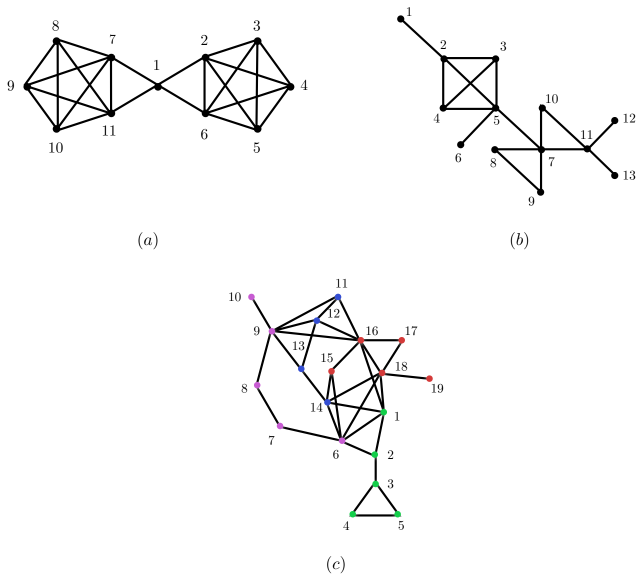


Figure 3: (a) Changing Intercentrality example from Ballester et al. (2006), (b) Changing Bonacich Centrality, (c)Krebs 9/11 network

The first example is borrowed from Ballester et al. (2006). The purpose of this example is to show what happens in a situation in which Intercentrality node ranking changes depending on the magnitude of θ, even though Bonacich Centrality node ranking is preserved. For $\theta = 0.1$ Bonacich and Intercentrality node ranking are equivalent. For $\theta = 0.2$ Intercentrality node ranking changes even though Bonacich remains the same.

The next two networks are examples of graphs in which the Bonacich ranking changes with θ but Bonacich and Intercentrality rankings are equivalent for all values of θ. Hence we present the rankings for $\theta = 0.1$ & $\theta = 0.2$ but only present probability profiles for $\theta = 0.1$.

The second example is constructed to show that Bonacich Centrality can also be sensitive to the magnitude of θ. In this particular example Intercentrality moves in line with Bonacich.

Finally, the third example maps the 9/11 terrorist network, as identified by Krebs(2002) and shows how policing authorities would have placed different nodes on their targeting list as a function of their skill level.

4 Comparative Statics

The main three ingredients that make up the equilibrium are Proposition 1, Proposition 3 and Proposition 7. We can find the optimal π through Proposition 7, which in our setup builds a relationship between expected punishment and network structure. Proposition 1 shows that equilibrium crime levels are a function of network structure and the θ with which it operates. In equilibrium even the network structure is a function of θ as shown in Proposition 5. Finally Proposition 3 gives us a criteria for evaluating the stability of a network.

Taken from Proposition 3 let

$$\omega_\gamma = \frac{\sqrt{\omega(\delta - \pi\tilde{\delta})}}{1-\pi}$$

denote the resulting outside option adjusting for policing level and skill.

ω_γ provides a restriction on the graph structures that can arise in equilibrium and depends not only on the exogenous outside option and policing skill and resources but also on the degree of global substitutability in the network. This means that through his choices on (λ, δ) the network planner can also influence the adjusted outside option individual criminals face.

In its formulation θ also depends on policing skill and expected punishment along with model primitives. This is also a parameter that can be influenced by both the choices of the network planner and by the choices of the policing authority.

It is clear that ω_γ is directly proportional to ω. The relationship between (ω_γ, θ) and other model parameters is made explicit with the following results.

Global Substitutability

The network planner can directly influence equilibrium outcomes through his choices of (λ, δ). His choice of the degree of global substitutability[13] will influence both ω_γ and θ in a similiar manner which can be stated in the following lemma.

Lemma 4. *Let $z = \{\omega_\gamma, \theta\}$. For a given configuration of $(\omega, \pi, \lambda, s)$ and $\delta' > \delta$:*
(a) If $\delta' \leq \delta_o$ then $z(\delta') \geq z(\delta)$.
(b) If $\delta \geq \delta_o$ then $z(\delta') \leq z(\delta)$.
(c) If $\delta \leq \delta_o \leq \delta'$ then sign$(z(\delta') - z(\delta))$ depends on $|\delta' - \delta|$. Where

$$\delta_o = \left(\frac{1}{s\pi}\right)^{\frac{1}{s-1}}$$

When substitutability is low enough a network planner can induce a higher θ by increasing δ slightly. But he faces a cost of doing so as he also raises the adjusted outside option in the process. When substitutability is high enough raising

[13]Imagine the Network planner along with designing a network also agrees upon a "work-plan" for the criminals, the nature of actions chosen for them can induce different degrees of substitutability in the network

substitutability can have the opposite effect. When deciding between very high and very low substitutability the network planner needs to look at the magnitude of the difference between the two to ascertain how his decisions will influence (ω_γ and θ.

Local Complementarities

The network planner can only influence θ through his choices of λ. The relationship between λ and θ is analogous to that presented in Lemma 2 and follows a similiar intuition.

Lemma 5. *For a given configuration of (π, δ, s) and $\lambda' > \lambda$:*
(a) If $\lambda' \leq \lambda_o$ then $\theta(\lambda') \geq \theta(\lambda)$.
(b) If $\lambda \geq \lambda_o$ then $\theta(\lambda') \leq \theta(\lambda)$.
(c) If $\lambda \leq \lambda_o \leq \lambda'$ then $sign(\theta(\lambda') - \theta(\lambda))$ depends on $|\lambda' - \lambda|$. Where

$$\lambda_o = \left(\frac{1}{s\pi}\right)^{\frac{1}{s-1}}$$

Expected Punishment

The choice of expected punishment will influence both the adjusted outside option ω_γ and θ. The following two results relate the relevant comparative statics.

Lemma 6. *For a given configuration of (ω, δ, s) if $\pi' > \pi$ and $\pi > \frac{\omega\delta - \tilde{\delta}}{(1+\omega)\tilde{\delta}}$ then $\omega_\gamma(\pi') > \omega_\gamma(\pi)$.*

Given an outside option ω and policing skill s, if expected punishment is high enough then a further increase in π will lead to a more restrictive adjusted outside option. For low π an increase in expected punishment can actually reduce the outside option. Note that $\lim_{s\to 0} \frac{\omega\delta - \tilde{\delta}}{(1+\omega)\tilde{\delta}} =$

The relationship between π and θ is more straightforward. As the following result shows increasing π always raises θ.

Lemma 7. *For any (λ, δ) and $s \in (0, 1)$ if $\pi' > \pi$ then $\theta(\pi') > \theta(\pi)$.*

The result holds specifically for[14] $f(a, s) = a^{\frac{1}{s}}$ and may seem slightly counter-intuitive at first but this is not so. Given that $\lambda \leq \delta$, by definition, for any skill level $s \in [0, 1]$ a policing authority will always form a closer estimate of δ to its true value compared to λ. Increasing π favours this lopsidedness in policing ability.

[14]We are trying to formulate a proof for the general class of functions that satisfy Assumption 3.

The criminal network uses local links more intensely to compensate for the fact that the policing authority has a much better understanding of global patterns relative to local ones.

Policing Skill

The relationship between policing skill and ω_γ is immediate. Higher skill leads to a better estimate of δ which increases the outside option. This is shown formally in the following result.

Lemma 8. *For a given configuration of* (ω, π, δ) *if* $s' > s$ *then* $\omega_\gamma(s') > \omega_\gamma(s)$.

The relationship between s and θ is slightly more complicated. For a predetermined choice of (λ, δ) the network planner T would like to operate his network using $\theta_o = \frac{\lambda}{\delta}$ but, in fact, is induced by the policing authority to operate using $\theta = \frac{\lambda - \pi\tilde{\lambda}}{\delta - \pi\tilde{\delta}}$. The next lemma shows that a good policing authority will induce the network to operate at $\theta \geq \theta_o$.

Lemma 9. $\theta(s) > \theta_o$ *for any* $s \in [0, 1]$ *and* $\theta(s)$ *is globally concave in* s.

Both $s = 1$ and $s = 0$ induce a $\theta = \theta_o$, but for opposing reasons made explicit in the proof of the lemma. Although a good policing authority cannot influence θ it compensates by assigning probability profiles based completely on node Intercentrality. Whereas the bad policing authority not only cannot influence θ but also assigns a probability profile based on Bonacich centrality and hence assigns a lower probability to each recruit.

For intermediate values of skill, relatively good policing authorities can induce a $\theta > \theta_o$ inducing a network to make more use of its local links relative to global interactions. But in order to capitalise on the higher induced θ a network planner will use Proposition 5 to design a smaller decentralised network and in fact will be forced to do so because of Lemma 6.

4.1 Model Inference

Because of the particular form our policing authority's objective function takes we are limited in our ability to pin down the precise form of the equilibrium in the one shot game of crime and punishment. However given our equilibrium characterisations and comparative statics results we can make the following inferences from the model.

Consider two policing skills $s_1 > s_2$ and let $\pi, \pi' \in [0, 1]$ be two values to be chosen such that they satisfy:

$$\left(\frac{1}{s_1\pi}\right)^{\frac{1}{s_1-1}} < \left(\frac{1}{s_2\pi}\right)^{\frac{1}{s_2-1}}$$

Let $\lambda_2 > \lambda_1$ and $\delta_2 > \delta_1$ and define:

$$\theta_1 = \frac{\lambda_1 - \pi\tilde{\lambda}_1}{\delta_1 - \pi\tilde{\delta}_1} \; \theta_2 = \frac{\lambda_2 - \pi\tilde{\lambda}_1}{\delta_1 - \pi\tilde{\delta}_1}$$
$$\theta_3 = \frac{\lambda_1 - \pi\tilde{\lambda}_1}{\delta_2 - \pi\tilde{\delta}_2} \; \theta_4 = \frac{\lambda_2 - \pi\tilde{\lambda}_1}{\delta_2 - \pi\tilde{\delta}_2}$$

Because we want to concentrate on the role of complementarities and wish to keep the exposition simple we will concentrate on the case where:

$$\left(\frac{1}{s_2\pi}\right)^{\frac{1}{s_2-1}} < \delta_1$$

Lemma 4 gives us the partial ordering $\theta_4 < \theta_2$ and $\theta_3 < \theta_1$. For local interactions we assume that:

$$\lambda_1 < \left(\frac{1}{s_1\pi}\right)^{\frac{1}{s_1-1}} < \lambda_2 < \left(\frac{1}{s_2\pi}\right)^{\frac{1}{s_2-1}}$$

Such that, in accordance with Lemma 5 $\theta_1 > \theta_2$ and $\theta_3 > \theta_4$ when $s = 1$ and $\theta_1 < \theta_2$ and $\theta_3 < \theta_4$ when $s = 2$.

The above structure allows us to make our first inference

Inference 1. *Given the structure given above*

(1) When facing a ***low skilled authority*** *the network planner T obtains a ranking:*
(a) $\theta_3 < \theta_1 < \theta_4 < \theta_2$ *if* $\frac{\hat{\lambda}_2}{\hat{\lambda}_1} > \frac{\hat{\delta}_2}{\hat{\delta}_1}$
(b) $\theta_3 < \theta_4 < \theta_1 < \theta_2$ *if* $\frac{\hat{\lambda}_2}{\hat{\lambda}_1} < \frac{\hat{\delta}_2}{\hat{\delta}_1}$

(2) When facing a ***high skilled authority*** *the network planner T obtains a ranking:*
(a) $\theta_4 < \theta_3 < \theta_2 < \theta_1$ *if* $\frac{\hat{\lambda}_2}{\hat{\lambda}_1} > \frac{\hat{\delta}_1}{\hat{\delta}_2}$
(b) $\theta_4 < \theta_3 < \theta_2 < \theta_1$ *if* $\frac{\hat{\lambda}_2}{\hat{\lambda}_1} < \frac{\hat{\delta}_1}{\hat{\delta}_2}$

One can think of a higher λ as reflecting activity based on exchange of higher information content. Whereas a higher δ reflects greater rivalry in actions which leads to less activity in equilibrium and directly lessens the probability of detection of individual recruits. Hence in the terminology of Baker and Faulkner (1993) a higher δ reflects less need to conceal the activities of the criminal network.

Inference 1 shows that the kind of network structures that arise in equilibrium are not just functions of this trade off between sharing information and concealment but also of policing skill and resources.

For example, take the high skilled authority. A criminal network that wishes to coordinate actions that require high information sharing but are not easily concealed will have to operate with θ_1, whereas it can operate with $\theta_4 < \theta_1$ if it wishes to coordinate actions that require low information sharing and are easily concealed. Using Proposition 5 and Lemma 8 we can stipulate ω_γ will be high for a given ω and that θ_1 will induce a small and decentralised network. On the other hand θ_4 is better suited for a larger decentralised network, but this may not be stable or feasible given the edge restrictions.

However when facing a low skilled authority and a low ω the network planner could potentially design a large centralised network for both cases. In particular because of the ranking we present above, these two networks are likely to operate at a similiar size, even without edge restrictions.

Inference 2. *For a given ω and the edge restriction E higher skill s limits criminal recruitment n and is more likely to induce decentralised criminal activity. The Policing authority will place a greater weight on a criminal's Intercentrality relative to his Bonacich Centrality when assigning a detection profile p_i. Given s large centralised networks arise when a combination of π and/or ω very low happens.*

5 Coordinated Policing: A One Shot Game

We now focus attention towards the issue of coordinated policing and present a very simple framework that captures the tension policing authorities face when deciding to coordinate. By coordinating policing local authorities gain access to a superior technology but incur a cost of doing so. Like standard coordination games in the literature the coordination decision of policing authorities exhibits strategic complementarities and can lead to multiple coordination equilibria. The network planner T is aware of this and we will show that he can influence the equilibria that arise through his choice of local network structures, in particular under certain conditions he can induce at least one local authority not to coordinate inducing all other policing authorities not to coordinate as well.

In order to highlight the coordination issue we study the equilibria of another one-shot game proposed below. The model takes on some of the features of the one-shot crime and punishment game, however we keep these as simple as possible to focus on coordination.

5.1 The Model

Let $\mathcal{L}$ denote the set of all locations. There are $l \geq 2$ locations such that the cardinality of $\mathcal{L}$ lies between 2 and some unspecified upper bound denoted by L.

Each location has its own policing authority endowed with skill $s_l \in [0,1]$ and its own criminal network g_l. Policing authorities either act alone in targeting local criminal networks or can coordinate their efforts. Coordinating incurs a cost c_l. We assume that coordination costs are decreasing in s_l.

Each local network g_l is set up by a central planner. The central planner faces a set of exogenous variables (π_l, ω_l, s_l) that determine the maximum number of criminals x_l he can recruit. The central planner T also faces a global edge restraint E that needs to be distributed across the different locations.

In order to avoid confusion with the previous model we first explicitly state the payoff functions for the game players and assumptions made.

Individual Recruits

We are not interested in explicitly considering the Individual recruit problem here. We will consider a simplified utility form for the recruit that will allow us to concentrate on the choices of T. The utility of the individual recruit is still given by the general formulation of equation 3, but with the following caveats:

Assumption 4. *We normalise $f = 1$ and assume that the probability of detection is only a function of police resources π and individual efforts x_i, so that $p_i = \pi x_i$.*

This assumption allows us to write the utility of recruit i as:

$$u_i(\mathbf{x}(g_l, \theta_l)) = \max\{(1-\pi)x_i - \delta \sum_{j=1}^{n} x_i x_j + \lambda \sum_{j=1}^{n} g_{ij} x_i x_j, \omega\}$$

We also normalise $\omega_l = 0$ for all $l \in \mathcal{L}$. In conjunction with Proposition 3 this assumption ensures that all potential recruits will be part of one of the local networks. Using Proposition 2 we can write the equilibrium output of recruit i as:

$$x_i^*(g_l, \theta_l)) = \frac{(1-\pi)b_i}{\delta(1+b)} \geq 0 \tag{6}$$

for $\pi \leq 1$ and $\theta = \frac{\lambda}{\delta}$.

The Network Planner

The network planner T can choose any network size from $g \in \{c, d\}$ as described in previous sections. The network planner connects, with a direct link, to the locally Intercentral player in each g_l. For a centralised network this is always the hub. For a decentralised network T randomly picks one recruit to connect with.

In each local network we denote l_t the locally Intercentral player with a link to T. If all policing authorities coordinate l_t is captured with certainty. In order to avoid strategic considerations between T and each l_t we make the following assumption.

Assumption 5. *If all l_t are captured then T is also captured. If this happens T suffers a loss of $-\infty$.*

This assumption guarantees that T will not willingly commit to designing local network structures that induce the local policing authorities to coordinate.

Let κ_l be exogenously given and denote the marginal benefit T receives from operating in locality l.

We partition the set $\mathcal{L}$ into $C \cup D$. Where C denotes the set of locations in which T designs a centralised network, and D denotes the set of locations that operate decentralised networks.

The expected payoff to the central planner can be written as

$$\sum_{l \in C} (1-\tau)\kappa_l \left(\left(1 - \frac{\pi(1-\pi)b_t}{\delta(1+b_c)} \right) b_l(n-1,\theta) \right) +$$
$$\sum_{l \in D} \left((1-\tau)\kappa_l \left(1 - \frac{\pi(1-\pi)b_t}{\delta(1+b_d)} \right) b_l(n-1,\theta) + \tau\kappa_l(b_d(n-1,\theta) \right)$$

The Policing Authorities

Policing authorities can act alone or coordinate resources. Coordinating incurs a cost that is increasing in police resources π_l but decreasing in policing skill s_l, and is given by:

$$c_l = \frac{c\pi_l}{s_l}$$

Let $\tau_l \in [0,1]$ denote the time l spends coordinating with other policing authorities. The trade-offs of coordination are captured in the following minimum effort game[15] where γ_l denotes the payoff to l:

$$\gamma_l = (1-\tau_l)(a_g) + z_g \min(\tau_l, \tau_{-l}) - \tau_l \frac{c\pi_l}{s_l}$$

Acting alone i.e. $\tau_l = 0$ a policing authority gets a benefit of a_g. As before we assume that this benefit is $\max(\mathbf{p})$, by direct substitution we have that:

$$a_g \equiv p_m = \frac{\pi(1-\pi)b_m}{\delta(1+b)}$$

Definition 1. *Let $\underline{\tau} = min(\tau_{-i})$ and $\tau = min(\tau_i, \underline{\tau})$*

[15] Van Huyck et al. point out how the minimum effort game is a generalisation of the stag hunt game. This analogy seems quite apt for the problem we are studying.

By coordinating and conditional on $\tau > 0$ the policing authority can target the locally Intercentral recruit with certainty. In the case of a locally centralised network coordination allows l to target the hub, leaving all other nodes disconnected, therefore $z_c = b_c(n, \theta)$.

When facing a decentralised network the targeting of the Intercentral node still leaves a smaller decentralised network operational, but the policing authority captures not only the effort level of the Intercentral recruit but also his influence on the effort levels of all other recruits, therefore $z_d = b_d(n, \theta) - b_d(n-1, \theta)$.

When the benefits from coordinating outweigh the costs a local policing authority has a clear incentive to cooperate. We make the following assumption that does not allow the policing authority to coordinate in certain cases.

Assumption 6. *When the benefits of coordinating z_g are exactly offset by the costs c_l will always choose not to coordinate and set $\tau_l = 0$.*

5.2 The Coordination Game

Each local policing authority l has one opportunity to target and disrupt the output of their local network g_l. The magnitude of disruption depends on whether local policing authorities decide to cooperate with one another or not. The local policing authorities simultaneously decide to coordinate or not and T simultaneously decides on the network structure in each location and the θ with which to operate.

We impose a certain level of symmetry by requiring (κ, π, δ) to be the same in all locations and known to all players. Policing skill $s_l \in [0, 1]$ varies by location but is common knowledge to all players.

Let n_l be the total number of recruits available to T in location l. We will consider the case in which T has $E_l \leq L(n_l - 1)$ number of edges available for each location, which are not globally transferable. Again n_l is common knowledge for all players.

The general game considered with the above payoffs and structure can be described the couplet $\Gamma = (\mathcal{G}, \mathcal{L})$.

5.3 Equilibrium

Coordination

We allow for differences in c_l stemming from differences in s_l in each locality. This leads us to consider the following cost structures and the games defined over these specific cost structures.

High Cost. $C_h = (\max\{c_i, i \in I\} > b_s(\delta_m)) \in \mathcal{C}$ *and* $\Gamma_h = (C_h, g, \mathcal{I})$

Low Cost. $C_l = (\max\{c_i, i \in I\} < b_s(\delta_0)) \in \mathcal{C}$ *and* $\Gamma_h = (C_l, g, \mathcal{I})$

Intermediate Cost. $C_s = (b_s(\delta_0) < \max\{c_i, i \in I\} < b_s(\delta_m)) \in \mathcal{C}$ *and* $\Gamma_h = (C_s, g, \mathcal{I})$

We relate a proposition for each cost game. However before doing so we state the following lemma which will be used in proving the subsequent propositions.

Lemma 10. *(a) If l_i faces a regular network and $c_i \geq 0$ then the best response for l_i is to set $e_i = 0$.*
(b) If $c_i > b^s_i(\delta)$ then the best response for local policing authority l_i is to set $e_i = 0$.
(c) If $c_i \leq b^s_i(\delta)$ then the best response requires that l_i set $e_i = \underline{e}$.
(d) If there exists k such that $e_k = 0$ then all other agents set $e_i = 0$.

Proof. For ease of exposition we will ignore subscripts on the quadruplet (π, a, b, c) for the time being. We recall that the payoff from coordinating is $\pi = (1 - e_i)(a - b) + a\min(e_i, \underline{e}) - ce_i$.
a) When $c_i =$ this follows from the assumption made on coordination. When $c_i > 0$ the policing authority incurs cost $c_i > 0$ of coordinating but still receives the same benefits $a^r_i - b^r_i$. b) Start by taking $\underline{e}$ as given. Setting $e_i > \underline{e} > 0$ yields $\pi = (1 - e_i)(a - b) + a\underline{e} - ce_i$. Expansion gives us $\pi = (a - b) + a(\underline{e} - e_i) - e_i(b + c)$. The second term on the LHS shows that l_i can do strictly better by setting $e_i = \underline{e}$. Lets assume $e_i = \underline{e} > 0$, the expression simplifies to $\pi = (a - b) + (b - c)\underline{e}$. Since $c > b$ the second term subtracts from π and lowering $e_i < \underline{e}$ improves π. When $e_i < \underline{e}$, we have $\min(e_i, \underline{e}) = e_i$ and l_i payoff adjusts to $\pi = (a - b) + (b - c)e_i$. Again it is straightforward to see that as long as $e_i > 0$ π is decreasing in e_i. Therefore, the maximum is achieved by setting $e_i = 0$ to obtain $\pi = a - b$.
c) Starting with $\underline{e}$ as given and setting $e_i > \underline{e} > 0$ obtains $\pi = (a - b) + a(\underline{e} - e_i) - e_i(b + c)$. Once again π is decreasing in $(\underline{e} - e_i)$ hence l_i has an incentive to decrease e_i until $e_i = \underline{e}$ leading to $\pi = (a - b) + (b - c)\underline{e}$. Since $b > c$ π is increasing in $\underline{e}$. Next we show that l_i does not have an incentive to undercut $\underline{e}$. Setting $e_i < \underline{e}$ leads to payoffs $\pi = (a - b) + (b - c)e_i$. Again payoff is increasing in e_i and reaches its maximum at $e_i = \underline{e}$. d) This follows straight from the fact that if $e_k = 0$ for some k then $\min(e_i, \underline{e}) = 0$ for all i. □

Network Design

T faces the same recruitment pool, n_l in each locality l. We know that $b(g_l, \theta)$ is rising in n, so T when deciding between (c, d) for each location T will include as many recruits as possible in the given network structure, subject to his edge restriction.

Because of the symmetry in $T's$ problem we drop the subscript l. In each location T can design a centralised network with n recruits. Alternatively he can design a decentralised network with n_d recruits such that $n_d(n_d-1) \leq 2(n-1)$.

For the purposes of the proposition we will always use k to denote the locality with the highest cost.

Proposition 8. *a) In game Γ_h the unique SPE of the game allows T to design star networks in each locality and operate each local network at δ_m. Each local policing authority l_i sets $e_i = 0$.*

b) In game Γ_l there are multiple SPE of the game in which $e_i > e*$ for all i and T designs a regular network in any one locality and star networks in every other locality and operates each local network at δ_m. And another SPE in which $e_i \leq e^*$ for all i and T designs star networks in each locality and operates them at δ_m.

c) The game Γ_s has multiple SPE. If $e_i \leq e^*$ for all i then T designs star networks in each locality and operates at δ_m. If $e_i > e^*$ then T designs a star network in every locality but in k operates at δ_s^k such that $b_k^s(\delta_s^k) = c_k$ and operates at δ_m everywhere else.

Proof. a) Denote by k the local policing authority that satisfies $c_k = \max C_h$. Lemma BR implies that l_k sets $e_k = 0$ for any configuration of g. From the same lemma it follows from part (c) that given $e_k = 0$ all other $i \in I, i \neq k$ will set $e_i = 0$. Given this T will design star networks in each location and maximises total output by operating at δ_m.

b) This follows straight from the fact that if $c_k < b_s(\delta_0)$ then $c_i < b_s(\delta)$ for all $i \in I$ for all $\delta \in (0, \delta_m)$. When this is the case every $i \in I$ has an incentive to cooperate if facing a star network and will set $e_i = \underline{e}$. Any $\underline{e} \in [0,1]$ constitutes a NE of this subgame. If T does design star networks in each locality he will operate at δ_m and obtain $(1-e)Ib_s$ depending on the equilibrium that arises in the minimum effort game. The other option for T is to design a regular network in any one location, let this location be k, and induce $e_k = 0$. Given $e_k = 0$ we have $e_i = 0$ for all $i \neq k$ allowing T to design stars everywhere else. This will lead to a pay off of $(I-1)b_s + b_r$. Designing star networks in each locality dominates only for small levels of coordination equilibria $e \leq e_t \equiv \frac{b_x - b_r}{Ib_s}$.

c) Denote δ_s^k such that $b_s(\delta_s^k) = c_k$. This allows us to partition the δ-space into two regions. In the first region we have that $b_s^k \leq c_k$. In the second region we have that $c_k < b_s^k$. In the first region by operating on a low δ, T can induce

$e_k = 0$ and thus induce $e_i = 0$ for all $i \in I$ and $i \neq j$. But since b_k is increasing in δ, T can do strictly better by setting $\delta = \delta_s^k$ in locality k. At this point the benefit of coordinating is exactly offset by the costs since $b_s^k = c_k$ so k will set $e_k = 0$. Alternatively T can operate in region 2. Since $c_k < b_s^k$ in this region T would have to design a regular network to induce no cooperation. Again b_r^k is increasing in δ and output is maximised when $\delta = \delta_r^k$. However we know that $b_s^k(\delta_s^k) > b_s^k(\delta_0) > b_r^k(\delta_m)$ implying that T will never want to induce no coordination in k by designing a regular network. Given T operates at c_k in locality k there is no coordination and T can operate star networks in all other localities at δ_m. Total output to T from this setup is $c_k + (I-1)b_s(\delta_m)$. Alternatively, T can design star networks, operating at δ_m, in all localities. This will induce cooperation and total expected output to T will be $Ib_s(\delta_m)$. T is indifferent between these two configurations when $e = 1 - \frac{b_s(\delta_m)-c_k}{Ib_s(\delta_m)} \equiv e_t$. This leads to the multiple SPE stated in the proposition.

□

6 Discussion and Future Research

We present two models to analyse the strategic interaction between a criminal network planner and a set of independent policing authorities. In the first model we provide a micro-foundation for rationalising an individuals decision to join a criminal network, he should earn as much utility from being in the network as he would from his outside option. This forms one of the constraints the network planner faces when designing an optimal network. Individual utility in the network is hetergenous and a function of network centrality. In order for the network to be stable and supported at equilibrium the network planner needs to make sure that minimum utility within the network is greater than the outside option.

Policing authorities monitor the network activities but do so with a bias. Depending on network activity and policing bias the policing authority assigns a probability of detection to each individual criminal which influences his willingness to participate. In equilibrium the detection profile is increasing in policing skill and an individuals probability of detection is a weighted average of his Bonacich and Intercentrality relative to total Bonacich centrality. A high skilled policing authority places more weight on Intercentrality relative to a low skilled. Optimal policing time and punishment are a function of policing skill and network primitives.

Through its monitoring a relative high skilled policing authority can influence the network primitives and induce the network participants to put more weight on local complementarities relative to global substitutabilities when deciding optimal effort levels. The network planner factors this in when designing a network and is forced to design a smaller network. In equilibrium centralised structures can arise when either the outside option or policing resource and punishment is extremely low or both. When either is high decentralised structures arise. The final size of the network structure depends on the resource endowment of the network planner measured in terms of how many links he can manage.

In the second model we allow policing authorities to coordinate with one another and show that when the cost of coordination is a function of policing time and skill, the network planner will induce one policing authority not to coordinate by designing a decentralised network and maximise global output by designing centralised structures everywhere else. When the expectation of coordination is very low the network planner may also design centralised structures in all locations.

The title of this paper is not meant to be misleading. We would like to take our findings towards a unified analysis of the two models presented here. What we have in mind is a rather more complicated but interesting structure that uses the major features of both models.

We propose the study of a two stage game with incomplete information. In the first stage policing authorities act independently and gather information on the network characteristics. In the second stage they are given the option to either act independently again or coordinate with one another. During the first stage of the game policing authorities will draw a skill level from a known distribution, but the precise skill level is private knowledge. Only if policing authorities coordinate in the second stage does this skill level become common knowledge amongst policing authorities. To add to the problem each policing authority will be able to choose from a high or low level of policing resources devoted to policing. Choosing a low level allows the policing authority to upgrade to a higher skill level in the next stage. Potentially the network planner can choose from a set of network primitives.

We think this problem is interesting as it reflects reality. It will allow us to study how a criminal organisation distributes actions across different locations depending on the skill set it faces and its own limitations. How policing authorities endogenously decide to coordinate or act independently and whether this decision influences the centralisation of the criminal network.

7 Appendix - Proofs

Proposition 1. Notice that our utility specification is identical to that provided in Ballester et al. (2006) when

$$\alpha = 1 - \pi \ , \ \beta = \hat{\beta} \ , \ \gamma = \hat{\delta} \ , \ \lambda = \hat{\lambda}$$

Let $\theta = \frac{\hat{\lambda}}{\hat{\delta}}$. As long as $\theta\rho(g) < 1$ we can use Theorem 1 from**?** to the write the general effort level profile in equilibrium as

$$x_i^*(g,\theta) = \frac{1-\pi}{\hat{\beta} + \hat{\delta}b(g,\theta)} b_i(g,\theta) \tag{7}$$

In the special case where $\beta = \delta \Leftrightarrow \hat{\beta} = \hat{\delta}$ for any $\pi, s \in [0,1]$ we have

$$x_i^*(g,\theta) = \frac{1-\pi}{\hat{\delta}(1 + b(g,\theta))} b_i(g,\theta) \tag{8}$$

Also note that when $\beta \geq 0$ and $\delta = 0$, Assumption 3 implies that $\hat{\delta} = 0$ and we have that

$$x_i^*(g,\theta) = \frac{1-\pi}{\hat{\beta}} b_i(g,\theta) \tag{9}$$

□

Proposition 2. In evaluating u_i in equilibrium we will make use of one of the properties of the matrix $\mathbf{M}$ which we state explicitly. Note that

$$\mathbf{b} = (\mathbf{I} - \theta\mathbf{G})^{-1}.\mathbf{1} \Leftrightarrow (\mathbf{I} - \theta\mathbf{G})^{-1}.\mathbf{b} = \mathbf{1}$$

The above identity allows us to write $b_i - \theta\sum_{j=1}^{n} g_{ij}b_j = 1$. Rearranging we obtain

$$\theta\sum_{j=1}^{n} g_{ij}b_j = b_i - 1 \tag{10}$$

Substituting 8 into 3 and making use of 10 we obtain

$$u_i(\mathbf{x}^*(g,\theta)) = \frac{(1-\pi)^2}{\hat{\delta}(1 + b(g,\theta))^2} b_i(g,\theta)^2$$

In equilibrium we require that $u_i(\mathbf{x}^*(g,\theta)) \geq \omega$ and rearranging gives us the desired result. □

Lemma 2. Given the network structure d and a pre-determined choice of $(\beta, \delta, \lambda, \pi, s)$ an individual recruit chooses an optimal effort level x_i to maximise equation 3. The F.O.C of this maximisation w.r.t. x_i yields

$$(1-\pi) - (\hat{\beta} - \hat{\delta})x_i - \hat{\delta}\sum_{j=1}^{n} x_j + \hat{\lambda}\sum_{j=1}^{n} d_{ij}x_j = 0$$

Because of the symmetry in structure of a decentralised graph, in equilibrium we have $x_i^* = x_j^* = x^*$ for all $j \neq i \in d$. Rearranging the above expression after direct substitution of x^* gives us

$$x^* = \frac{1-\pi}{\hat{\beta}\left(1 + n\frac{\hat{\delta}}{\hat{\beta}} - \theta(n-1)\right)}$$

Where $\theta = \frac{\hat{\lambda}}{\hat{\beta}}$. In the special case we are considering where $\hat{\beta} = \hat{\delta}$ we have the desired result stated in the proposition.

Using Proposition 1 we have that

$$\frac{(1-\pi)b_i(d,\theta)}{\hat{\delta}(1+b(d,\theta))} = \frac{1-\pi}{\hat{\delta}(n+1-\theta(n-1))}$$

Noting that in the symmetric equilibrium we can write $b(d,\theta) = nb_i(d,\theta)$, rearranging the above in terms of b_i gives us the desired result

$$b_i(d,\theta) = \frac{1}{1-(n-1)\theta}$$

□

Lemma 3. Given the network structure c and a pre-determined choice of $(\beta, \delta, \lambda, \pi, s)$ an individual recruit chooses an optimal effort level x_i to maximise equation 3. The F.O.C of this maximisation w.r.t. x_i yields

$$(1-\pi) - (\hat{\beta} - \hat{\delta})x_i - \hat{\delta}\sum_{j=1}^{n} x_j + \hat{\lambda}\sum_{j=1}^{n} c_{ij}x_j = 0$$

Denote x_h the equilibrium level effort level exerted by the hub, connected to all other nodes. We exploit the structure of c and note that the symmetry of c for all periphery nodes $x_j\{j \neq h\}$ will induce all periphery nodes to exert the same effort level in equilibrium. Denote x_p the equilibrium level effort level exerted by an individual periphery. Then the F.O.C for the hub can be written as

$$(1-\pi) - (\hat{\beta} + \hat{\delta})x_h + (\hat{\lambda} - \hat{\delta})(n-1)x_p = 0$$

Rearranging we have

$$x_h = \frac{(1-\pi) + (\hat{\lambda} - \hat{\delta})(n-1)x_p}{\hat{\beta} + \hat{\delta}} \tag{11}$$

Similarly the F.O.C for a periphery can be written as

$$(1-\pi) - (\hat{\beta} + (n-1)\hat{\delta})x_p + (\hat{\lambda} - \hat{\delta})x_h = 0$$

Direct substitution of equation 11 for x_h, into the above and rearranging in terms of x_p gives

$$x_p = \frac{(1-\pi)(\hat{\beta} + \hat{\delta})}{(\hat{\beta} + \hat{\delta})(\hat{\beta} + (n-1)\hat{\delta}) - (n-1)(\hat{\lambda} - \hat{\delta})^2} \tag{12}$$

In the special case where $\hat{\beta} = \hat{\delta}$ and $\theta = \frac{\hat{\lambda}}{\hat{\delta}}$ equation 12 simplifies to the proposition

$$x_p = \frac{(1-\pi)(1+\theta)}{\hat{\delta}((n+1) + 2(n-1)\theta - (n-1)\theta^2)} \tag{13}$$

Substituting equation 12 into equation 11 gives us, after some algebraic manipulation

$$x_h = \frac{(1-\pi)[\hat{\beta} + (n-1)\hat{\lambda}]}{(\hat{\beta}+\hat{\delta})(\hat{\beta}+(n-1)\hat{\delta}) - (n-1)(\hat{\lambda}-\hat{\delta})^2} \tag{14}$$

In the special case where $\hat{\beta} = \hat{\delta}$ and $\theta = \frac{\hat{\lambda}}{\hat{\delta}}$ equation 12 simplifies to the proposition

$$x_h = \frac{(1-\pi)(1+(n-1)\theta)}{\hat{\delta}((n+1) + 2(n-1)\theta - (n-1)\theta^2)} \tag{15}$$

Using Proposition 1 we note that

$$\frac{b_h}{1+b} = \frac{1+(n-1)\theta}{(n+1)+2(n-1)\theta-(n-1)\theta^2}$$

and

$$\frac{b_p}{1+b} = \frac{1+\theta}{(n+1)+2(n-1)\theta-(n-1)\theta^2}$$

We also note that $b = b_h + (n-1)b_p$ which allows us to write

$$\frac{b}{1+b} = \frac{n+2(n-1)\theta}{(n+1)+2(n-1)\theta-(n-1)\theta^2}$$

Using these three facts we can write the Bonacich centralities of the hub and periphery, respectively, as

$$b_h = \frac{1+(n-1)\theta}{1-(n-1)\theta^2}$$
$$b_p = \frac{1+\theta}{1-(n-1)\theta^2}$$

Thus proving the lemma. □

Proposition 4. In order to prove parts 1-3 of Proposition 4 we will make direct use of a result taken from Ballester et al. (2006), stated here as a lemma.

Lemma 11. *Let g and g' be symmetric networks such that $g \subset g'$. If $\theta\rho(g') < 1$ then in equilibrium total effort level under g' is strictly higher than under g i.e. $x(g', \theta) > x(g, \theta)$.*

Proof. See proof of Theorem 2 in Ballester et al. (2006) □

Consider two networks g and g'. We have that g is nested in g', $g \subset g'$, if (a) g and g' connect the same number of nodes and if it uses less direct links than g', or, (b) If g' connects additional nodes to the network g.

(1) Let $(c(n), d(n))$ respectively denote a star and a regular network with n nodes. Without loss of generality fix $n' > n$. So that $s(n) \subset r(n)$ and $s(n') \subset r(n')$ are specific cases of (a) and we use Theorem 2 from Ballester et al. (2006).

(2) $s(n) \subset s(n')$ and $r(n) \subset r(n')$ are special cases of (b) and again we use Theorem 2 to conclude.

However notice that taken together $s(n) \subset r(n)$, $s(n) \subset s(n')$, $r(n) \subset r(n')$ and $s(n') \subset r(n')$ do not imply $r(n) \subseteq s(n')$. It is no longer clear that a centralised with more nodes induces higher aggregate effort compared to a decentralised network with less nodes. This issue becomes increasingly important when the network architecture is limited by the number of feasible edges but not the number of available nodes.

In order to prove part (3) we will first show that there exists a $\theta^*(n, n')$ such that $x(c(n'), \theta^*) = x(d(n), \theta^*)$. We will then proceed to show that this θ^* is always smaller than the inverse of $\rho(d(n))$ - the maximum bound on the decentralised network. Finally we compare aggregate effort levels from the two types of graphs when $\theta \neq \theta^*$.

In equilibrium aggregate effort level for a given g is

$$x(g, \theta) = \frac{1-\pi}{\hat{\delta}} \sum_{i=1}^{n} \frac{b_i(g, \theta)}{1 + b(g, \theta)} = \frac{1-\pi}{\hat{\delta}} \frac{b(g, \theta)}{1 + b(g, \theta)}$$

Note that $x(g, \theta)$ is concave in $b(g, \theta)$ i.e.

$$\frac{\partial x}{\partial b} = \frac{1}{(1+b)^2} > 0 \quad \frac{\partial x}{\partial b} = -\frac{2}{(1+b)^3} < 0$$

for any $b(g, \theta) \geq 0$. Hence evaluating θ such that $x(g', \theta) = x(g, \theta)$ is equivalent to evaluating θ such that $b(g', \theta) = b(g, \theta)$.

With $c(n')$ we denote the centralised network with n' nodes, and $d(n)$ denotes a decentralised network with $n \leq n'$ nodes. Using Lemmas 4 and 5 we have that

$$b(c(n'), \theta) = b(d(n), \theta) \Leftrightarrow \frac{n' + 2(n'-1)\theta}{1 - (n'-1)\theta^2} = \frac{n}{1 - (n-1)\theta}$$

Simple algebraic manipulation shows that $b(c(n'), \theta^*) = b(d(n), \theta^*)$ when

$$\theta^* = \frac{n' - n}{(n'-1)(n-2)} \tag{16}$$

To show how $b(c(n'), \theta)$ and $b(d(n), \theta)$ compare when $\theta \neq \theta^*$ we first write aggregate Bonacich Centrality as a sum of its components

$$b(g, \theta) = \sum_{i=1}^{n} b_i(g, \theta)$$

and individual Bonacich centrality as a sum of its components

$$b_i(g,\theta) = m_{ii}(g,\theta) + \sum_{j\neq i}^{n} m_{ij}(g,\theta)$$

remembering that

$$m_{ij}(g,\theta) = \sum_{k=1}^{+\infty} \theta^k g_{ij}^k$$

It is clear that $b_i(g,\theta)$, and therefore $b(g,\theta)$ is an infinite polynomial in θ with real and positive coefficients. This implies that $b(g,\theta)$ is continuous and strictly monotone in θ.

By definition $\lim_{\theta\to 0} m_{ii}(g,\theta) = 1$ and $\lim_{\theta\to 0} m_{ij}(g,\theta) = 0$. Hence $\lim_{\theta\to 0} b_i(g,\theta) = 1$ and $\lim_{\theta\to 0} b(g,\theta) = n$. Since $n' > n$ the above argument with continuity implies that $b(c(n'),\theta) > b(d(n),\theta)$ for $\theta < \theta^*$.

We note that θ^* is unique in (n,n'). This along with the properties of continuity and strict monotonicity imply that $b(c(n'),\theta) < b(d(n),\theta)$ for all $\theta > \theta^*$.

We now show that under certain conditions θ^* is always less than the inverse of $\rho(d(n))$. θ^* is strictly monotone in n'

$$\frac{\partial\theta^*}{\partial n'} = \frac{(n-2)(n-1)}{((n'-1)(n-2))^2} > 0$$

for any $n > 2$ and using Lemma 1

$$\frac{n'-n}{(n'-1)(n-2)} \leq \frac{1}{\rho(d(n))} \Leftrightarrow \frac{n'-n}{(n'-1)(n-2)} \leq \frac{1}{n-1} \Leftrightarrow n' \leq n^2 - 2n + 2$$

In order to conclude the proof we draw attention to the edge distribution of $c(n')$ and $d(n)$. We note that for a given n the number of edges used up by $d(n)$ in connecting all nodes is $\frac{n(n-1)}{2}$. Equivalently, by using at most the same number of edges in $d(n)$ we can connect $n' \leq \frac{n(n-1)}{2} + 1$ in a centralised network $c(n')$. For $n > 2$

$$n' \leq \frac{n(n-1)}{2} + 1 \Rightarrow n' \leq n^2 - 2n + 2$$

We can potentially use more edges than in $d(n)$ and connect $n' > \frac{n(n-1)}{2} + 1$. However if we use more than an extra n edges we can also move from $d(n)$ to $d(n+1)$ which is a strictly preferred configuration to $d(n)$ given Proposition 2. So a natural bound for comparing $d(n)$ and $c(n')$ is $n' \leq \frac{n(n-1)}{2} + n$. For $n > 4$

$$n' \leq \frac{n(n-1)}{2} + n \Rightarrow n' < n^2 - 2n + 2$$

And we have that for $n > 4$ and $n' \in [n, \frac{n(n+1)}{2}]$

$$\theta^* \in [0, \frac{1}{n-1})$$

□

Corollary 1. *For the special case of* $n' = \frac{n(n-1)}{2} + 1$*, we have that* $\theta^* = \frac{1}{n}$*.*

Proof. Substitute $n' = \frac{n(n-1)}{2} + 1$ in equation 16 and simplify. □

References

Atkinson, Scott E, Todd Sandler, and John Tschirhart, "Terrorism in a Bargaining Framework," *Journal of Law & Economics*, April 1987, *30* (1), 1–21.

Baker, Wayne E. and Robert R. Faulkner, "The Social Organization of Conspiracy: Illegal Networks in the Heavy Electrical Equipment Industry," *American Sociological Review*, 1993, *58* (6), 837–860.

Ballester, Coralio, Antoni Calvó-Armengol, and Yves Zenou, "Who's Who in Networks. Wanted: The Key Player," *Econometrica*, 09 2006, *74* (5), 1403–1417.

Bandiera, Oriana, "Land Reform, the Market for Protection, and the Origins of the Sicilian Mafia: Theory and Evidence," *Journal of Law, Economics and Organization*, April 2003, *19* (1), 218–244.

Bar-Isaac, Heski and Mariagiovanna Baccara, "How to Organize Crime," 2006, (06-07).

Becker, Gary S., "Crime and Punishment: An Economic Approach," *Journal of Political Economy*, 1968, *76* (2), 169.

Benmelech, Efraim and Claude Berrebi, "Attack Assignments in Terror Organizations and The Productivity of Suicide Bombers," NBER Working Papers 12910, National Bureau of Economic Research, Inc February 2007.

Berman, Eli, "Hamas, Taliban and the Jewish Underground: An Economist's View of Radical Religious Militias," NBER Working Papers 10004, National Bureau of Economic Research, Inc September 2003.

_ **and David Laitin**, "Hard Targets: Theory and Evidence on Suicide Attacks," NBER Working Papers 11740, National Bureau of Economic Research, Inc November 2005.

Bonacich, Phillip, "Power and Centrality: A Family of Measures," *The American Journal of Sociology*, 1987, *92* (5), 1170–1182.

Burdett, Kenneth, Ricardo Lagos, and Randall Wright, "Crime, Inequality, and Unemployment," *American Economic Review*, December 2003, *93* (5), 1764–1777.

C., Calvo-Armengol A. Zenou Y. Ballester, "Delinquent Networks," *Working Paper*, June 2008.

Garicano, Luis and Esteban Rossi-Hansberg, "Organization and Inequality in a Knowledge Economy," *The Quarterly Journal of Economics*, November 2006, *121* (4), 1383–1435.

Garoupa, Nuno, "The Theory of Optimal Law Enforcement," *Journal of Economic Surveys*, September 1997, *11* (3), 267–95.

_ , "Optimal law enforcement and criminal organization," *Journal of Economic Behavior & Organization*, July 2007, *63* (3), 461–474.

Glaeser, Edward L., Bruce Sacerdote, and Jose A. Scheinkman, "Crime and Social Interactions," NBER Working Papers 5026, National Bureau of Economic Research, Inc February 1995.

Hart, Oliver and John Moore, "On the Design of Hierarchies: Coordination versus Specialization," ESE Discussion Papers 117, Edinburgh School of Economics, University of Edinburgh March 2004.

Huang, Chien-Chieh, Derek Laing, and Ping Wang, "Crime And Poverty: A Search-Theoretic Approach," *International Economic Review*, 08 2004, *45* (3), 909–938.

Klerks, P, "The network paradigm applied to criminal organisations: Theoretical nitpicking or a relevant doctrine for investigators? Recent developments in the Netherlands," *Connections*, 2001, *24* (3), 53–65.

Lapan, Harvey E and Todd Sandler, "To Bargain or Not to Bargain: That Is the Question," *American Economic Review*, May 1988, *78* (2), 16–21.

Maskin, Eric, Yingyi Qian, and Chenggang Xu, "Incentives, Information, and Organizational Form," *Review of Economic Studies*, April 2000, *67* (2), 359–78.

Mesquita, Ethan Bueno De and Catherine Hafer, "Public Protection Or Private Extortion?," *Economics and Politics*, 03 2008, *20* (1), 1–32.

Polinsky, A. Mitchell and Steven Shavell, "The Economic Theory of Public Enforcement of Law," NBER Working Papers 6993, National Bureau of Economic Research, Inc March 1999.

Radner, Roy, "Hierarchy: The Economics On Managing," 1990, (90-07).

_ , "The Organization of Decentralized Information Processing," *Econometrica*, September 1993, *61* (5), 1109–46.

Rantakari, Heikki, "Governing Adaptation," *Review of Economic Studies*, October 2008, *75* (4), 1257–1285.

Sandler, Todd and Daniel G. Arce, *Terrorism: A Game-Theoretic Approach*, Vol. 2 of *Handbook of Defense Economics*, Elsevier,

_ **and Kevin Siqueira**, "Global terrorism: deterrence versus pre-emption," *Canadian Journal of Economics*, November 2006, *39* (4), 1370–1387.

Verdier, T. and Y. Zenou, "Racial Beliefs, Location and the Causes of Crime," Discussion Paper Series In Economics And Econometrics 0101, Economics Division, School of Social Sciences, University of Southampton January 2001.

W., Alonso R. Dessein and N. Matouschek, "When does Coordination require Centralization," *American Economic Review*, March 2008, *98* (1), 145–179.