How to Make People Work Without Direct Supervision (Short version)

Chenyang Li

June 25, 2022

Abstract

Can you make people work without directly supervising them?The answer is yes if you have multiple agents by creating an architecture where they supervise one another. Complication arises from the fact that, the agents may collude and jointly deviate to no effort and no peer supervision. This paper models the collusion formation process and characterizes the conditions under which the collusion may or may not occur. The central insight is: if the principal can limit the communication among the agents, it is much easier to deter collusion.

I study two ways of joint deviation: voting and commitment. When all agents are directly connected in a communication network, deviation by voting can be stopped if and only if the threshold of passing the vote is sufficiently high. Deviation by commitment, however, cannot be stopped. However, if the principal limits the initial communication network to a "ring", the joint deviation can be deterred when the passing threshold is at least three people no matter the total number of players. Commitment can be stopped when there are at least six agents in the department. The negotiation power of an arbitrary individual in an arbitrary network can also be calculated by an algorithm. The findings give us insights into firm management, and political control, and also designing mechanisms for controlling corruption.

1 Introduction

This paper is about the art of managing and controlling people without directly supervising them. The key lies in designing how people relate to and monitor one another. I shall start with an example that shows why at times the principal has difficulty contracting punishment or rewards based on the agents effort or the outcome. Think of a Governor choosing between two policies: the first one is optimal for the society, and the second one serves his private interest. In other words, the governor needs to exert a costly effort to choose the first policy over the second one. Assume that both actions are legal. The governor makes hundreds of such choices every day, so the general public (the principal) cannot supervise each decision directly because of the high cost. The people want to design a political system in which the governors supervise one another (because they can easily observe one another's decision) and exert such a socially beneficial effort without the principal's direct involvement. An example is the Commission for Discipline Inspection in China.

A standard result is that there are subgame perfect equilibria in which peer supervision can sustain positive effort level. However, these equilibria seem unstable because exerting effort and punishing the peers reduces a player's utility and thus pushes them off the Pareto Optimal boundary. Since the agents supervise each other, they should also be able to communicate and coordinate a joint deviation to a better equilibrium. For instance, no one exerts any effort, and no one carries out any punishment.

There are good reasons to question the feasibility of using peer supervision to reinforce Pareto-dominated outcomes. Farrell and Maskin (1989) has this comment on the renegotiation process: "unless players somehow cut the line of communication, it seems possible that they can renegotiate after the game begins, that they will not follow mutually-unpleasant subgame-equilibrium path when there is a Pareto dominating alternative available, even if they agree to do so when the game begins." Though intuitive, this comment implicitly assumes that the renegotiation stage is not a part of the game, so naturally, the players would renegotiate to reach efficiency. I find that if the principal can limit the communication network among the players, the effort provision equilibrium can be much more robust against collusion [1].

In this paper, I study an mechanism design problem in which the principal cannot commit to punishment or reward that is based on the principal's direct observation of the agent's effort level or the outcome (No direct supervision[2]). So the contract has to rely on peer supervision. The game has infinite periods and every period includes two stages. On the first stage, each agent chooses an effort level which is costly. The effort level is observed

[1] I will define the robustness measure later in this paper

[2] This is an extreme assumption. If the principal can exert some monitoring, the resulting equilibrium will be more robust

by every other agent. In the second stage, the agents choose whether to punish each other. Thus by the Folk Theorem, there is a set of equilibria such that the agents exert a strictly positive level of effort.

I assume the principal can choose an initial equilibrium for the agents through pre-job training or other similar ways to set an initial equilibrium. However, as the department starts to work, the agents can talk to each other and plan for a collusion. I study two potential paradigms for endogenous equilibrium selection: 1. Deviation by voting. 2. Deviation by commitment. They are just two of many potential ways the players may bargain for equilibrium, but they are commonly used in reality, and we can think about real-life examples corresponding to each of them. More discussion of the examples are provided in the last section.

The first paradigm allows all the agents to vote for joint deviation. There is an exogenously given threshold of m. If at least m out of n agents vote for yes, then the entire department deviates to a new equilibrium chosen by the initiator of this vote. For instance, the law requires that half of the workers vote yes before a labor union can be established. In such a case, $m = \frac{n+1}{2}$. The second paradigm is "stronger": Players can commit to strategies that are not incentive compatible. Some players might have an exogenous reputation, so others trust her commitment. Other commitments are like legally binding contracts that force each player to carry out certain activities at each point in history.

There is a quick preview of the results. First, when all the agents can directly communicate with one another: 1. Deviation by voting can only be stopped when the threshold for passing the vote (m) is higher than some cutoff value. 2. The principal cannot stop deviation by commitment. However, suppose the principal can limit the initial communication network among the agents, both types of deviation can be stopped under much more relaxed conditions. There are two constraints on the communication network: 1. if one agent can punish another, both of them must be directly connected in the communication network. 2. There must be a direct of indirect communication path between any two agents in the game because utility transfers and information concerning the history of game has to go through a communication path. If there are two disconnected sections of the communication network, they do not affect each other and has to be studied separately.

Under the two constraints, the principal can put all the agents in a ring facing the center, and every agent starts with only connecting to the two agents on his left and right. Interestingly, the architectural structure is quite similar to Jeremy Bentham's famed panopticon (Bentham, 1791), even though the system of monitoring I am discussing here is quite different from that of Bentham. Then, if $m \leq 3$, all deviation by voting can be deterred no matter how many agents are there. I also show that the ring is the most robust peer supervision network among all possible communication networks in which the agents exert full effort. In the deviation by commitment, the principal can stop the corruption if the department has at least six agents by putting them into a ring. In general, when the people are less connected, the deviation to the agent-optimal equilibrium is less likely to occur.

Limiting communication to deter joint deviation has been used in the real world. A dictator may enjoy the benefit of exploiting his people by putting them into a peer supervision structure. To maintain such a desirable equilibrium, the dictator usually imposes strict limitations on communication. Methods include news censorship, cover-up, or other information manipulation. Attempts of upheaval are cracked down. Thus a revolution would be unlikely. Though all the rewards and punishments come from the people and are inefficient, people under a dictatorship would not be able to deviate from the undesirable equilibrium and overthrow the dictator.

Similarly, some large companies stop their front-line workers from forming a labor union by limiting communication among them. Managers get rewards if he or she reports on the flyers that attempt to unionize the workers. Companies hire union-busting services to crack down on union groups on social media.

There are also positive uses of limiting communication. For instance, to prevent corruption of the cashier, the accountant, and the warehouse managers, employers usually adopt the practice called: "separation of duties", which means different people are in charge of different positions. Thus, no single agent can fake evidence and cover up his or her illegal actions. Attempts to establish a personal connection would also be punishable, resulting in violators being dismissed from the position.

This paper has the following contributions: 1. I propose a new way to model joint deviation. Instead of using equilibrium refinement, I directly model the equilibrium selection process as a part of the repeated game. 2. I study two paradigms of how players may reach an agreement in each case and deter Pareto-improving collusion. 3. I find that limiting communication among the players can significantly improve the robustness of effort provision equilibrium, and this conclusion has many real-world applications, as presented below.

There are two main strands of application of this paper: 1. Construct "autonomous apartments" in which the agents exert the optimal effort and are robust against collusion. 2. For groups of people stuck in inefficient equilibrium but unable to negotiate away from it, this paper also provides internal or external intervention methods.

In the following subsection, I discuss the related literature. In section 2, I will set up the base model and derive the highest level of effort that a peer punishment structure can sustain. In sections 3 and 4, I study joint deviation by voting, commitment, and subgroup coalition. Section 5 introduces a more robust single ring supervision network. Finally, I conclude this paper with some extensions and a discussion of the findings and future projects.

Before getting into the formal model it is worth clarifying that, while the model developed in this paper is an exercise in pure theory, the results have important implications for the real world, from ideas for managing a corporation to managing political authority. The paper does not take a normative stance. The theory shows how these kinds of control can be used for both efficient control and also for exploitation. When the latter happens we hope that the theory can be used to develop more elaborate structures to prevent the principal from exploiting the workers.

1.1 Related literature (Under revision)

It is not surprising that ideas for managing and controlling large groups go far back into history, such as the eighteenth-century English philosopher Jeremy Bentham's idea of a panopticon meant to keep efficient control over a large group. We also have real-world examples of peer monitoring, such as the one used by Bangladesh's Grameen Bank, started by Yunus in 1976 (Stiglitz, 1990).

The aim of this section however is to briefly survey the related theoretical literature. The theoretical part of this project is related to the literature on monitoring or community enforcement. Most of them focus on the threat of withdrawal of the cooperation and does not consider costly punishments (Kandori, 1992, Ellison, 1994, Kranton, 1996, Wolitzky, 2013, Ali and Miller, 2014). A few papers do allow punishment, mostly focusing on enforcers' incentive to carry out punishment (Dixit, 2011, Masten and Prüfer, 2014, Levine and Modica 2016, Aldashev and Zanarone, 2017, Acemoglu and Wolitzky, 2019). The largest difference from this literature and this paper is that, all of them focus on reinforcing the efficient outcome, while I focus on reinforcing the private inefficient equilibrium to achieve a higher order target. I also study costly punishment and collusion as endogenous equilibrium selection process.

Another branch of studies concern the upper bound of welfare in repeated games that is achievable by punishment and rewards (Acemoglu and Wolitzky, 2017, 2018). They find relatively tight bounds in repeated games with an arbitrary number of individuals. My paper, on the other hand, focus on maintaining the "worst" outcome for the players. Because the more privately inefficient the outcome is, the more contribution the players can do for the rest of the society.

Akerlof (1976) first developed a model describing how "labeling" people and discriminatory social custom can result in sub-optimal outcome, and how these equilibria can be break by coalition of members. He argues the inefficient equilibrium vanishes when sufficient number of players deviates from it. His main theme was close to my paper, but he does model the coalition as an endogenous equilibrium selection process.

The equilibrium selection process in this paper is different from equilibrium refinement such as the concept of coalition proof equilibrium (Bernheim et al. (1987)) or renegotiation proof equilibrium (Farrell and Maskin (1989)). These criteria selects equilibrium when players can negotiate free of constraints, however, I put more structure on the negotiation process and model it as a part of the game.

Finally, I use the network structure to study more delicate structures of supervision. This is related to the network diffusion and network games literature. However, this paper studies how to optimally incentivize players to stop diffusion, which is different from most of the papers in this literature.

2 The Base Model: No communication

We first start with a base model in which communication among the players is not allowed. This is a perfect information finite-agent repeated game. Denote the set of players (agents) as I. There are n players in the game. I mainly focus on the cases in which $n \geq 2$, because peer supervision is not possible when there is only one agent.

The principal is not a player in this game because we assume she cannot exert reward or punishment on the agents. The only thing she can do is choose an equilibrium for the agents through a "pre-job training". In training, she tells each agent the other players' strategy and the corresponding best response. Thus, no one has the incentive for unilateral from the designated strategy when the game starts. The principal chooses the equilibrium to maximize the sum of discounted daily effort levels. Then the repeated game starts from period $t = 0$.

On each period t, there are two stages. You can think about the two stages as the morning and the afternoon of a working day. In the first stage of period $t > 0$, each players $i \in I$ can simultaneously choose an effort level e_{it} and a vector of transfers levels $\pi_{it} = (\pi_{i1t}, \pi_{i2t}, ..., \pi_{iNt})$, in which every $\pi_{ijt} \geq 0$ for all $i \in \mathcal{I}$, $j \in \mathcal{I}$ and $t \geq 0$ means the transfer from player i to player j. Here, assume $\pi_{iit} = 0$, because the transfer to oneself has no use. Denote the vector of effort $e_t = (e_{1t}, e_{2t}, ...e_{nt})$ and denote the set of transfer $\{\pi_{ijt}\}_{i,j \in \mathcal{I}}$ as $\pi(t)$. You can think about such a transfer as one player giving money to another. In reality, such transfer can be more diverse: giving praise and glory, improved friendship, giving other kinds of favor, or anything that costs one player some effort and makes the receiver happier. To avoid the Ponzi Scheme, assume the daily transfer of each agent is bounded: $\pi_{ijt} \leq \overline{\kappa}$ for all i and all j, where $\overline{\kappa}$ is a large positive number.

In the second stage of period $t > 0$, all players observe π_{ijt} and $e_{it}$for all i and $j \in \mathcal{I}$, and then each player $i \in I$ simultaneously chooses a vector $\phi_{it} = (\phi_{i1t}, \phi_{i2t}, ..., \phi_{iit}, ..., \phi_{int})$ in which for all $i \neq j$, $\phi_{ijt} \in \{0, 1\}$ and for all i, j and t. Here $\phi_{ijt} \in 0, 1$ represents the available punishment technology to each player. When $\phi_{ijt} = 0$, it mean player i does not want to initiate a fight with player j. If $\phi_{ijt} = 1$ It means player i initiates a punishment against player j, then whether j wants to fight or not, both parties lose one unit of utility. The punisher loses utility for many reasons. For instance, she may suffer emotionally for inflicting harm to her peers. This paper's main results still hold when the punisher suffers less than or equal to the punished person. Denote the set of punishment as $\{\phi_{ijt}\}_{i\in\mathcal{I}\&j\in\mathcal{I}} = \phi_t$ Also, we have $\phi_t \in \Phi$ which is the set of possible punishment in the second stage of each period t.

Let u_i denotes the per period utility that individual i receives. $u_i : \Pi \times \Phi \mapsto \mathbf{R}$. Thus we have:

$$u_{it}(\pi(t), \phi(t)) = -e_{it} - \sum_{j\in\mathcal{I}} \pi_{ijt} + \sum_{j\in\mathcal{I}\&j\neq i} \pi_{jit} - \sum_{j\in\mathcal{I}\&j\neq i} max\{\phi_{ijt}, \phi_{jit}\}$$

The first component is the loss from exerting effort. The second term is the sum of all the transfers that individual i gives away, and the third term is all the transfers he gets from the other players. The last component is player i's loss from exerting or receiving punishment on the second stage. I further assume every individual has a δ discounted utility function:

$$U_i = (1-\delta)\sum_{t=0}^{\infty} \delta^t u_{it}$$

The effort benefits the principal, so she wants to maintain an equilibrium in which the agents work as hard as possible. To maintain the effort, punishment is needed for peer supervision. Transfers are required to incentivize punishment when the agents are impatient or when an agent does not want to carry out the punishment. Furthermore, the transfers are especially important when we discuss joint deviation. In some cases, the initiator needs to promise "bribery transfers" to make another peer vote yes. I assume there is no discount from the first to the second stage of a period because this assumption gives us a more straightforward characterization of the maximum sustainable effort $\overline{e_i} = n-1$. The main results do not change if the agents discount between the first and second stages of the game.

I assume that the entire game history is perfect information for each player in this game. For tractability, I focus on pure strategy equilibria.

Denote the history on the first stage of period t as the collection of all player actions that happened before period t: $h_{1t} = \{\pi_o, \phi_0, \pi_1, \phi_1, ..., \pi_{t-1}, \phi_{t-1}\}$. And denote the history at the second stage of period t as the collection of all player actions that happened before day t plus all the player actions on the first stage of day t. $h_{2t} = \{\pi_0, \phi_0, \pi_1, \phi_1, ..., \pi_{t-1}, \phi_{t-1}, \pi_t\}$. I distinguish between h_{1t} and h_{2t} is that at the second stage of the day, all players can base their decision on what happened in the first stage of the day π_t.

Define a pure strategy of a individual $i \in \mathcal{I}$ as two functions $s_{1,i} : h_{1t} \mapsto \pi_{it}$ and function $s_{2,i} : h_{2t} \mapsto \phi_{it}$. Let $s_{1,i} \in S_1$ and $s_{2,i} \in S_2$ be the two sets of maps. Define $\mathcal{S} = \{S_1, S_2\}$

as the strategy set. This paper focuses only on pure strategy equilibria because the pure strategies alone can sustain the minmax payoff, which covers all the extreme cases that are interesting.

A complete game is defined as $\Gamma = \{I, \mathcal{S}, U, \Pi, \Phi\}$. I is the set of players, $\mathcal{G}$ is the set of strategy, U is the payoff function. Π and Φ are the action sets in the first and second stages of each period.

The principal chooses a subgame perfect equilibrium ($eq \in EQ$) to maximize the sum of discounted effort:

$$U_p(eq) = (1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}\sum_{i\in I} e_{it} \tag{1}$$

In this model, the principal's welfare improves as each $\pi_{i,i}$ increases while fixing the other variables. However, effort is privately costly to the agents, making any equilibrium with a positive effort level privately inefficient for the agents. In fact, any equilibrium that has strictly positive $\pi_{i,i}$ in this model is Pareto dominated by another equilibrium that sets $\pi_{i,i} = 0$.

2.1 Effort-Provision and Corruption

In the basic model, a natural solution concept is the Subgame Perfect Nash Equilibrium. This section establishes a benchmark of making players exert a positive amount of effort without concerning coalition or collusion. The following two terms are frequently used.

Definition: An effort provision equilibrium is any equilibrium in which there exist at least one player $i \in \mathcal{I}$ who chooses $e_{i,t} > 0$ for some periods $t \geq 0$.

Definition: A corruptive equilibrium, on the other hand, is any equilibrium in which all player $i \in \mathcal{I}$ choose $e_{i,t} = 0$ for all periods $t \geq 0$.

These two types of outcomes are the focus of this paper. The principal wants to maximize the sum of total effort, but putting the players into effort provision equilibrium as shown in the following lemma 1 However, the agents lose utility from exerting effort, and thus, if they can coordinate a joint deviation, they will never specify any positive effort level. In the base model, we assume the joint deviation and establish a benchmark model of how much effort the principal can harvest from the department.

Lemma 1. *Given the number of player $n \geq 2$ and a discount factor δ larger than $\frac{1}{2}$, the largest daily effort level for each agent is $n-1$*

Proof of lemma 1: The proof can be directly derived from the Folk theorem using a trigger strategy. When there are n agents in a department, the minmax payoff for each is $-(n-1)$. Thus, all effort level smaller than $n-1$ is individually rational and can be sustained in the repeated game. The detail of such an SPNE is the appendix A1.

Proposition 1. *Every effort provision equilibrium is Pareto dominated by a corruptive equilibrium.*

Based on this proposition, the effort provision equilibria is only stable when the agents in the department cannot talk to each other and thus cannot form coalitions or joint deviation. If all the players can jointly renegotiate, intuitively they can deviate to a Pareto Improving corruptive equilibrium. Since the workers share the same office space and can freely communicate, such joint deviation to corruption is very likely. The following corollary is another way to illustrate this problem of instability.

Corollary 1. *Any effort provision equilibrium is neither a coalition proof equilibrium (Bernheim et al. (1987)) nor weakly renegotiation proof (Farrell and Maskin (1989)).*

The detailed proof is in the Appendix A2.

However, in this paper, I find that if we explicitly model the re-negotiation process, then limiting the communication channels among the agents can significantly increase the difficulty of coordinating a joint deviation (collusion). I study two potential ways that the players may reach an agreement for joint deviation: voting and commitment. 1. Voting: If a fraction of players agrees to the deviation plan, then all the players in the entire department switch to the new equilibrium. 2. Commitment: A player can commit to any strategy whether it is incentive compatible or not. Those who do not join the commitment plan have to best respond to the commitment and non commitment group.

First, I study a model in which all the players can directly communicate with each other. If the agents choose equilibrium by voting, when the threshold of passing the vote is sufficiently high, then the deviation to corrupt equilibrium cannot form. If the players can make a binding commitment, corruption always occurs.

In the second model, suppose the principal can limit the initial communication network among the agents. Then, both deviation paradigms can be stopped. In this model, a receiver has to be reached by the initiator so that she can cast a vote. The players can ask for each other's "phone numbers" and organize joint deviations. However, the principal would set up rewards for rejecting sharing contact information. Thus, if most agents can be shielded from the initiator, the joint deviation does occur.

3 Voting for Deviation

This section provides a model of voting for equilibrium selection. Assume that all the agents follow an equilibrium selection rule: an initiator can propose a plan of new equilibria. If m out of n agents vote yes, the entire department follows the initiator's proposal. Otherwise, they keep the principal's plan. The threshold m is exogenously given.

The threshold m is usually the majority rule, but sometimes more essential decisions may require $\frac{2}{3}$ of the voters to agree. In some other cases, a few players (leaders) may have higher weights than others, and thus, a bill may be passed even when only a small fraction of players agree. In this paper, I study the effect of different level of threshold m so the result applies to more real world cases.

The formal model of the voting procedure is the following: The principal first chooses a subgame perfect equilibrium $eq_{default}$ for all the agents. All the agents start with playing this equilibrium before the voting stage.

Everyday, with a small probability p_0[3], one agent is randomly chosen to be the initiator and she comes up with a plan of deviation. Then, the initiator chooses whether to start the voting stage and pass the deviation proposal. If not, she remains silent, and all the players continue with the default equilibrium. [4]

When the voting stage starts, for individual i, her vote is denoted as $v_i \in \{yes, no, NA\}$, where NA stands for not being able to vote because the initiator does not reach the player. NA has the same effect as voting for no. Let EQ denotes the entire set of subgame perfect equilibria in the repeated game after the voting stage. A specific element is the set is denoted as eq. Define the equilibrium selection function of the principal as a mapping from the identity of the initiator and the realized votes to an equilibrium outcome:

$$eq_p : I \times V \rightarrow EQ$$

Define the equilibrium selection function of the initiator similarly as:

$$eq_i : I \times V \rightarrow EQ$$

The equilibrium selection function eq_i allows the initiator to propose favorable equilibria for those who vote yes; similarly, ep_p allows the principal to choose favorable equilibria for those who vote no. Such flexibility is crucial to the analysis.

The principal specifies the equilibrium selection function eq_p, so the agents understand if

[3]We start with small p_0 so that the possibility of deviation does not collapse all the incentives for punishments and rewards before it happens. In the extension part, I will discuss the case when p_0 is large. Those cases will be too different from the analysis here.

[4]Since coordinating the vote is costly, not everyone is willing to do so every day. Only when someone with a preference shock that makes her willing to coordinate the vote will be the initiator.

they reject a deviation proposal they follow eq_p. The initiator can observe the eq_p. Then she chooses whether to initiate the vote. If the vote begins, she chooses the eq_i to best respond to the principal and calling the other agents one by one to collect their votes. Each receiver chooses whether to vote for yes or no upon receiving the phone call. This is called sequential voting. The main results are the same if all the agents vote together after the negotiation stage is done. So I focus on the sequential voting here.

If the initiator chooses to be silent, that is equivalent to no initiator being selected for the period. All the players continue with the default equilibrium $eq_{default}$. Finally, $eq_{default}$ has to be an equilibrium under any proposal of the initiator, and it also has to be consistent with eq_p when the deviation vote is rejected. If the initiator starts the voting stage, she automatically votes for yes, and then all the other players would know who the initiator is. [5].

In this section, I assume that all the players are directly connected. In later sections, we allow the principal to design the initial communication among the agents and how such limited information helps increase the robustness of the effort provision equilibrium. Thus, any two agents can directly transmit information and money to anyone else. Also, all the communication about the voting stage is fully observed by all the agents. The voting stage ends when at least m yes votes are collected, or the initiator has contacted all her colleagues, yet not enough yes votes can be collected. We may also think about a simultaneous vote if all the receivers after the communication stage ends. However, the sequential vote is identical to the simultaneous vote if the principal specifies the order of reward given to the players who reject the vote. So in this paper, I only focus on the sequential voting paradigm.

Assumption 1. *Assume that once a deviation proposal is passed, there will be no future initiator whether the voting fails or passes.*

This assumption is strong. It represents an extreme case where the initiator does not need to worry about her proposal being overturned by future initiators. So she can promise the yes voters more desirable equilibria than what she can do without this assumption. Results derived with this assumption represents an upper bound of the initiator's negotiation power. Consequently, if the principal can stop the deviation with this assumption, the principal can also stop all potential deviations without this assumption.

Let $R \in pass, fail$ denotes the result of the vote. If there are at least m players vote for yes, then all the players follow the equilibrium selection function of the initiator, and if there are insufficient yes votes, then the equilibrium selection function of the principal would be followed. Assume all the agents have the same patience level δ, and they choose the strategy to maximize the discounted total welfare. Let $u_{i,t}(eq)$ denotes the payoff to player i at day t

[5] I assume that the players can freely pass around the information of who is the initiator

when the equilibrium is eq. Let $U_i(eq) = \sum_{t=1}^{\infty} \delta^{t-1} u_{i,t}(eq)$ denote utility of player i when the equilibrium in the repeated game turn out to be eq. By the Folk theorem, $u_i(eq) \geq -(n-1)$[6] for all i and all eq.

I allow both the principal and the initiator to choose an equilibrium selection function instead of a single equilibrium because the former setting enables both parties to discriminate between the agents who vote for yes or no. Such discrimination is crucial for the successful deterrence of undesired deviation.

The Principal's Goal Assume the principal still wants to maximize the discounted sum of effort of the agents. However, since the agents can now coordinate a joint deviation, and no effort would be provided after that, the principal also cares about deterring the collusion. We can further assume that the principal would suffer from an additional loss $C \geq 0$. once the agents collude. A larger C means the principal is more collusion averse. Thus, in the following section, I solve the model and explore conditions under which it is possible to deter collusion.

3.1 Solving the Equilibrium

In this model, I use the standard practice in the voting literature and assume that players would not choose weakly dominated strategies in the voting stage. This assumption will eliminate equilibria in which the agents reject Pareto Improving proposals because too many other players decide to vote for no. The resulting equilibrium will thus be unique.

Lemma 2. *If a joint deviation is passed, no player exerts a strictly positive effort in the new equilibrium.*

The proof is the following: if a deviation to a new equilibrium eq can be passed, yet some players still need to exert a positive amount of effort, then eq cannot be an optimal deviation proposal of the initiator. The initiator can be strictly better off choosing another equilibrium eq' with the same punishment and reward structure, but whenever the effort level is $e_j > 0$ for each player j in eq, the alternative equilibrium eq' has strategy $e'_j = 0$ and $\pi'_{ji} = e_j$ for all agents whose $e_j > 0$ and i is the initiator. This change from eq to eq' means, the initiator collects all the benefit of effort and keep them to herself. Though this change, all the other players has the same utility level in all the sub-games, but the initiator is strictly better off. Consequently, it is important for the principal to know when and how to stop the joint deviation.

This section discusses how the principal may stop joint deviation using conditional punishment and reward. The main result of this section is the following proposition:

[6] This proof is in Appendix B1.

Proposition 2. *The joint deviation to no effort can be passed if and only if .*

$$m \leq \frac{n^2+n+1}{2n-1} \tag{2}$$

The detailed proof is in appendix B2. The intuition is that to stop the joint deviation (coalition), the principal must make at least $n-m+1$ players vote for no. On the other hand, the initiator needs to make precisely m players vote for yes. Thus, the initiator could choose an equilibrium selection function to promise favorable equilibria to the yes voters and unfavorable equilibria for the no and NA voters. So if the deviation is passed, the yes voters get transfers from those who vote for no and NA. Expecting the initiator's proposal, the principal would choose the opposite reward schedule: reward those who vote for no, and such reward has to come from those who vote for yes and NA. If m is large, there would not be enough reward to incentivize enough people to vote yes. So corruption does not occur. On the other hand, if m is small, the principal would have trouble incentivizing the players to reject the deviation. So, we can derive the cutoff m by comparing the disposable rewards. Notice that the proposition 2 also implies that deviation to corruption always happens when majority voting rule is used. The next section finds that the principal can make the effort provision equilibrium more robust by limiting the communication.

3.2 Voting under Limited Communication

The previous section discusses how the agents endogenously choose equilibrium when all the players are fully connected. However, when the communication network is limited, a critical receiver can shield the initiator from reaching the rest of the department by rejecting the deviation proposal. Thus, the initiator needs to promise a very desirable equilibrium to the critical voter. As a result, the initiator may exhaust all the rewards before enough people vote for yes. Here, I provide a formal model of the voting stage under limited communication.

First, I assume the history of the voting stage is perfectly observed only by agents who receive a direct contact from the initiator. When the voting stage ends, all the other players then observe the voting result and the equilibrium to follow. This assumption implies that an agent cannot vote without receiving direct contact from the initiator because he does not know the vote is going on. Then, I define supervisors and supervisees.

Definition 1. *A player i is called the supervisor of a player j if and only if player i can initiate a punishment against j. Also, player j is called the supervisee of the player i. The set of supervisor and supervisee is listed in a $n \times n$ square matrix $\mathbf{S}$. A entry $\mathbf{S}_{i,j} = 1$ if i is the* *supervisor of j, and $s_{i,j} = 0$ if i is not the supervisor of j.* [7]

Think about the supervision network as specified by a contract. If a punishment is not allowed in the labor contract, the player cannot carry out the penalty. So the supervision network does not expand over time. Let $d_i(S)$ denotes the number of agent i's supervisor given network S. Then, we can define the communication network and the constraints on it.

Definition 2. *Let $\mathbf{C}$ be a $n \times n$ symmetric matrix that represents the communication network. If $\mathbf{C}_{i,j} = 1$, it means player i can directly send any message to player j. If $\mathbf{C}_{i,j} = 0$ then the players i and j are not directly connected.*

The communication in this model is always bilateral. This is a realistic assumption: if two players are connected, one should be able to call the other and vice versa. Ideally, the principal wants the agents to have as little communication as possible. However, it is natural to assume that each supervisor needs to communicate with their supervisees to make monitoring and punishment possible. Thus, we have the following constraint on the communication network:

Assumption 2. *For all $i, j \in N$ and $i \neq j$* [8] *If $s_{i,j} = 0$ and $s_{j,i} = 0$, then $\mathbf{C}_{i,j} = \mathbf{C}_{j,i} = 0$, otherwise $\mathbf{C}_{i,j} = \mathbf{C}_{j,i} = 1$*

Finally, I assume that all the transfers between agents have to go through some communication paths. A department with two separate components should be studied as two different departments.

The timing of the voting stage is the following: When the initiator calls a receiver and proposes a deviation plan, the receiver can choose between vote yes or no. If the receiver chooses yes, she shares all of her contact information with the initiator. It means the initiator then directly connects to all the neighbors of the receiver in the communication network. If the receiver chooses no, she does not share any contact information with the initiator. No new communication link is formed. [9] Let i denotes the initiator, and j denotes the receiver. If j votes yes, then for all $\{r \in N | \mathbf{C}_{j,r} = 1\}$, we also have $\mathbf{C}_{i,r} = 1$. If the receiver j rejects the offer, the communication network stays the same. The effect of i asking j to "pass round" the voting ticket to j's neighbors has the same effect. So, I only study the case of the initiator directly calling each receiver.

The negotiation stage stops when the initiator has called all her neighbors (including those newly connected neighbors) or if m "yes" votes are collected. Under this modified model, it

[7] $\mathbf{S}$ is a directed graph, and thus it does not need to be symmetric.

[8] N is the set of players in the department.

[9] I do not consider cases in which the receiver votes no but shares contact information or the case where the receiver votes "yes" but refuses to share the contact information because these strategies are inherently inconsistent. In the former case, sharing the contact information facilitates the joint deviation, but voting for no undermines the coalition. It is always weakly better for the receiver to not share the contact information if the receiver expects the coalition to fail or vote for yes if the receiver expects the coalition to pass. The same reasoning applies to the second case.

is harder for the initiator to organize the collusion because of the communication constraint. Each receiver has more power to reject the vote because it can block the initiator from reaching other individuals if he chooses no. We can then construct a measure of robustness for any supervision network using the following notations:

Definition 3. *let $\underline{m}_i(\mathcal{S})$ be the largest voting cutoff, such that if players i is the initiator and $\mathcal{S}$ is the supervision network, the deviation vote can be passed. $\underline{m}_i(\mathcal{S})$ is called the negotiation power of player i in supervision network S.*

Intuitively, m_i measures the number of yes votes the initiator can get given the supervision network when the principal chooses the eq_p to minimize this number. The larger the number m_i, the player i has more negotiation power because when m_i is large, it means this agent can successfully coordinate a joint deviation even when the vote is hard to pass.

Definition 4. *Let $\underline{m}(\mathcal{S})$ denotes the robustness of the supervision network $\mathcal{S}$. $\underline{m}(\mathcal{S}) =$* $\max_{i \in N}\{m_i(\mathcal{S})\}$

The robustness of a supervision network is determined by the player with largest $\underline{m}_i$. If there is at least one initiator who can successfully coordinate a deviation, then the effort provision equilibrium vanishes with probability 1 as time goes to infinity. The principal has the same goal as before. Given the voting threshold m, the principal chooses the supervision network S and equilibrium selection function eq_p first to deter collusion and then maximize the discounted sum of effort.

3.3 Network characterization

In this section, since both the robustness of a network $m(S)$ and the effort level depend on each other, we focus on the following two characterizations: 1. what is the most robust equilibrium and supervision network when players can sustain the full effort? 2. what is the most robust equilibrium when the full effort is not required?

Definition 5. *A full effort equilibrium is one that on the equilibrium path, each player i can exert effort $e_0(i) = d_S(i) - \epsilon$ in the default equilibrium, for all $\epsilon > 0$. Here $d_S(i)$ is the number of supervisors of player i S.*

We are especially interested in the full effort equilibrium for two reasons: 1. this type of equilibrium fully utilizes each individual's supervision. 2. If we assume the magnitude of damage is independent of the number of supervisors that an agent has, then the effort levels are the same for all full effort equilibria. Yet, we can find a unique, most robust supervision structure.

Reducing the effort level in the default equilibrium could further reduce each agent's incentive to coordinate a deviation. We are also interested in the most robust network without the full effort constraint. In later sections, I separately characterize the most robust network structures with or without the constraint of full effort.

I show that a single ring supervision network is the most robust network under full effort. To further increase the robustness, a star network can be used. However, the sustainable effort level decreases significantly to achieve the increment of robustness.

Definition 6. *A single ring supervision network $S \in \mathbf{S}$ is one that $s_{i,i+1} = 1$, $s_{i+1,i} = 1$ for all $1 \leq i \leq n-1$, $s_{n,1} = s_{1,n} = 1$ and $s_{i,j} = 0$ otherwise. In other word, every player is both the supervisor and supervisee of her two neighbors.*

Then, we have the following propositions:

Proposition 3. *When the threshold of passing is m and full effort is required, the principal can stop joint deviation for all $m \geq 3$ and all $n \geq m$, using a single ring supervision network and a corresponding equilibrium selection function eq_p.* [10]

The proof is in Appendix C1. This proposition means that no matter how large the department is, the principal could stop all deviation attempts as long as the threshold of passing (m) is no less than three. Thus, for all $m \geq 3$, there is a corresponding full effort supervision network that can deter joint deviation.

The main difference between the fully connected communication network and the single circle supervision network is that: in the former one, even if one receiver votes for no, the imitator can still contact the player 'behind' this rejecter, and thus the coalition may still form. When the coalition forms, the first rejecter could suffer greatly. So, if the principal wants to stop a collusion, she has to stop $n-m+1$ voter from saying yes. Such reward may be too costly when m is small. However, in the single ring supervision network, the initiator can only contact her "neighbors". If the two "neighbors" do not agree to the deviation plan, they can stop the initiator from reaching the rest of the department. Thus, I call players who can shield the initiator from the rest of the department the "gatekeepers". They have a "combined veto power", which means if they reject the offer, then the deviation cannot occur. Thus, the gatekeepers can ask for large compensation to say yes. The bribery could be so large that the initiator cannot credibly promise them to get it even if the deviation is passed. Anticipating this, the initiator would thus keep silent and stay at the effort provision equilibrium. The following two corollaries describes why the single ring supervision network is a nice combination of robustness and full effort.

Corollary 2. *Given a size of department $n \geq 3$, The single ring supervision network is weakly more robust (has weakly smaller $\underline{m}(\mathcal{S})$) than all supervision network that can sustain full effort.*

[10] When $n = 2$ and $m = 1$, then the deviation always occurs. If $n = 2$ and $m = 2$, the deviation does not occur. When $n = m = 1$, there is no effort provision equilibrium.

Corollary 3. *For all supervision networks S with robustness $m(S) = 2$, the single ring supervision network (a player is both the supervisor and supervisee of her two neighbors) generates the highest total daily effort level.*

However, the single ring is not the most robust network when full effort is not required.

Definition 7. *A most robust networks S has a robustness measure $m(S) = 1$. (In other words, deviation can be stopped if $m = 2$)*

The most robust networks is a star: there is a central player and all the other peripheral players only connects with the center. However the central node does not exert effort in the default equilibrium. She gets large transfers from the peripheral players to reduce her incentive for deviation.

3.4 Arbitrary Exogenous Supervision Network

We are also interested in finding the negotiation power of individuals in an arbitrary network. The communication network might have been exogenously given so the principal cannot put the players into a ring. Or, the principal may have a different trade-off between the robustness and the effort level and thus wants to reduce the robustness in return for a higher effort level. In this section, I provide an algorithm to solve this problem. The problem of finding the negotiation power in an arbitrary network can be transformed into a linear programming program.

3.4.1 Maximum Effort

Let S denotes the supervision network. Let r denotes the total number of supervisors in this network. After the vote fails, we can think about the peer transfers as every player giving a transfer $\sum_{j \in N} \pi_{ij}$ to a central pool, and the pool allocates the transfer to each player.

Definition 8. *Call the daily transfer from the central pool to a player i as the "reward to player i".*

Let r_p denotes the size of the central pool that principal can allocate. For every players to be incentive compatible in the default equilibrium, we must have $r_p + \sum_{i \in N} e_{it} = r$. So, to maximize the robustness, the principal chooses $e_{it} = 0$ for all i and all t after the voting stage, so $r_p = r$.

The principal can make the disposable reward of each initiator slightly smaller than r by making the initiator's effort level in the default equilibrium ϵ smaller than the number of her supervisors. Here, $\epsilon > 0$ can be arbitrarily close to zero. Thus, when the initiator designs the new equilibrium, she must keep at least ϵ amount of reward to herself. So the remaining rewards to the other players are at most $r-\epsilon$. For simplicity, I take ϵ to zero. The result is the following tie-breaking rule:

Assumption 3. *When maximum effort is required, if the principal and the initiator promise the same rewards to a group of players, the last receiver in the communication protocol always votes no to the deviation plan.*

Denotes the principal's plan of daily reward to player j as $r_p(j|i)$ if j votes no while i is the initiator; similarly, let $r_i(j|i)$ denotes the initiator's plan of daily reward to player j for voting yes, while i is the initiator. The principal never gives the initiator transfer in the new equilibrium ($r_p(i|i)=0$ for all i) because doing so has no effect on deterring collusion but increases the probability of the initiator starting the voting stage. [11]

Then, we define the communication protocol. A communication protocol with respect to a threshold m and the initiator i is denoted $\sigma_i(m)$. It is a ordered set of $m-1$ receivers who can be reached through communication if all of them saying yes. Let $A_i(m)$ denotes the set of all such communication protocols when the initiator is i. Let $n(\sigma_i(m))$ be the set of all the neighbors of the players in $\sigma_i(m)$ (if the neighbor is in the set $\sigma_i(m)$, then this person is not counted in $n(\sigma_i(m))$)

According to the tie-breaking assumption 3, the initiator does not start the vote if the initiator needs to promise the receivers at least r units of transfers for all the communication protocols that reach $m-1$ receivers. More formally, the vote cannot be passed if for all $\sigma_i(m) \in A_i(m)$:

$$\sum_{j \in \sigma_i(m)} r_p(j|i) \geq r \tag{3}$$

This condition means the principal chooses the equilibrium selection function such that for all communication protocols, the initiator cannot incentivize $m-1$ receivers to vote for yes. Thus, he will not choose to start the vote. I call inequality 3 the blocking condition.

The principal's promised rewards also have to satisfy the feasibility constraint: under each feasible communication protocol of the initiator, the principal's promised rewards to all the rejecters given the communication protocol cannot exceed r. More formally, for each

[11] If the initiator gets a reward when the deviation proposal is rejected, the initiator needs to promise himself the same reward in each of his deviation proposals eq_i. So, the disposable amount of reward for both the principal and the initiator reduces by the same amount. So, rewarding the initiator has no effect when the full effort is required. When full effort is not required, and $eq_p(i|i)$ is smaller than the default reward to the initiator, then a higher $eq_p(i|i)$ only decreases the disposable reward of the principal. Thus, a higher $eq_p(i|i)$ makes the network less robust. If the $eq_p(i|i)$ is larger than the default reward, such a reward schedule is dominated by only rewarding the initiator in the default equilibrium and set $eq_p(i|i)=0$. Because in the former case, the principal has strictly more disposable rewards after the voting stage to deter collusion.

$m' \leq m$, and all $\sigma_i(m') \in A_i(m')$ the following inequality holds:

$$\sum_{j \in n(\sigma_i(m'))} r_p(j|i) \leq r \tag{4}$$

This condition means the principal's promised transfers to all the rejecters must be feasible given any protocols.

Lemma 3. *Let $m+1$ denote the smallest voting threshold such that there exists a set of $r_p(j|i)$ that satisfies both the blocking condition 3 and the feasibility condition 4. The negotiation power of the initiator i is equal to m.*

With the above definition, we can describe the algorithm of solving for the bargaining power of each individual i in an arbitrary supervision network.

1. Starting with $m = 2$ and initiator i. Find if there exists a vector of $r_p(j|i)$ that satisfies both the blocking condition and the feasibility condition.
2. If the solution exists, then this m is the bargaining power of j in this network. Otherwise, increase m by one and repeat this step.
3. Do this for each individual in the supervision network and derive their negotiation power. Then, take the maximum of all the negotiation power is the robustness of this network.

Here, I provide some examples of optimal reward allocation in some supervision networks. In the following figure, each arrow points from a supervisor to and supervisee. Here, we consider the case that player A is chosen to be the initiator. The letter in each circle is the name of the agent. Suppose A is the initiator, so she is labeled as a red square. The number next to each node j is the corresponding transfer chosen by the principal: $r_p(j|A)$.

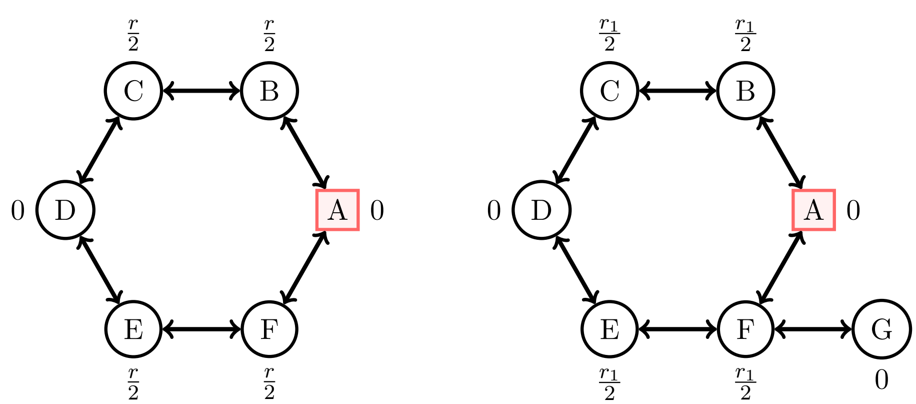

Figure 1: Single ring supervision network and modification

In the left figure, $m_A(S) = 2$, which means the initiator A can get at most two yes votes, including herself. $r = 12$ is the total transfer in this network. Players B, C, E, and F each can get a $r_p(j|A)$ of $\frac{r}{2}$ for voting no, so the initiator will not have enough reward to make two receivers vote yes votes. On the right figure, $r' = 14$. The allocation of rewards is still the same. However, now $m_A(S) = 4$ because of the additional player G. Though G is not connected to any players other than F, the principal would find no need to give G any reward for rejecting the offer. Intuitively, a player with only one connection has very weak negotiation power. If we do the calculation, we find $m_G(S) = 1$. It is achieved by $r_P(F|G) = r$

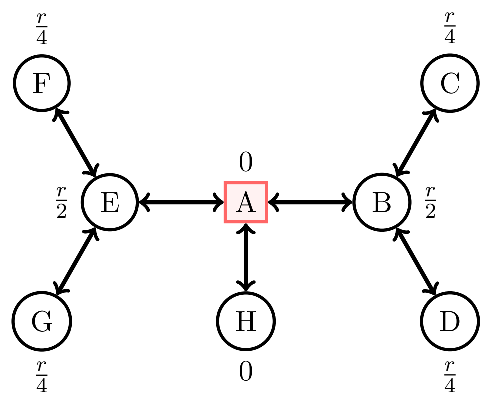

Figure 2: Single ring supervision network and modification

In figure 2, $m_A(S) = 4$ and $m(S) = 4$. A is the central player and has many connections, so she naturally has the highest negotiation power. This network is a tree, which means there is only one unique path connecting any two individuals. The feasibility constraint in a tree has special propriety: Let $r_p(j|i)$ be the principal's reward to player j for saying no. Then let $n_{sub-tree}(j)$ be the set of players who are directly connected to j in the sub-tree. We must have:

$$\sum_{k \in n_{sub-tree}(j)} r_p(k) = r_p(j|i)$$

If the players in the sub-tree get more than $r_p(j|i)$ in total, such a reward cannot be feasible. If the players in the sub-tree get less than $r_p(j|i)$ in total, then the principal can give them more rewards, making the network more robust yet still feasible.

We can repeat the above practice for all the individuals in a network, and then we can get the negotiation power for each of them $m_i(S)$, which means the number of yes votes needed to block the deviation vote from being passed. The following figure lists the negotiation power of each individual in the networks.

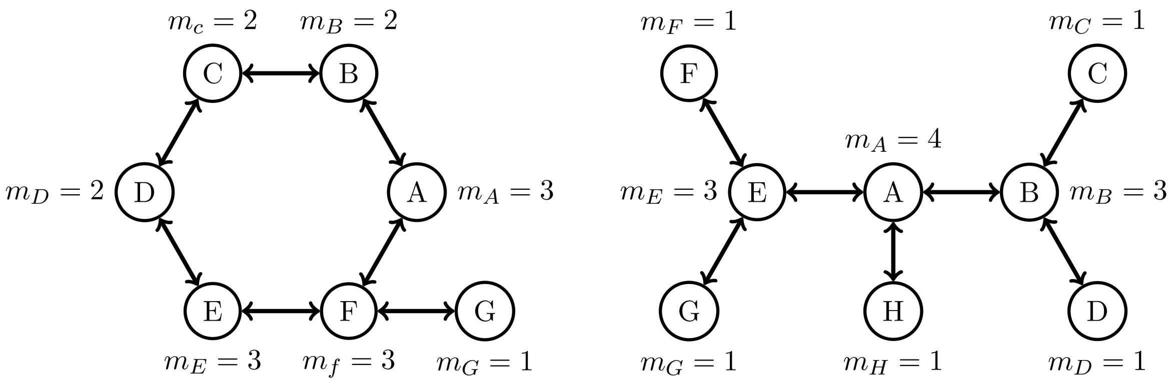

Figure 3: The number next to each node represents the player's negotiation power $m_i(S)$

3.4.2 Maximum Robustness

The previous algorithm finds the negotiation power of each individual when the full effort is required in the default equilibrium. However, the principal can further reduce the negotiation power by reducing central individuals' workload or even giving them positive transfers in the default equilibrium. However, that comes with the cost of reducing the total effort level in the default equilibrium. Thus in this section, I also develop an algorithm that calculates the most robust allocation given a supervision network.

Let $r_0(i)$ be the transfer an individual i received from other agents in the default equilibrium. It also equals the total transfers that all the players give away in the default equilibrium. Let $e_0(i)$ be the effort level of the individual i in the default equilibrium. The resource constraint requires that:

$$\sum_{i \in N} (r_0(i) + e_0(i)) \leq r \tag{5}$$

Call this inequality the anti-corruption reward constraint.[12] Recall that r is the size of the central reward pool if every player gives a daily transfer to the pool that equals the number of their supervisors. The reward $r_0(i)$ has an anti-corruption purpose because the more rewards a player receives, the less motivation she has to initiate a joint deviation. In this section, the principal wants to maximize the robustness without the full effort constraint, so the algorithm is the following:

1. For each initiator i, starting from $m = 2$, check if there exists are two vectors $\{r_0(1), r_0(2), ..., r_0(n)\}$ and $\{e_0(1), e_0(2), ..., e_0(n)\}$, such that for each initiator i, there exists a vector of $r_p(j|i)$,

[12]The name comes from a policy of the Hong Kong government. The city uses a policy called "high salary for clean government". The idea is to give government employees high wages to reduce their incentive for corruption and thus keep the government clean.

so that for each $\sigma_i(m) \in A_i(m)$, the following modified blocking condition holds.

$$\sum_{j \in \sigma_i(m)} r_p(j|i) \geq \sum_{j \neq i} d_j(S) + e_0(i) - r_0(i) \tag{6}$$

The right hand side is the disposable reward of the initiator. It is the sum of the total transfers from all the other players plus the effort he can save by deviating subtracts $r_0(i)$, which is the amount of reward the initiator needs to keep to himself. Moreover, for each $m' \leq m$, and all $\sigma_i(m') \in A_i(m')$ the following inequality feasibility constraint holds holds:

$$\sum_{j \in n(\sigma_i(m'))} r_p(j) \leq r. \tag{7}$$

2. If the solution exists, then choose the effort level in the default equilibrium to maximize $\sum_{i \in n} e_0(i)$. The robustness of the initiator i is then $m-1$
3. If the solution does not exists, then increase m by 1 and repeat the first step.

Using this algorithm, I can solve for the most robust supervision network. The label next to each node in the following graph is the $r_0(i)$ for the corresponding individual.

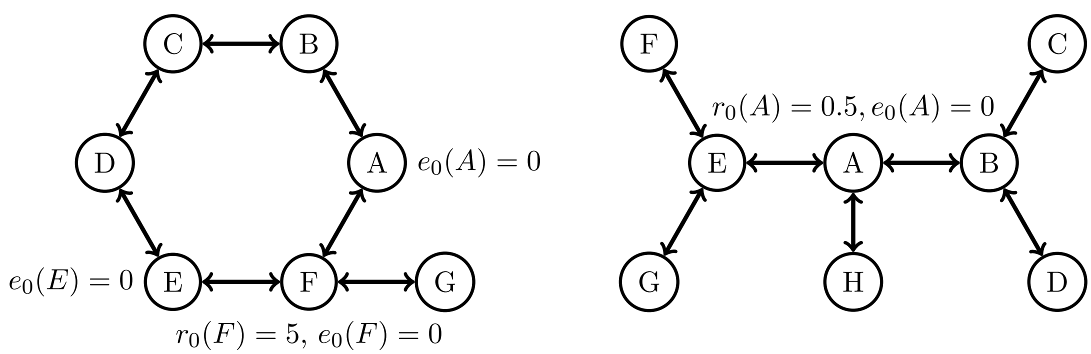

Figure 4: $r_0(i)$ for players in most robust network. The left networks $\underline{m}(S) = 3$ and the right network $\underline{m}(S) = 4$

Figure 4 indicates the $r_0(i)$ and $e_0(i)$ of each agent to maximize the robustness of the network. Players without a number on the side exert full effort and receive no transfer in the default equilibrium. Compared to the full effort cases, the left network's robustness goes from 4 to 3, and the right network's robustness goes from 5 to 4. The total effort decreases significantly to achieve a more robust outcome. In the left network, the effort loss from A, E, and F sum to 11. It means the effort level that the principal can have is $14 - 11 = 3$. On the right network, the total transfer to A is 3.5, which means the total daily effort reduces to 10.5.

4 Discussion

1. Further motivate why the principal cannot contract on the outcome or the effort.

    - There is no explicit principal in the game. Players are stuck into inefficient tradition because of the general environment, however, they are unable to coordinate away from it. This model provides an possible explanation of why they are stuck there.
    - There are laws or regulates that prevents the principal from imposing certain punishment. Such as the law prevent the employer punish the workers for organizing labor unions.
    - A large coalition can manipulate all the contractible signals and thus making the reward or punishment based on the signal ineffective. The collusion of civil servants, the police department and so on. (let the coalition be able to not only choose the equilibrium, but also manipulate the signal). It is not always the case that one non-cooperator can raise enough evidence accuse the entire department for corruption (The Tokyo Olympic scandal). Colluding agents and make up fake evidence against the non-complier, forcing them to concede.

2. Explain the advantage of setting up agents into a peer supervision than let them freely achieve the highest private optimality. See the newly added section there.

    - Risk sharing
    - Liquidity constraint.
    - Externality control.

3. Some examples of organizations actually use the peer supervision structure.

    - Turkish Coup (Why coup is hard to do. Fail because it only takes a small amount of people who refuse to join the coup to leak the information and the plot would fail. Not enough reward to compensate all the people contacted. Also, it is hard for them to recruit enough people to over power the existing government. Something like not reaching the coalition size to be exempt from external source of punishment).
    - The separation of duties for critical jobs in a firm.
    - Separation of power in the US government.
    - A managerial implication is: make it clear who should not be able to contact whom else. Periodically "restart the department" because you cannot guarantee

the initiator do not have other ways to find the contact information of the other agents.

- Second managerial implication: construct ways for small fraction of agents to reliably report collusion. So, rewarding the truth teller would be much more useful. Make faking evidence harder so it requires more people to participate.

4. Extension of future model

- There agents have possibility of like the "effort" and thus are more likely to reject the deviation upon contact (like a heroic police who is not corruptible). Such uncertainty makes it harder for the initiator to coordinate a joint deviation, and thus need to be more strategic.
- If there is uncertainty in the overall benefits that the agents can achiever. Assume the agents are risk averse. Formal discussion on the risk sharing.
- Uncertainty about the collusion success threshold. Unlike the form of labor union which has clear labor law on what counts as a successful vote, the success threshold in some other cases are much less clear. Still taking the Turkish Coup as an example. It is unclear what is the fraction of people need to join the coup for it to be a success. When the threshold is unclear, the success probability of the coalition also becomes uncertain. That makes the coalition harder to occur.
- Possibility of field experiment: whether limiting contact can really deter collusion.

There is another advantage of the peer supervision structure. When the principal cannot contract punishment or rewards based on performance or outcome, a conventional solution is to "rent" the entire department to the agents. Let them achiever their optimal outcome, and the principal charges a fixed rent or pay them fixed wage so the agents is indifferent between doing the job and choose the outside option. However, there are three disadvantages of the rental scheme. 1. Agents may be more risk averse than the principal, so letting the agents bear all the risk can be sub-optimal. 2. The agents might be liquidity constrained and thus unable to pay the rent. Thus, some beneficial departments might not be established. 3. Most importantly, when the agents acquire benefits without supervision, such activity might generate large negative externality. For instance, privatizing the police department can be a very bad idea. Suppose the policeman can fake evidence and cover one another up for their criminal activities, the principal (government) cannot contract punishment or rewards based on the evidence raised by the police department. Suppose the principal give up supervision on the police department and totally privatize it. Let u_i denotes the utility of the outside option of each individual policeman. Let π denotes the total profit of the department if they can freely abuse their power. So the government can charge a fixed rent r for the company who runs the department, such that $\sum_i u_i = \pi - r$. Thus the private police department is willing to function. Then, we can anticipate the private police department to abuse their power and do all kinds of illegal things to achieve the profit π. However, the illegal activities may generate large negative externality: for each dollar of profit the policeman acquire, the society may lose 10 dollars or more. A much better option is to let the police supervise each other and exert an effort not to abuse their power. Let e_i be the individual effort level, $\sum_i u_i = \pi - \sum_i e_i$. Still each individual policeman is willing to work. But the amount of $\sum_i e_i$ negative externality generating action can be saved. The total social welfare can thus be much higher under the peer supervision scheme than in the no supervision case. The government can also compensate the policeman by a fixed amount so higher effort level is sustainable. Such compensation is collected through less harmful channels such as tax or other state owned business. This example shows the importance of having a proper peer supervision structure.

5 Conclusion

This paper models the detailed negotiation process of equilibrium selection. The model suggests that the players might not always be able to coordinate to a Pareto optimal outcome and thus may be stuck in an inefficient equilibrium: the more sparse the communication network, the more complex the players to coordinate a joint deviation. Thus, a principal can take advantage of this finding and make a group of agents exert stable effort without direct supervision. In general, the principal can make the agents supervise each other and limit the communication among them to deter collusion. This peer supervision model has many real-world applications in firm management, political systems, and more.

This endogenous equilibrium selection model can also be applied to more general settings. We need to specify the communication network corresponding to the stage game and the criteria for joint deviation. Then, we can model each equilibrium's stability and what equilibrium the players will choose through the endogenous equilibrium selection process.

In the future, it would also be interesting to study other endogenous equilibrium selection processes, such as sub-group coalition. However, the strategy space for such a coalition is so ample that I leave it for future research.

References

Abreu, D. (1988). On the theory of infinitely repeated games with discounting. *Econometrica: Journal of the Econometric Society*, 383–396.

Abreu, D., P. K. Dutta, and L. Smith (1994). The folk theorem for repeated games: a neu condition. *Econometrica: Journal of the Econometric Society*, 939–948.

Acemoglu, D. and A. Wolitzky (2018). A theory of equality before the law. Technical report, National Bureau of Economic Research.

Akerlof, G. (1976). The economics of caste and of the rat race and other woeful tales. *The Quarterly Journal of Economics 90*(4), 599–617.

Aldashev, G. and G. Zanarone (2017). Endogenous enforcement institutions. *Journal of Development Economics 128*, 49–64.

Ali, S. N. and D. Miller (2013). Enforcing cooperation in networked societies. *Unpublished paper.[282]*.

Ali, S. N. and D. A. Miller (2016). Ostracism and forgiveness. *American Economic Review 106*(8), 2329–48.

Baliga, S., E. B. de Mesquita, and A. Wolitzky (2019). Deterrence with imperfect attribution. Technical report, Working Paper.

Baliga, S. and T. Sjöström (1998, December). Decentralization and Collusion. *Journal of Economic Theory 83*(2), 196–232.

Basu, K. (2018). *The Republic of Beliefs*. Princeton university press.

Bawn, K., M. Cohen, D. Karol, S. Masket, H. Noel, and J. Zaller (2012). A theory of political parties: Groups, policy demands and nominations in american politics. *Perspectives on Politics*, 571–597.

Bentham, J. (1791). *Panopticon: Postscript: Part II: Containing a Plan of Management for a Panopticon Penitentiary-house. By Jeremy Bentham,...*, Volume 2. T. Payne.

Bernheim, B. D., B. Peleg, and M. D. Whinston (1987). Coalition-proof nash equilibria i. concepts. *Journal of economic theory 42*(1), 1–12.

Bikhchandani, S., D. Hirshleifer, and I. Welch (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy 100*(5), 992–1026.

Chen, C. X. and T. Sandino (2012). Can Wages Buy Honesty? The Relationship Between Relative Wages and Employee Theft. *Journal of Accounting Research 50*(4), 967–1000. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1475-679X.2012.00456.x.

Ching-to Ma (1988). Unique Implementation of Incentive Contracts with Many Agents. *The Review of Economic Studies 55*(4), 555–571. Publisher: Review of Economic Studies Ltd.

Chow, C. W., J. C. Cooper, and W. S. Waller (1988). Participative Budgeting: Effects of a Truth-Inducing Pay Scheme and Information Asymmetry on Slack and Performance. *The Accounting Review 63*(1), 111–122. Publisher: American Accounting Association.

Choy, J. P. (2018). Social division with endogenous hierarchy. *The Economic Journal 128*(615), 2711–2742.

Chung, A. F.-T. and C. A. Welcome (2004). Coalition-stable equilibria in repeated games. In *Econometric Society 2004 North American Summer Meetings*, Volume 581. Econometric Society.

Deb, J. (2007). Cooperation and community responsibility: A folk theorem for repeated matching games with names. *Available at SSRN 1213102*.

Deb, J. and J. González-Díaz (2014). Enforcing social norms: Trust-building and community enforcement. Technical report, Working paper, Stanford University.

Dixit, A. K. (2011). *Lawlessness and economics: Alternative modes of governance*. Princeton University Press.

Ellison, G. (1994). Cooperation in the prisoner's dilemma with anonymous random matching. *The Review of Economic Studies 61*(3), 567–588.

Elster, J. (1989). Social norms and economic theory. *Journal of economic perspectives 3*(4), 99–117.

Farrell, J. and E. Maskin (1989). Renegotiation in repeated games. *Games and economic behavior 1*(4), 327–360.

Fudenberg, D. and E. Maskin (2009). The folk theorem in repeated games with discounting or with incomplete information. In *A Long-Run Collaboration On Long-Run Games*, pp. 209–230. World Scientific.

Gambetta, D. (1993). The sicilian mafia.

Gromb, D. and D. Martimort (2007, November). Collusion and the organization of delegated expertise. *Journal of Economic Theory 137*(1), 271–299.

Hansen, S. C. (1997). Designing Internal Controls: The Interaction between Efficiency Wages and Monitoring*. *Contemporary Accounting Research 14*(1), 129–163. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1911-3846.1997.tb00522.x.

Hörner, J., M. Morelli, and F. Squintani (2015). Mediation and peace. *The Review of Economic Studies 82*(4), 1483–1501.

Kandel, E. and E. P. Lazear (1992). Peer pressure and partnerships. *Journal of political Economy 100*(4), 801–817.

Kandori, M. (1992). Social norms and community enforcement. *The Review of Economic Studies 59*(1), 63–80.

Kranton, R. E. (1996). Reciprocal exchange: a self-sustaining system. *The American Economic Review*, 830–851.

Levine, D. K. and S. Modica (2016). Peer discipline and incentives within groups. *Journal of economic behavior & organization 123*, 19–30.

Masten, S. E. and J. Prüfer (2014). On the evolution of collective enforcement institutions: communities and courts. *The Journal of Legal Studies 43*(2), 359–400.

Moser, D. V. Michigan State University. pp. 25.

Olatunji, B. O., K. Naragon-Gainey, and K. B. Wolitzky-Taylor (2013). Specificity of rumination in anxiety and depression: A multimodal meta-analysis. *Clinical Psychology: Science and Practice 20*(3), 225–257.

Rubinstein, A. (1982). Perfect equilibrium in a bargaining model. *Econometrica 50*(1), 97–109.

Stiglitz, J. E. (1990). Peer monitoring and credit markets. *The World Bank Economic Review 4*(3), 351–366.

Webb, R. (2002, May). The impact of reputation and variance investigations on the creation of budget slack. *Accounting, Organizations and Society 27*(4-5), 361–378.

Wolitzky, A. (2013). Cooperation with network monitoring. *Review of Economic Studies 80*(1), 395–427.

Zhang, Y. (2008). The Effects of Perceived Fairness and Communication on Honesty and Collusion in a Multi-Agent Setting. *The Accounting Review 83*(4), 1125–1146. Publisher: American Accounting Association.

6 Appendix

A1. An example effort provision equilibrium of lemma 1

Here, I provide an example of effort provision equilibrium. The players may use minmax punishment to threat those who does not exert enough effort.

At the first stage (morning) of each period, for every player $i \in \mathcal{I}$, choose effort level $e_{it} = n - 1$ and zero transfer $\pi_{ijt} = 0$ for all $j \in \mathcal{I}$ and $j \neq i$. If no deviation happens at the morning, choose $\phi_{i,j} = 0$ for all $i and j \in \mathcal{I}$, in other word, no punishment if no deviation. The game repeats for all future periods. If a player i deviates at the morning of a period t, then the game changes to a punishment stage. In the second stage (afternoon) of the day t, all the other players $j \neq i$ choose $\phi_{jit} = 1$ to punish the deviater. If everyone carries out the punishment accordingly, then the game restarts on the next morning, and everyone is expected to exert effort $e_{i(t+1)} = n - 1$.

If a player $k \in \mathcal{I}$ deviates in the punishment state, then all the other players switch to punishing this player k infinitely by choosing $\phi_{j,k} = -1$ for all $j \neq k$ for all later periods. Deviation in the punishment state will result in a payoff stream of $0, -n+1, -n+1, ...,$ for the deviater, while not deviating will result in $-1, -1, -1, ...$ There is no profitable deviation in the punishment state if $\frac{\delta(n-1)}{1-\delta} \geq 1$. The inequality simplifies to $\delta > \frac{1}{n}$. Thus, if $n > 2$, the equilibrium holds when $\delta \geq \frac{1}{2}$. If there is a new deviater in the punishment stage, then all the players (including the original deviater) switch to punish the new deviater. Deviation in the punishment stage results in more severe long term punishment because the other players can only respond to such deviation in the following day, and thus the loss to the deviater is discounted.

Finally, ignore multiple simultaneous deviations. Because we are looking at the subgame perfect Nash equilibrium, so long as the strategy can deter all unilateral deviations such strategy should be a SPNE.

In this paper, we mainly focus on patient players, which means $\delta \geq 1/2$. However, the main results of this paper also holds when the patience level is even smaller. The main problem with impatient agent is that they might lack the incentive to carry out punishment. To deal with this issue, the principal can reduce the required effort level of each agents accordingly. So less number of peers are needed to punish those who fall short of effort. Consequently, those who carry out punishment can be rewarded by transfers from those who are not required to exert punishment on the previous day. If someone deviates from the punishment path, she not only lose the reward, she would also be punished by the others on the next day. Using such an equilibrium, positive effort level can be sustained with very small δ as long as n is sufficiently large.