CIRJE-F-201

Multi-Group Incentives

Hitoshi Matsushima
The University of Tokyo

March 2003

Multi-Group Incentives[+]

Hitoshi Matsushima[*]
Faculty of Economics, University of Tokyo

March 10, 2003

Abstract

This paper investigates the agency problem with moral hazard, where the principal hires multiple agents, and can only imperfectly monitor their action choices by observing their correlated public signals. The principal will penalize any detected deviant only by firing her and other agents. The key assumption of the paper is that agents are divided into multiple distinct groups. Within each group, all its members can make the binding commitments to achieve their collusive action choices. It is shown that it may be easier to provide the agents with the incentive to make the most desired action choices when multiple groups are established than when either no group or only the grand group is established. It is also shown that in terms of uniqueness, relative performance evaluation through inter-group competition will work better than that through inter-individual competition.

Keywords: Multiple Agents, Moral Hazard, Multiple Groups, Relative Performance Evaluation, Efficiency, Uniqueness.

[+] The earlier version of this paper is a part of Matsushima (1994), which was written in Japanese.

[*] Faculty of Economics, University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113, Japan. Fax: +81-3-3818-7082. e-mail: hitoshi@e.u-tokyo.ac.jp

1. Introduction

This paper investigates the agency problem in which the principal hires multiple agents. The key assumption of the paper is that agents are divided into multiple distinct groups. Within each group, all its members can make the binding commitments to achieve their collusive action choices that maximize the sum of their expected payoffs. Since their collusive action choices do not necessarily maximize the principal's welfare, the principal have to write incentive schemes carefully in order to equalize each group's collusive action choices to the most desirable choices to the principal. Otherwise, their collusive action choices may even extract the expected surplus from the principal in more severe ways than what they will independently choose in order to maximize their own self-interests.

This paper assumes that the principal has the limited capability to penalize deviating agents in the following two ways. First, the principal cannot directly observe agents' action choices. She can only imperfectly monitor their action choices by observing their respective random public signals, the distributions of which depends on their action choices. Second, the principal cannot provide agents with monetary fines and rewards. Hence, the only way of penalizing each deviant is to fire her and other agents.

When grouping is permitted, each agent will take into account any other agent who belongs to the same group. This implies that the principal can penalize any detected deviant in more powerful ways than when grouping is not permitted. Suppose that the principal writes an incentive scheme in ways that if any agent deviates and is detected, then the principal will fire not only this agent but also the other agents who belong to the same group. Then, this incentive scheme can provide any agent with the stronger incentive not to deviate than any incentive scheme that fires only detected deviants.

This paper assumes that the realizations of the public signals depend on either a random macro shock or a random private shock profile, but not on both. Which works between the macro shock and the private shock profile is randomly determined, and is unknown to the principal and the agents. Hence, the random public signals as monitoring instruments are imperfectly correlated, as far as the probability of the macro shock working is positive. This implies that relative performance evaluation through competition between groups would be a powerful tool to incentivize the agents, where agents who belong to any group that has low performance relatively to the other groups will be fired. This point is in contrast with the case where all agents are permitted to establish only the grand group, because by definition there exists no counter-group that can provide the principal with information about whether the grand group deviates or not.

The previous works such as Holmstrom (1982) have studied relative performance evaluations in the models where grouping is not permitted and any competing individual who has low performance relatively to the other individuals will either be penalized or not be rewarded.[1] The present paper will show what is the point of differences between

[1] See also Lazear and Rosen (1981), Green and Stokey (1983), and Nalebuff and Stiglitz (1983). These works assumed that agents are risk averse and showed that relative performance evaluation provides for better risk sharing than independent evaluation. In contrast, the present paper does not assume that agents are risk averse, and assumes that the principal's tools to incentivize agents are restricted. Legros and Matsushima (1991) investigated partnerships where

inter-group competition and inter-individual competition.

When grouping is not permitted, incentive schemes on the basis of relative performance evaluation have the weak point in terms of uniqueness that all agents' neglecting their duties together might be another equilibrium behavior (See Mookherjee (1984), Demski and Sappington (1984), and Ma (1988), for example). It is hard for the principal to detect any deviant when the other agents all deviate, because this deviant's performance is not necessarily low relatively to the other agents' performances in this case. Hence, the probability of any deviant's being detected is left small, which implies that relative performance evaluation does not work on agents' incentives.

When grouping is permitted, on the other hand, the principal can easily provide any group with the incentive to induce at least one agent belonging to this group to make the desired action choice whenever the other groups deviate. Suppose that the principal write an incentive scheme in ways that this agent, by making the desired action choice instead of deviating from it, can prevent not only herself but also all other members of the same group from being fired. Since this agent will take into account the welfare of any other agent who belongs to the same group, she will have the stronger incentive not to deviate than she will do when grouping is not permitted. In this way, each group is willing to provide the principal with information about the fact that the other group deviate. This is the driving force of eliminating unwanted equilibrium behaviors.

As Aoki (1990, Chapter 8) has written, it is well-accepted view that not individuals, but groups, are the robust cores at various levels of organizations in Japanese and Asian societies. In the studies of Japanese economy and society, it has long been a puzzle how to show theoretical foundations on what is the point of difference between inter-individual competition and inter-group competition. For instance, Aoki wrote: '… although the small-group values inherited from the cultural tradition have played a significant role in shaping Japanese organizational practices, in order to be efficient the J-firm (Japanese firm) had to consciously design and develop efficient intergroup coordination mechanisms and an accompanying incentive structure. Groupism is not a sufficient condition for the competitive performance of the J-firm. I argue further that it is not a necessary condition either' (Aoki, 1990, Chapter 8, pp.299). His statement above might rely on the presumption that there exist no substantial differences between inter-individual competition and inter-group competition. In contrast, the present paper will show that inter-group competition can provide for better match between agents' incentives and their welfare loss than inter-individual competition, mainly because each agent will take into account the welfare of not only herself but also the other members.

It is an implicit assumption of this paper that within each group, its members can mutually observe their action choices and can write a side contract contingent on their action choices, which is enforceable in non-judicial ways such as a 'word of honor'. See Tirole (1992) for the issue on how to specify the hidden side-contracting technology.

The previous works such as Varian (1990), Holmstrom and Milgrom (1990), Ramakrishnan and Thankor (1991), Ito (1993), and Macho-Stadler and Perez-Castrillo (1993) investigated the agency problem with multiple risk averse agents where only the grand group is permitted. These works commonly showed that grand grouping may provide for better risk sharing when the agents share information that is not held by the

---

agents are risk neutral and the principal's tools to incentivize agents are restricted in that agents' liability is limited.

principal. In contrast, this paper does not assume that agents are risk averse and assumes that the principal's instruments to incentivize agents are restricted. Although the main concern of this paper is the case that multiple groups are established, the paper will even show an alternative explanation on why grand grouping is welfare improving compared with when it is prohibited.

The organization of the paper is as follows. Section 2 shows the model. Section 3 investigates the case that no groups are established, and investigates also the case that only the grand group is established. Section 4 is the main part of the paper, which investigates the case that multiple groups are established and compete each other. Section 5 concludes.

2. The Model

We consider the following agency problem with multiple agents and a single principal. Let $N \equiv \{1,...,n\}$ denote the set of agents, where $n \geq 2$. Each agent $i \in N$ chooses an action $a_i \in A_i = \{0,1\}$. We assume that the principal prefers action 0 to action 1 for each agent. The payoff for each agent $i \in N$ does not depend on the other agents' action choices and is given by

$$u_i(a_i) = a_i.$$

Hence, the choice of action 0 for each agent provides her with the higher payoff than the choice of action 1. Let $A = \prod_{i \in N} A_i$. Agents' action choices are denoted by $a = (a_i)_{i \in N} \in A$. Let $a^0 = (a_i^0)_{i \in N} = (0,...,0)$ and $a^1 = (a_i^1)_{i \in N} = (1,...,1)$, where a^0 (a^1) is regarded as the action profile that is the most (least) desired by the principal.

The principal cannot observe the agents' action choices, but can observe a random public signal ω_i for each agent $i \in N$, the distribution of which depends on all agents' action choices $a \in A$. Let $\Omega_i = \{0,1,2\}$ denote the set of possible public signals for each agent $i \in N$. Let $\Omega = \prod_{i \in N} \Omega_i$ and $\omega = (\omega_i)_{i \in N} \in \Omega$. Let $p(\cdot \mid a) : \Omega \to [0,1]$ denote the conditional distribution over the signal profile, where $\sum_{\omega \in \Omega} p(\omega \mid a) = 1$ for all $a \in A$.

We assume that there exist two random *macro* shocks denoted by λ and θ_0, and n random *private* shocks denoted by θ_i, $i \in N$, which are independently drawn and cannot be observed by the principal and the agents. Let $\Gamma = \{0,1\}$ denote the set of possible λ. For every $h \in \{0,...,n\}$, let $\Theta_h = \{0,1\}$ denote the set of possible θ_h. We assume that when $(\lambda, \theta_0, ..., \theta_n)$ is the shock profile realized, the principal certainly observes the public signal for each agent $i \in N$ given by

(1) $$\omega_i = \lambda \max[a_i + \theta_i, 1] + (1 - \lambda)(a_i + \theta_0).$$

The implication of the specification of public signals above is the following. The realization of the public signal for each agent $i \in N$, ω_i, depends either on the second macro shock θ_0 or on the private shock θ_1 for her, but not on both. The realization of the first macro shock λ determines which actually works between the second macro shock θ_0 and the private shock profile $(\theta_1,...,\theta_n)$. When $\lambda = 0$, the second macro shock θ_0 works for all signals, and the principal observes

$$\omega_i = a_i + \theta_0 \quad \text{for all } i \in N.$$

When $\lambda = 1$, the private shock profile $(\theta_1,...,\theta_n)$ works for all signals, and the principal observes

$$\omega_i = \max[a_i + \theta_i, 1] \quad \text{for all } i \in N.$$

Note that when the second macro shock works, it holds that for every $i \in N$, and every $j \in N / \{i\}$,

$$\omega_i = \omega_j \quad \text{if and only if } a_i = a_j,$$

and

$$\omega_i = \omega_j + 1 \text{ if and only if } a_i = 1 \text{ and } a_j = 0.$$

Note that when the private shock profile works, it holds that

$$\omega_i \neq 0,$$

and

$$\omega_i \neq 2 \text{ if } a_i = 0.$$

Fix $q \in (0,1)$ arbitrarily. Let

$$\lambda = 0 \text{ with probability } q,$$

and

$$\theta_h = 0 \text{ with probability } \frac{1}{2} \text{ for all } h \in \{0,...,n\}.$$

Based on the shock profile $(\lambda, \theta_0,...,\theta_n)$ above, we will specify $p(\omega \mid a)$ as the conditional probability of the occurrence of any shock profile $(\lambda, \theta_0,...,\theta_n)$ satisfying that

$$\omega_i = \lambda \max[a_i + \theta_i, 1] + (1-\lambda)(a_i + \theta_0) \text{ for all } i \in N.$$

Before agents' choosing their actions, the principal will write an *employment scheme for each agent* $i \in N$ denoted by $x_i : \Omega \to \{0,1\}$, where "$x_i(\omega) = 1$" ("$x_i(\omega) = 0$") means that agent i will be fired (not be fired) when the principal observes the signal profile $\omega \in \Omega$. Let $x = (x_i)_{i \in N}$ denote an *employment scheme*. The loss for each agent from being fired is given by $H > 0$. The payoff for each agent $i \in N$ associated with the employment scheme x when she chooses the action a_i and the principal observes the signal profile ω is given by

$$a_i - Hx_i(\omega).$$

The purpose of the principal is to provide every agent $i \in N$ with the incentive to choose the desired action $a_i^0 = 0$ without firing any agents, i.e., with the constraint that for every $\omega \in \Omega$,

(2) $$x_i(\omega) = 0 \text{ for all } i \in N \text{ if } p(\omega \mid a^0) > 0.$$

Note that

$$p(\omega \mid a^0) > 0 \text{ if and only if } \omega_1 = \omega_i \neq 2 \text{ for all } i \in N.$$

Hence, an employment scheme x satisfies the equalities (2) if and only if for every $\omega \in \Omega$,

(3) $$x_i(\omega) = 0 \text{ for all } i \in N \text{ whenever } \omega_1 = \omega_i \neq 2 \text{ for all } i \in N.$$

3. Basic Results

This section investigates the following two situations, in which either no group or only the grand group is permitted to be established, and shows several properties on the possibility of the desired action profile a^0 being played, which will be regarded as the benchmark results of this paper.

3.1. Case I: Individual Incentives

This subsection considers the situation named *Case I*, in which there exists no possibility of agents' establishing the binding commitments to achieve their collusive action choices. Hence, in Case I, each agent chooses the action that maximizes her expected payoff associated with the employment scheme x. In addition to the equalities (2), we will require x to satisfy that for every $i \in N$,

(4) $$u_i(a_i^0) - H\sum_{\omega\in\Omega} x_i(\omega)p(\omega \mid a^0) \geq u_i(a_i^1) - H\sum_{\omega\in\Omega} x_i(\omega)p(\omega \mid a_{-i}^0, a_i^1).$$

The inequalities (4) imply that each agent $i \in N$ has incentive to choose the desired action a_i^0 when the other agents choose the desired action choices a_{-i}^0, i.e., a^0 is a Nash equilibrium in the *game associated with the employment scheme* x. Note that with the constraint of the equalities (2), an employment scheme x satisfies the inequalities (4) if and only if for every $i \in N$,

(5) $$H\sum_{\omega\in\Omega} x_i(\omega)p(\omega \mid a_{-i}^0, a_i^1) \geq 1.$$

The following proposition shows the necessary and sufficient condition under which there exists an employment scheme, in the game associated with which, the desired action profile a^0 is a Nash equilibrium.

Proposition 1: *There exists an employment scheme* x *that satisfies the equalities (2) and the inequalities (4) if and only if*

(6) $$H \geq \frac{2}{1+q}.$$

Proof: Note that we can choose an employment scheme x satisfying the equalities (3) and

$$x_i(\omega) = 1 \text{ whenever } \omega_i > \omega_j \text{ for all } j \in N/\{i\}.$$

Note that the specified employment scheme x above satisfies the equalities (2). Note also that any agent will be fired whenever she is the unique deviant and the principal detects her deviation.

Whenever the second macro shock works, then any single deviant will be certainly detected and fired. Whenever the private shock profile works, then any single deviant will be detected and fired only with probability half. Hence, the probability of a single deviant's being fired equals

$$\sum_{\omega\in\Omega} x_i(\omega)p(\omega \mid a_{-i}^0, a_i^1) = q + \frac{1-q}{2}.$$

This implies that x satisfies the inequalities (4) if and only if

$$H(q+\frac{1-q}{2})\geq 1\text{, i.e., } H\geq \frac{2}{1+q},$$

which equals the inequality (6).

Note that for every $i \in N$, both $p(\omega \mid a^0) = 0$ and $p(\omega \mid a^0_{-i}, a^1_i) > 0$ hold if and only if

$$\omega_i > \omega_j \text{ for all } j \in N/\{i\}.$$

This implies that the specified employment scheme above makes the severest punishment on any deviant when the other agents choose the desired actions. Hence, we have proved that the inequality (6) is not only sufficient but also necessary for the existence of employment scheme x satisfying the equalities (2) and the inequalities (4).

Q.E.D.

In addition to the equalities (2) and the inequalities (4), we will require x to satisfy that for every $a \in A/\{a^0\}$, there exists an agent $i \in N$ who has no incentive to choose a_i when the other agents conform to a_{-i}, that is,

(7) $$u_i(a_i) - H\sum_{\omega\in\Omega} x_i(\omega)p(\omega \mid a) < u_i(a'_i) - H\sum_{\omega\in\Omega} x_i(\omega)p(\omega \mid a_{-i}, a'_i) \text{ for some } i \in N\text{, where } a'_i \neq a_i.$$

This implies that there exists no (pure strategy) Nash equilibrium other than a^0 in the game associated with x. The following proposition shows the necessary and sufficient condition under which there exists an employment scheme, in the game associated with which, the desired action profile a^0 is the *unique* Nash equilibrium.

Proposition 2: *There exists an employment scheme x that satisfies the equalities (2), the inequalities (4), and the inequalities (7), if and only if*

(8) $$H > 2.$$

Proof: Note that we can choose x satisfying the inequalities (3) and that for every $i \in N$, and every $\omega \in \Omega$,

$$x_i(\omega) = 1 \text{ whenever either } \omega_i = 2\text{, or } \omega_i > \omega_j \text{ for some } j \in N/\{i\}.$$

Note that the specified employment scheme x above satisfies the inequalities (2). Note also that any agent will be fired whenever she deviates and the principal detects her deviation. Since $2 > \frac{2}{1+q}$, the inequality (8) implies the inequality (6), and therefore, it follows, in the same way as in Proposition 1, that the inequality (8) implies the inequalities (4).

The probability of any agent who chooses action 0 being fired equals zero irrespective of the other agents' action choices, i.e., for every $i \in N$, and every $a \in A$,

$$\sum_{\omega\in\Omega} x_i(\omega)p(\omega \mid a) = 0 \text{ if } a_i = a^0_i.$$

Whenever there exists an agent who chooses the desired action 0 and the second macro shock works, then any deviant will certainly be detected and fired. On the other hand, whenever there exist no such agents and the second macro shock works, then any deviant will be detected and fired only with probability half. Whenever the private shock profile works, then any deviant will be detected and fired with probability half. Hence, the probability of a deviant's being fired equals

$$\sum_{\omega\in\Omega} x_i(\omega)p(\omega \mid a^1) = \begin{cases} q + \dfrac{1-q}{2} & \text{if } a \neq a^1 \text{ and } a_i = a_i^1 \\ \dfrac{1}{2} & \text{if } a = a^1 \end{cases}.$$

Since $q + \frac{1-q}{2} \geq \frac{1}{2}$, it follows that only the inequalities (7) for $a = a^1$ are binding. Hence, x satisfies the inequalities (7) if and only if $1 - \frac{\mathrm{H}}{2} < 0$, i.e., if and only if the inequality (8) holds.

Note that for every $a \in A$, if both $p(\omega \mid a_i^0, a_{-i}) = 0$ and $p(\omega \mid a_i^1, a_{-i}) > 0$ hold, then it holds that either $\omega_i = 2$, or $\omega_i > \omega_j$ for some $j \in N/\{i\}$. Note also that for every $a \in A$, $p(\omega \mid a_i^0, a_{-i}) = 0$ holds whenever either $\omega_i = 2$, or $\omega_i > \omega_j$ for some $j \in N/\{i\}$. These imply that the specified employment scheme above makes the severest punishment on any deviant irrespective of the other agents' action choices. Hence, we have proved that the inequality (8) is not only sufficient but also necessary for the existence of the employment scheme satisfying the equalities (2), the inequalities (4), and the inequalities (7).

Q.E.D.

3.2. Case G: Grand Group Incentives

This subsection considers the situation named Case G, in which all agents establish the *grand* group that is binding in the sense that they can be committed to choose the action profile that maximizes the sum of their expected payoffs associated with the employment scheme x. Instead of the inequalities (4) and (7), we will require x to satisfy that for every $a \in A/\{a^0\}$,

(9) $$\sum_{i\in N}\{u_i(a_i^0) - H\sum_{\omega\in\Omega} x_i(\omega)p(\omega \mid a^0)\} > \sum_{i\in N}\{u_i(a_i) - H\sum_{\omega\in\Omega} x_i(\omega)p(\omega \mid a)\}.$$

The inequalities (9) imply that the desired action profile a^0 is the unique maximizer of the sum of their expected payoffs associated with the employment scheme x. Note that with the constraint of the equalities (2), an employment scheme x satisfies the inequalities (9) if and only if for every $a \in A/\{a^0\}$,

$$H \sum_{i\in N,\omega\in\Omega} x_i(\omega)p(\omega \mid a) > \sum_{i\in N} a_i.$$

The following proposition shows the necessary and sufficient condition under which there exists an employment scheme, in the game associated with which, the grand group has strict incentive to be committed to choose the desired action profile a^0.

Proposition 3: *There exists an employment scheme x that satisfies the equalities (2) and the inequalities (9) if and only if*

$$(10) \qquad H > \frac{1}{1-\frac{1}{2^n}-q(\frac{1}{2}-\frac{1}{2^n})}.$$

Proof: Note that we can choose x satisfying the equalities (3) and

$$x_i(\omega)=1 \quad \text{whenever} \quad p(\omega \mid a^0)=0.$$

Note that the specified employment scheme x above satisfies the inequalities (2). Note that all agents will be fired whenever the principal can see that there exists at least one agent who deviates. Note also that x makes the severest punishment on the grand group for deviating from a^0. Hence, all we have to do in this proof is to show the necessary and sufficient condition under which the specified employment scheme x above satisfies the inequalities (10).

Suppose that all agents deviate from the desired action profile a^0, i.e., choose the action profile a^1. Whenever the second macro shock works, then at least a single agent's deviation will be detected with probability half. Whenever the private shock profile works, then at least a single agent's deviation will be detected with probability $1-\frac{1}{2^n}$. Since all agents will be fired whenever there exists at least one detected deviant, the expected total loss for the grand group from being fired equals

$$H \sum_{i\in N,\omega\in\Omega} x_i(\omega)p(\omega \mid a^1) = Hn\{\frac{q}{2}+(1-q)(1-\frac{1}{2^n})\}.$$

Hence, x satisfies the inequality (9) for $a=a^1$ if and only if

$$(11) \qquad Hn\{\frac{q}{2}+(1-q)(1-\frac{1}{2^n})\} \geq n.$$

Next, suppose that all agents choose any action profile $a \in A/\{a^0,a^1\}$, where there exists at least one agent who chooses the desired action. Whenever the second macro shock works, then at least a single agent's deviation will be certainly detected. Whenever the private shock profile works, then at least a single agent's deviation will be detected with probability $1-\frac{1}{2^n}$. Hence, the expected total loss for the grand group from being fired equals

$$H \sum_{i\in N,\omega\in\Omega} x_i(\omega)p(\omega \mid a) = Hn\{q+(1-q)(1-\frac{1}{2^m})\},$$

where $m \in \{1,...,n-1\}$ is the number of agents i choosing $a_i = a_i^1$. Hence, it follows that x satisfies the inequalities (9) for all $a \in A/\{a^0,a^1\}$ if and only if

$$(12) \qquad Hn\{q+(1-q)(1-\frac{1}{2^m})\} \geq m \quad \text{for all} \quad m \in \{1,...,n-1\}.$$

For every $m \in \{1,...,n-1\}$, let

$$B(m) = \frac{1}{m+1}\{q+(1-q)(1-\frac{1}{2^{m+1}})\} - \frac{1}{m}\{q+(1-q)(1-\frac{1}{2^m})\}.$$

Note

$$B(m) = \frac{(1-q)(m+2) - 2^{m+1}}{2^{m+1} m(m+1)}.$$

If $B(m) \le 0$, then,

$$(1-q)(m+2) - 2^{m+1} \le 0,$$

which implies

$$B(m+1) = \frac{(1-q)(m+3) - 2^{m+2}}{2^{m+2}(m+1)(m+2)}$$
$$= \frac{(1-q)(m+2) - 2^{m+1} + 1 - q - 2^{m+1}}{2^{m+2}(m+1)(m+2)}$$
$$\le \frac{2\{(1-q)(m+2) - 2^{m+1}\}}{2^{m+2}(m+1)(m+2)}$$
$$\le 0.$$

Hence, it follows that for every $m \in \{1,...,n-1\}$, and every $m' \in \{m,...,n-1\}$,

$$B(m') \le 0 \text{ if } B(m) \le 0.$$

This implies that the inequalities (12) hold for all $m \in \{1,...,n-1\}$ if and only if the inequalities (12) hold for $m = 1$ and $m = n$, i.e., if and only if

$$\frac{1}{2}(1+q) \ge \frac{1}{Hn} \text{ and } \frac{1}{n}\{q + (1-q)(1 - \frac{1}{2^n})\} \ge \frac{1}{Hn}.$$

The latter inequality is implied by the inequality (11). The former inequality is implied by the inequality (11), because

$$\frac{1}{n}\{\frac{q}{2} + (1-q)(1 - \frac{1}{2^n})\} < \frac{1}{2}(q+1).$$

Hence, it follows that the inequality (11) is the necessary and sufficient condition under which x satisfies the inequalities (9). Since the inequality (11) equals the inequality (10), we have proved Proposition 3.

Q.E.D.

3.3. Discussion

Note

$$2 > \frac{1}{1 - \frac{1}{2^n} - q(\frac{1}{2} - \frac{1}{2^n})}.$$

This implies that the inequality (8) is more restrictive than the inequality (10), and therefore, Case G is easier to implement the desired action profile a^0 as the unique maximizer of the sum of all agents' expected payoffs than Case I to implement it as the unique Nash equilibrium. In Case G, each agent will take into account the loss for the other agents from being fired, while in Case I, she will not do so. Hence, in Case G, the principal can strengthen each agent's incentive not to deviate from the desired action, by firing not only her but also the other agents whenever her deviation is detected. In Case I, on the other hand, the principal cannot do this way.

As the number of agents n increases, the right hand side of (10) approaches $\frac{2}{2-q}$. Hence, in the limit of the increase of n, the necessary and sufficient condition for Case G to implement a^0 equals

(13) $$H \geq \frac{2}{2-q}.$$

Since

$$\frac{2}{1+q} \geq \frac{2}{2-q} \text{ if and only if } q \leq \frac{1}{2},$$

it follows that if the probability of the second macro shock working is less than half, i.e., $q < \frac{1}{2}$ (if the probability of the private shocks working is less than half, i.e., $q > \frac{1}{2}$), then, in this limit, Case G is easier (more difficult) to implement a^0 than Case I to implement it as a Nash equilibrium that is not necessarily unique.

Suppose that the second macro shock works. Then, in Case G, the grand group's overall deviation from a^0 to a^1 will be detected only with probability half. In Case I, on the other hand, at most a single agent will deviate at one time and she will certainly be detected. This implies that whenever the second macro shock works, then Case G is more difficult to implement a^0 than Case I.

Next, suppose that the private shock profile works. Then, in Case G, the grand group's overall deviation will (almost) certainly be detected because of the Law of Large Numbers. A single agent's deviation can be detected only with probability half in both cases, but only in Case G any agent will take into account the loss for the other agents from being fired. These observations imply that whenever the private shock profile works, then Case G is easier to implement a^0 than Case I to implement it as a Nash equilibrium.

The idea of firing not only the deviant but also the other agents the welfare of whom the deviant will take into account is related to the idea of 'sphere of influence' originated in Bernheim and Whinston (1990) in the context of repeated oligopoly with multimarket contact. Matsushima (2001) is more closely related to the present paper from technical aspects, which investigated the imperfect monitoring case of repeated oligopoly with multimarket contact a la Bernheim and Whinston, and showed, by using the Law of Large Numbers in the same way, the efficiency result in the limit of the increase of the number of distinct markets that firms encounter each other.

4. Case MG: Multi-Group Incentives

This section considers the situation named Case MG, in which the agents are divided into two distinct groups, i.e., group $A = \{1,...,\hat{m}\}$ and group $B = \{\hat{m}+1,...,n\}$, where we assume $n \geq 4$, and that the size of group A is not larger than that of group B, i.e.,

$$2 \leq \hat{m} \leq n - \hat{m}.$$

All agents belonging to each group can make the binding commitment to choose their actions that maximize the sum of their expected payoffs associated with the employment scheme x.

4.2. Desired Action Profile As An Equilibrium

Instead of the inequalities (4) and (9), we will require x to satisfy that for every $C \in \{A, B\}$, and every $a_C = (a_i)_{i \in C} \in A_C = \prod_{i \in C} A_i$,

$$(14) \qquad \sum_{i \in C} \{u_i(a_i^0) - H \sum_{\omega \in \Omega} x_i(\omega) p(\omega \mid a^0)\} \geq \sum_{i \in C} \{u_i(a_i) - H \sum_{\omega \in \Omega} x_i(\omega) p(\omega \mid a_{N/C}^0, a_C)\}.$$

The inequalities (14) imply that each group $C \in \{A, B\}$ has incentive to choose their desired actions a_C^0 whenever the other group N/C choose their desired actions $a_{N/C}^0$, i.e., $a^0 = (a_A^0, a_B^0)$ is a Nash equilibrium associated with x when the two groups A and B are regarded as players. Note that with the constraint of the inequalities (2), an employment scheme x satisfies the inequalities (14) if and only if for every $C \in \{A, B\}$, and every $a_C \in A_C / \{a_C^0\}$,

$$(15) \qquad H \sum_{i \in C, \omega \in \Omega} x_i(\omega) p(\omega \mid a_{N/C}^0, a_C) \geq \sum_{i \in C} a_i.$$

The following theorem shows the necessary and sufficient condition under which $a^0 = (a_A^0, a_B^0)$ is a Nash equilibrium associated with x in the above sense.

Theorem 4: *There exists an employment scheme x that satisfies the equalities (2) and the inequalities (14) if and only if*

$$(16) \qquad H \geq \frac{1}{q + (1-q)(1 - \frac{1}{2^{\hat{m}}})}.$$

Proof: Note that we can choose x satisfying the equalities (3) and that for every $C \in \{A, B\}$, every $i \in C$, and every $\omega \in \Omega$,

$$x_i(\omega) = 1 \text{ whenever } \omega_h > \omega_j \text{ for some } h \in C \text{ and all } j \in N/C.$$

Note that it satisfies the equalities (2). Note that all members of any group will be fired whenever the principal can see that there exists at least one member of this group who deviates and all members of the other group do not deviate. Whenever the second macro shock works, then any group's deviation will certainly be detected. Whenever the private shock profile works, then any group's deviation will be detected only with probability $1-\frac{1}{2^m}$, where m is the number of the agents i belonging to the group who deviate. Since all agents belonging to the same group will be fired whenever there exists at least one detected deviant who belongs to this group, it follows that for every $C \in \{A, B\}$, and every $i \in C$, the probability of agent i being fired when group C chooses $a_C \in A_C / \{a_C^0\}$ equals

$$\sum_{\omega \in \Omega} x_i(\omega) p(\omega \mid a_{N/C}^0, a_C) = q + (1-q)(1-\frac{1}{2^m}),$$

where m is the number of agents $j \in C$ choosing $a_j = a_j^1$. Hence, it satisfies the inequalities (14) if and only if for every $m \in \{1,\dots,n-\hat{m}\}$,

$$H(n-\hat{m})\{q+(1-q)(1-\frac{1}{2^m})\} \geq m,$$

and for every $m \in \{1,\dots,\hat{m}\}$,

$$H\hat{m}\{q+(1-q)(1-\frac{1}{2^m})\} \geq m.$$

In the same way as in the proof of Proposition 3, it follows from the inequality $\hat{m} \leq n-\hat{m}$ that the above inequalities hold if and only if

$$H\hat{m}(q+\frac{1-q}{2}) \geq 1 \quad \text{and} \quad H\{q+(1-q)(1-\frac{1}{2^{\hat{m}}})\} \geq 1.$$

Since $\hat{m} \geq 2$, it follows that the first inequality holds if $H \geq 1$. Since the latter inequality implies $H \geq 1$, it follows that only the latter inequality is binding. Note that for every $C \in \{A, B\}$, and every $a_C \in A_C / \{a_C^0\}$, both $p(\omega \mid a^0) = 0$ and $p(\omega \mid a_{N/C}^0, a_C^1) > 0$ hold if and only if

$$\omega_h > \omega_j \quad \text{for some } h \in C \text{ and all } j \in N/C.$$

This implies that with the restriction of the equalities (2), the specified employment scheme x above makes the severest punishment on any deviating group when the other group chooses their desired actions. Hence, we have proved this theorem.

Q.E.D.

Note

$$\frac{2}{1+q} > \frac{1}{q+(1-q)(1-\frac{1}{2^{\hat{m}}})}.$$

This implies that the inequality (6) is more restrictive than the inequality (16), and therefore, Case MG is easier to implement a^0 as a Nash equilibrium in the above sense than Case I to implement it as a Nash equilibrium in the standard sense. In Case MG, each agent will take into account the loss for any other member of the same group from being fired. Hence, in Case MG, especially when the private shock profile works, the principal can strengthen each agent's incentive not to deviate by firing not only her but also the other agents belonging to the same group as her. In Case I, on the other hand, the principal can not do this way. Moreover, in Case MG, whenever the macro shock works, then any single group's deviation will certainly be detected, because the other group making the desired action choices can inform the principal that the group deviates, by showing higher signals. These observations imply that Case MG is easier to implement a^0 than Case I irrespective of the probability $q \in (0,1)$ of the second macro shock working.

As the number $\hat{m}$ of the agents belonging to group A increases, the right hand side of the inequality (16) approaches 1. Hence, in the limit of the increase of $\hat{m}$, the necessary and sufficient condition for Case MG to implement a^0 as a Nash equilibrium in the above sense equals

(17) $\quad H \geq 1$.

Since $\dfrac{2}{2-q} > 1$, it follows that, in this limit, the inequality (13) is more restrictive than the inequality (17), and therefore, Case MG is easier to implement a^0 as a Nash equilibrium in the above sense than Case G to implement it as one of the maximizers of the sum of all agents' expected payoffs. Not only in Case G but also in Case MG, whenever the private shock profile works, then any single group's overall deviation will almost certainly be detected. Whenever the macro shock works, then the grand group's overall deviation will be detected only with probability half in Case G. In Case MG, however, whenever the macro shock works, then any single group's deviation will certainly be detected, because the other group can inform the principal that the second macro shock works and equals zero. These observations imply that Case MG is easier than Case G in the limit.

4.2. Uniqueness

In addition to the equalities (2) and the inequalities (14), we will require x to satisfy that for every $a \in A/\{a^0\}$, there exist $C \in \{A, B\}$ and every $a'_C \in A_C/\{a_C\}$ such that

(18)
$$\sum_{i \in C} \{u_i(a_i) - H \sum_{\omega \in \Omega} x_i(\omega) p(\omega \mid a)\}$$
$$< \sum_{i \in C} \{u_i(a'_i) - H \sum_{\omega \in \Omega} x_i(\omega) p(\omega \mid a_{N/C}, a'_C)\}.$$

The inequalities (18) imply that there exists no (pure strategy) Nash equilibrium other than a^0 when the two groups A and B are regarded as players. The following theorem shows a sufficient condition under which the desired action profile is the *unique* Nash equilibrium in the above sense.

Theorem 5: *There exists an employment scheme x that satisfies the equalities (2), the inequalities (14), and the inequalities (18), if the inequalities (16) hold with strict inequality and*

(19)
$$H \geq \frac{2}{(n-\hat{m})\{q + (1-q)\dfrac{1}{2^{n-\hat{m}-1}}\}}.$$

Proof: For every $i \in A$, let

$$x_i(\omega) = 0 \;\text{ if }\; \omega_h \neq 2 \;\text{ and }\; \omega_h \leq \omega_j \;\text{ for all }\; h \in A \;\text{ and all }\; j \in B,$$

and

$$x_i(\omega) = 1 \quad \text{otherwise.}$$

For every $i \in B$, let

$$x_i(\omega) = 0 \ \text{ if } \ \omega_h \neq 2 \ \text{ and } \ \omega_h \leq \omega_j \ \text{ for all } \ h \in B \ \text{ and all } \ j \in A,$$

$$x_i(\omega) = 1 \ \text{ if } \ \omega_h > \omega_j \ \text{ for some } \ h \in B \ \text{ and some } \ j \in A,$$

and

(20) $$x_i(\omega) = 0 \ \text{ if } \ \omega_h < \omega_j \ \text{ for some } \ h \in B \ \text{ and all } \ j \in A.$$

Note that the specified employment scheme x above satisfies the equalities (3), and therefore, the equalities (2). In the same way as in Theorem 4, it follows that the inequality (16) implies the inequalities (14). From the specification (20), it follows that group B will not be fired whenever the principal can see that all members of group A deviate and at least one member of group B chooses the desired action, even though there exist detected deviant who belong to group B. On the other hand, group A will be fired whenever it deviates and the principal detects its deviation.

Whenever all agents choose the undesired action profile a^1, then the expected total loss for group B from being fired equals

$$H \sum_{i \in B, \omega \in \Omega} x_i(\omega) p(\omega \mid a^1) = H(n - \hat{m})\{\frac{q}{2} + (1-q)(1 - \frac{1}{2^{n-\hat{m}}})\},$$

where it must be noted that group $B's$ deviation will be detected only with probability half when the second macro shock works. Consider any action choices (a_A^1, a_B), according to which, all members except one of group B choose the undesired actions. The expected total loss for group B from being fired equals

$$H \sum_{i \in B, \omega \in \Omega} x_i(\omega) p(\omega \mid a_A^1, a_B) = H(n - \hat{m})(1-q)(1 - \frac{1}{2^{n-\hat{m}-1}}),$$

where it must be noted from the specification (20) that whenever the second macro shock works, then group B will never be fired in this case. Hence, from the inequalities (19), it follows that

$$\begin{aligned}
&\sum_{i \in B} \{u_i(a_i^1) - H \sum_{\omega \in \Omega} x_i(\omega) p(\omega \mid a^1)\} \\
&- \sum_{i \in B} \{u_i(a_i) - H \sum_{\omega \in \Omega} x_i(\omega) p(\omega \mid a_A^1, a_B)\} \\
&= n - \hat{m} - H \sum_{i \in B, \omega \in \Omega} x_i(\omega) p(\omega \mid a^1) - \{n - \hat{m} - 1 - H \sum_{i \in B, \omega \in \Omega} x_i(\omega) p(\omega \mid a_A^1, a_B)\} \\
&= (n - \hat{m})\{1 - H(\frac{q}{2} + (1-q)(1 - \frac{1}{2^{n-\hat{m}}}))\} \\
&- \{n - \hat{m} - 1 - (n - \hat{m}) H (1-q)(1 - \frac{1}{2^{n-\hat{m}-1}})\} \\
&= 1 - (n - \hat{m}) H (\frac{q}{2} + (1-q)(\frac{1}{2^{n-\hat{m}}})\} < 0,
\end{aligned}$$

which implies the inequality (18) for $a = a^1$.

Fix $a \in A / \{a^0\}$ arbitrarily. Suppose $a_A \neq a_A^0$ and $a_B \neq a_B^1$. Let m denote the number of agents $i \in A$ choosing $a_i = a_i^1$. Whenever the second macro shock works, then group A will certainly be fired. Hence, the expected total loss for group A from being fired equals

$$H \sum_{i\in A,\omega\in\Omega} x_i(\omega)p(\omega\,|\,a) = H\hat{m}\{q+(1-q)(1-\frac{1}{2^m})\}.$$

Whenever group A chooses the desired action choices a_A^0, then it will never be fired irrespective of group $B's$ action choices. Hence,

$$\begin{aligned}
&\sum_{i\in A}\{u_i(a_i) - H\sum_{\omega\in\Omega} x_i(\omega)p(\omega\,|\,a)\}\\
&-\sum_{i\in A}\{u_i(a_i^0) - H\sum_{\omega\in\Omega} x_i(\omega)p(\omega\,|\,a_A^0, a_B)\}\\
&= m - H\sum_{i\in A,\omega\in\Omega} x_i(\omega)p(\omega\,|\,a) + H\sum_{i\in A,\omega\in\Omega} x_i(\omega)p(\omega\,|\,a_A^0, a_B)\}\\
&= m - \hat{m}H(q+(1-q)(1-\frac{1}{2^m})\}.
\end{aligned}$$

In the same way as the proof of Proposition 3, it follows that

(21) $$\frac{1}{m}\{q+(1-q)(1-\frac{1}{2^m})\} > \frac{1}{H\hat{m}} \quad \text{for all } m\in\{1,...,\hat{m}\},$$

if and only if this inequality holds for each $m\in\{1,\hat{m}\}$, i.e., if and only if

$$H > \frac{2}{\hat{m}(1+q)} \quad \text{and} \quad H > \frac{1}{q+(1-q)(1-\frac{1}{2^{\hat{m}}})}.$$

Since $\hat{m}\geq 2$, the former inequality automatically holds. Hence, the latter inequality, which equals the inequality (16) with strict inequality, implies the inequalities (18) for all $a\in A$ satisfying $a_A \neq a_A^0$ and $a_B \neq a_B^1$.

Next, suppose $a_A = a_A^0$. Let m denote the number of agents $i\in B$ choosing $a_i = a_i^1$. Since any signal profile satisfying the specification (20) never takes place in this case, it follows that whenever the second macro shock works, then group B will certainly be fired. Hence, the expected total loss for group B from being fired equals

$$H \sum_{i\in B,\omega\in\Omega} x_i(\omega)p(\omega\,|\,a) = H(n-\hat{m})\{q+(1-q)(1-\frac{1}{2^m})\}.$$

Whenever group B chooses the desired action choices a_B^0, then it will never be fired irrespective of group $A's$ action choices. Based on these observations, and in the same way as the above, it follows from the inequalities (16) with strict inequality that

$$\begin{aligned}
&\sum_{i\in B}\{u_i(a_i) - H\sum_{\omega\in\Omega} x_i(\omega)p(\omega\,|\,a)\}\\
&-\sum_{i\in B}\{u_i(a_i^0) - H\sum_{\omega\in\Omega} x_i(\omega)p(\omega\,|\,a^0)\}\\
&= m - H\sum_{i\in B,\omega\in\Omega} x_i(\omega)p(\omega\,|\,a) + H\sum_{i\in B,\omega\in\Omega} x_i(\omega)p(\omega\,|\,a^0)\}\\
&= m - (n-\hat{m})H\{q+(1-q)(1-\frac{1}{2^m})\} < 0.
\end{aligned}$$

Hence, we have proved the inequalities (18) for all $a\in A/\{a^0\}$.

Q.E.D.

Note

$$2 > \frac{1}{q + (1-q)(1 - \frac{1}{2^{\hat{m}}})} \quad \text{and} \quad 2 > \frac{2}{(n-\hat{m})\{q + (1-q)\frac{1}{2^{n-\hat{m}-1}}\}}.$$

These inequalities imply that both the inequalities (16) and (19) are less restrictive than the inequality (8). Hence, Case MG is easier to implement the desired action profile a^0 as the unique Nash equilibrium in the above sense than Case I to implement it as the unique Nash equilibrium in the standard sense.

In Case MG, whenever all members of group B deviate, then at least one member of group A has incentive to choose the desired action, which informs the principal that group B deviate when the second macro shock works. The specification (20) implies that by choosing the desired action, any single agent belonging to group B can prevent the other members of group B from being fired, even though they deviate and are detected by the principal. In Case MG, any agent will take into account the loss for all other members of the same group as her from being fired. In Case I, one other hand, no agents will do so. This makes the incentive for each member of group B to choose the desired action stronger in Case MG than in Case I. This is the driving force of the fact that it is easier to eliminate unwanted equilibria in Case MG than in Case I.

As the number $\hat{m}$ of members of group A increases, the number $n-\hat{m}$ of members of group B as well as the number n of all agents diverges into infinity. Hence, as $\hat{m}$ increases, the right hand side of the inequality (19) approaches 0. In the limit of the increase of $\hat{m}$, the inequality (19) becomes trivial and only the inequality (16) with strict inequality becomes binding. This implies that in this limit, the sufficient condition for Case MG to implement the desired action profile a^0 as the unique Nash equilibrium in the above sense equals

$$H > 1.$$

Since $\frac{2}{2-q} > 1$, it follows that Case MG is easier to implement a^0 as the unique Nash equilibrium in the above sense than Case G to implement it as the unique maximizer of the sum of all agents' expected payoffs.

In Case MG, whenever the private shock profile works, then any group's overall deviation will certainly be detected because of the Law of Large Numbers. Even when the second macro shock works, group $A's$ overall deviation will certainly be detected, because the specification (20) implies that at least one member of group B has incentive to inform the principal that group A deviates, by choosing the desired action. In Case G, however, this incentive device does not work, because any member of group B will take into account the loss for, not only the other members of group B, but also all members of group A, from being fired. These observations imply that it is easier to eliminate unwanted equilibria in Case MG than in Case G.

5. Concluding Remarks

This paper has investigated the agency problem with multiple agents, where the principal could only imperfectly monitor their action choices by observing their correlated public signals, and could penalize any detected deviant only by firing her and other agents. We have assumed that agents were divided into multiple distinct groups, and that within each group, all its members could make the binding commitments to achieve their collusive action choices. We have shown that it was easier to incentivize the agents to make the desired action choices when multiple groups were established than when either no group or only the grand group was established. Although we have investigated only the special cases with signal correlations, most results of this paper might hold with minor modifications in a more general class of environments with signal correlations.

This paper has investigated only the special cases that at most two distinct groups were established. In the same way as in this paper, we can investigate the cases that three or more distinct groups are established. We can then prove that it is the easiest to incentivize the agents to make the desired action choices in the case that two groups with the same size are established among all possible cases whether multiple groups are established or not. In contrast to the other results of the paper, however, this result depends on the specialty of the model. For instance, if there exist partial macro shocks that influence only subsets of public signals, then the establishment of three or more groups might be welfare improving more than that of only two groups. Moreover, if the principal cannot fire agents but can promote a limited number of agents, then it might be the case that the principal had better not make the size of each group too large. Hence, it would be important future researches to show what is the optimal number of groups, and to show what is the optimal size of groups, in more general models.

References

Aoki, M. (1990): *Information, Incentives, and Bargaining in the Japanese Economy*, Cambridge: Cambridge University Press.

Bernheim, D. and M. Whinston (1990): "Multimarket Contact and Collusive Behavior," *Rand Journal of Economics* 21, 1-25.

Demski, J. and D. Sappington (1984): "Optimal Incentive Constraints with Multiple Agents," *Journal of Economic Theory* 33, 152-171.

Green, J. and N. Stokey (1983): "A Comparison of Tournaments and Contracts," *Journal of Political Economy* 91, 349-364.

Holmstrom, B. (1982): "Moral Hazard in Teams," *Bell Journal of Economics* 13, 324-340.

Holmstrom, B. and P. Milgrom (1990): "Regulating Trade among Agents," *Journal of Institutional and Theoretical Economics 146*, 85-105.

Ito, H. (1993): "Coalitions, Incentives, and Risk Sharing," *Journal of Economic Theory* 60, 410-27.

Lazear, E. and S. Rosen (1981): "Rank-Order Tournaments as Optimum Labor Contracts," *Journal of Political Economy* 98, 841-864.

Legros, P. and H. Matsushima (1991): "Efficiency in Partnerships," *Journal of Economic Theory* 55, 296-322.

Ma, C. T. (1988): "Unique Implementation of Incentive Contracts with Many Agents," *Review of Economic Studies* 55, 555-572.

Macho-Stadler, I. and J. D. Perez-Castrillo (1993): "Moral Hazard with Several Agents: the Gains from Cooperation," *International Journal of Industrial Organization* 11, 73-100.

Matsushima, H. (1994): "Rationality in Groupism: Labor Contracts and Corporate Groups," Discussion Paper No. 571, Institute of Socio-Economic Planning, University of Tsukuba (written in Japanese).

Matsushima, H. (2001): "Multimarket Contact, Imperfect Monitoring, and Implicit Collusion", *Journal of Economic Theory* 98, 158-178.

Mookherjee, D. (1984): "Optimal Incentive Schemes with Many Agents," *Review of Economic Studies* 51, 433-446.

Nalebuff, B. and J. Stigliz (1983): "Prizes and Incentives: Toward a General Theory of Compensation and Competition," *Bell Journal of Economics* 14, 21-43.

Ramakrishnan, R. and A. V. Thankor (1991): Cooperation versus Competition in Agency, *Journal of Law, Economics, and Organization* 7, 248-283.

Tirole, J. (1992): "Collusion and the Theory of Organizations," in *Advances in Economic Theory: Sixth World Congress*, ed. by J.-J. Laffont. Cambridge: Cambridge University Press.

Varian, H. (1990): "Monitoring Agents with Other Agents," *Journal of Institutional and Theoretical Economics* 146, 153-174.