Moral Hazard in Teams

Ram Singh

Department of Economics

September 23, 2009

Outline

1. Moral Hazard in Teams: Model
2. Unobservable Individual Output
3. First Best without Budget Breaker
4. Risk-Averse Teams

Model I

- Many agents; at least two agents
- Effort on the part of each agent affects the output;
- Effort is not observable or contractible;
- Cost of effort by an agent is private
- Risk-neutral parties

Example

- Firm as Team and Profit as Output;
- Cooperative (farm) as Team and Produce or profit as Output;
- Sales-persons as Team and sales as Outputs;
- Advocate as Team and Judicial judgement as Output

Model II

General Model: Holmstrom (1982, BJE)

- n Agents; $n \geq 2$
- $e = (e_1, ..., e_n)$
- Output $Q = (e_1, ..., e_n)$,
- $Q = \begin{cases} (q_1, ..., q_n) \in \mathcal{R}^n, & \text{or;} \\ Q \in \mathcal{R}, & . \end{cases}$
- Agents are weakly risk-averse.
- Team/partnership Contract: $\mathbf{w}(Q) = (w_1(Q), ..., w_n(Q))$ where $w_i(Q) = s_i(Q)$ is the output sharing rule such that $s_i \geq 0$. Typically, we have
$$\sum w_i(Q) = \sum s_i(Q) = Q.$$

Unobservable Individual Output I

Simple Model:

- $Q = Q(e_1, ..., e_n) \in \mathcal{R}$ is scalar deterministic output
- Q is increasing and concave; for all i, j,

$$\frac{\partial Q}{\partial e_i} > 0, \ \frac{\partial^2 Q}{\partial e_i^2} < 0, \ \frac{\partial^2 Q}{\partial e_i \partial e_j} \geq 0,$$

- Matrix of second derivatives $\mathcal{Q}_{ij}$ is *Negative Definite*
- Agent is risk neutral in wealth; $u_i(w_i, e_i) = u_i(w_i) - \psi(e_i) = w_i - \psi(e_i)$ and $\psi(e_i)$ is increasing and convex.
- $w_i(Q) = s_i(Q)$ is continuously differentiable and

$$(\forall Q)[\sum w_i(Q) = \sum s_i(Q) = Q]$$

Unobservable Individual Output II

The first best is solution to

$$\max_{e_1,\dots,e_n} \{Q(e_1,...,e_n) - \sum \psi_i(e_i)\}$$

s.t.

$$(\forall Q)[\sum w_i(Q) = Q] \tag{1}$$

Let $e_{-i} = (e_1,...,e_{i-1},e_{i+1},...,e_n)$.

Therefore, the first best effort e_i^* solves the following foc $\frac{\partial Q(e_i,e_{-i}^*)}{\partial e_i} = \psi'(e_i)$, for every $i = 1,...,n$. That is, for every $i = 1,...,n$

$$\frac{\partial Q(e_i^*,e_{-i}^*)}{\partial e_i} = \psi'(e_i^*) \tag{2}$$

Let $e^* = (e_1^*,...,e_i^*,...,e_n^*)$ solve system 2.

Is First Best Achievable? I

In SB, e is not contractible but Q is

Given $e_{-i} = (e_1, ..., e_{i-1}, e_{i+1}, ..., e_n)$, agent i solves

$$\max_{e_i}\{w_i(Q(e_i, e_{-i}) - \psi(e_i)\}.$$

Therefore, a (Nash) equilibrium is characterized by the following n equations

$$\frac{dw_i(Q(e_i, e_{-i}))}{dQ}\frac{\partial Q(e_i, e_{-i})}{\partial e_i} = \psi'(e_i), \tag{3}$$

for every $i = 1, ..., n$.

Now $e^* = (e_1^*, ..., e_i^*, ..., e_n^*)$ can solve (2) iff for every, we have

$$(\forall i \in \{1, ..., n\})[\frac{dw_i(Q(e_i^*, e_{-i}^*))}{dQ} = 1] \tag{4}$$

Is First Best Achievable? II

But from 1, we have

$$\sum \frac{dw_i(Q(e_i^*, e_{-i}^*))}{dQ} = 1 \tag{5}$$

4 and 5 give us a contradiction.

First Best with Budget Breaker I

Consider the following contract:

- BB demands an upfront payment of z_i and appropriate the output; and
- pays $(\forall Q)[w_i(Q) = Q]$ to each agent

Under this contract it is easy to see that

- BB and each agent is a residual claimant on the entire output; and
- $e^* = (e_1^*, ..., e_i^*, ..., e_n^*)$ is a N.E.

Is such a contract feasible?
Yes, if for all i

$$Q(e_1^*, ..., e_n^*) - \psi_i(e_i^*) \geq z_i, i.e.,$$

$$nQ(e_1^*, ..., e_n^*) - \sum \psi_i(e_i^*) \geq \sum z_i \tag{6}$$

and

$$\sum z_i + Q(e_1^*, ..., e_n^*) \geq nQ(e_1^*, ..., e_n^*) \tag{7}$$

First Best with Budget Breaker II

That is, if

$$Q(e_1^*, ..., e_n^*) - \sum \psi_i(e_i^*) > 0,$$

which is clearly true.

Is $e^* = (e_1^*, ..., e_i^*, ..., e_n^*)$ a unique N.E.?
Suppose $e_{-i} = (0, ..., 0)$. Agent i solves

$$\max_{e_i}\{Q(0, .., e_i, ..., 0) - \psi(e_i)\}.$$

Assuming $\frac{\partial Q(0,..,0,...,0)}{\partial e_i} > \frac{\partial \psi(0)}{\partial e_i}$, the agent i will choose a positive effort. Now $\frac{\partial^2 Q}{\partial e_i \partial e_j} \geq 0$ implies that other agents will also increase their effort. If $e^* = (e_1^*, ..., e_i^*, ..., e_n^*)$ a unique optimizer, iteration will continue till they reach e^*.

First Best without BB I

Consider the following 'Mirrlees' Contract:

$$w_i(Q) = \begin{cases} b_i \geq 0, & \text{if } \; Q = Q^*; \\ -k_i, & \text{if } \; Q \neq Q^*. \end{cases} \quad \text{where } b_i \geq 0 \text{ and } -k_i < 0$$

- BB pays b_i if output $Q = Q^*$, where $b_i \geq \psi_i(e_i^*)$; and
- imposes penalty of k_i if $Q \neq Q^*$
- Can choose $\sum b_i = Q^*$
- Do not need external intervention in equilibrium

First Best without BB II

Under this contract it is easy to see that $e^* = (e_1^*, ..., e_i^*, ..., e_n^*)$ is a N.E.

Multiple equilibria:
Let $\hat{e}_i$ solve

$$Q(0, ..., \hat{e}_i, ..., 0) = Q(e_1^*, ..., e_i^*, ..., e_n^*)$$

Now if

$$b_i - \psi_i(\hat{e}_i) \leq -k_i \tag{8}$$

holds $(0, ..., 0)$ is a N.E. If for some i,

$$b_i - \psi_i(\hat{e}_i) > -k_i \tag{9}$$

there exist N.E. $(\tilde{e}_1, ..., \tilde{e}_i, ..., \tilde{e}_n)$ such that for some j, $\tilde{e}_j < e_j^*$

Problematic Features I

Remark

Under Holmstrom scheme, the payoff of the BB is

$$w_{BB} = \sum z_i + Q(e) - \sum Q(e) = \sum z_i - (n-1)Q(e), i.e.,$$

$$\frac{dw_{BB}}{dQ} = -(n-1) < 0$$

Remark

- *Note the results do not depend on output being stochastic or*
- *Risk aversion of agents*

- BB want the scheme to 'fail'
- BB may collude with one of the agents

Problematic Features II

- A side contract between BB and an agent gives back original problem
- Agents may collude to borrow Q^* and game with BB

Deterministic Output and Finite Effort Space I

Legros and Matthews (1993)
Let

- Three agents, $i = 1, 2, 3$
- $Q = Q(e_1, e_2, e_3)$
- $e_i \in \{0, 1\}$, $i = 1, 2, 3$
- $\psi_i(e_i) = \psi_i(1) > \psi_i(0) > 0$, $i = 1, 2, 3$
- The FB solves $\max\{Q(e_1, e_2, e_3) - \sum \psi_i(e_i)\}$
- Let $(e_1^*, e_2^*, e_3^*) = (1, 1, 1)$
- $Q_i = Q(0, e_{-i})$, where $e_{-i} = (1, 1)$
- $Q_1 \neq Q_2 \neq Q_3$, a generic feature

Deterministic Output and Finite Effort Space II

Consider the following contract

$$w_i(Q) = \begin{cases} w_i^* & \text{if } Q = Q^*; \\ \frac{Q}{2} + \delta & \text{if } Q \neq Q^* \ \& \ Q \neq Q_{-i}; \\ -k_i, & \text{if } Q = Q_i. \end{cases}$$ where $w_i^* = w_i(Q^*) - \psi_i > 0$ and $\delta \geq 0$.

This contract implements the FB. However,

if $Q_1 = Q_2 = Q_3$ the FB cannot be implemented.

Approximating FB with Deterministic Output I

Legros and Matthews (1993) Let

- Two agents, $i = 1, 2$
- $Q = Q(e_1, e_2) = e_1 + e_2$
- $e_i \in [0, +\infty)$, $i = 1, 2$
- $\psi_i = \psi_i(e_i) = \frac{e_i^2}{2}$, $i = 1, 2$
- The FB solves

$$\max\{Q(e_1, e_2)_{e_1, e_2} - \sum \psi_i(e_i)\} = \max_{e_1, e_2}\{e_1 + e_2 - \frac{e_1^2}{2} - \frac{e_2^2}{2}\}$$

- Clearly $(e_1^*, e_2^*) = (1, 1)$

Consider the following contract

Approximating FB with Deterministic Output II

- If $Q \geq 1$
$$\begin{cases} w_1(Q) = & \frac{(Q-1)^2}{2} \text{ and ;} \\ w_2(Q) = & Q - w_1(Q). \end{cases}$$
- If $Q < 1$ $\begin{cases} w_1(Q) = & Q + k \text{ and ;} \\ w_2(Q) = & -k. \end{cases}$

Proposition

Under the above contract if agent acts as 'principal', then $((\epsilon, 1-\epsilon), (0, 1))$ is a N.E. in which the first agent plays $e = 0$ and $e = 1$ with probability ϵ and $1-\epsilon$, respectively; and agent two plays $e = 1$ with probability one.

Proof: Given $e_2 = 1$ opted by 2, agent 1 solves,

$$\max_{e_1}\{w_1(e_1+1) - \frac{e_1^2}{2}\} = \max_{e_1}\{\frac{e_1^2}{2} - \frac{e_1^2}{2}\} = 0$$

Approximating FB with Deterministic Output III

i.e., all effort levels are equally good. So, $(\epsilon, 1-\epsilon)$ is a best response for agent 1. Note agent 2 will never opt for $e_2 > 1$. Given that agent 1 opts for $(\epsilon, 1-\epsilon)$, a choice of $e_2 = 1$ gives agent 2,

$$(1-\epsilon)[2-\frac{1}{2}] + \epsilon[1-0] - \frac{1}{2} = 1 - \frac{\epsilon}{2}.$$

In contrast, when $e_2 < 1$ agent 2's payoff is,

$$(1-\epsilon)[1+e_2-\frac{e_2^2}{2}] - \epsilon k - \frac{e_2^2}{2} \leq 1 + e_2 - e_2^2 - \epsilon k,$$

which is uniquely maximized at $e_2 = \frac{1}{2}$. At $e_2 \frac{1}{2}$, agent 2's payoff is

$$\frac{5}{4} - \epsilon k$$

$e_2 = 1$ is the best response for 2, if

$$k \geq \frac{1}{2} + \frac{1}{4\epsilon}$$

Risk-Averse Team I

- $Q = Q(e_1, ..., e_n) \in \mathcal{R}$ is scalar deterministic output
- Q is increasing and concave; for all i, j,

$$\frac{\partial Q}{\partial e_i} > 0, \ \frac{\partial^2 Q}{\partial e_i^2} < 0, \ \frac{\partial^2 Q}{\partial e_i \partial e_j} \geq 0,$$

- Matrix of second derivatives $\mathcal{Q}_{ij}$ is *Negative Definite*
- Agents are risk-averse in wealth;

$$\tilde{u}_i(w_i, s_i(Q), e_i) = u_i(w_i, s_i(Q)) - \psi_i(e_i) = -e^{r_i s_i(Q)} - \psi_i(e_i)$$

  and $\psi_i(e_i)$ is increasing and convex.
- $(\forall Q)[\sum s_i(Q) = Q]$

Risk-Averse Team II

The First Best:

$$\max_{e_1,\ldots,e_i,\ldots,e_n;s_i}\{\sum \tilde{u}_i(s_i(Q),e_i)\}, i.e.,$$

$$\max_{e_1,\ldots,e_i,\ldots,e_n;s_i}\{\sum[u_i(s_i(Q))-\psi_i(e_i)]\}$$

s.t.

$$(\forall Q)[\sum s_i(Q)=Q]$$

Let $e^* = (e_1^*, ..., e_i^*, ..., e_n^*)$ along with a sharing scheme $s^*(Q)$ be the unique F.B. profile in this context.

Risk-Averse Team III

Remark

- *For a sharing scheme $s_i(Q)$ and a profile of efforts $(e_1, ..., e_i, ..., e_n) \neq (e_1^*, ..., e_i^*, ..., e_n^*)$, the following holds: There exists a sharing scheme $s^*(Q)$ such that*

$$(\forall i)[E(s_i^*, e_i^*) \geq E(s_i, e_i)] \tag{10}$$
$$(\exists j)[E(s_j^*, e_j^*) > E(s_j, e_j)] \tag{11}$$

- *If a sharing scheme $\hat{s}_i(Q)$ induces $e^* = (e_1^*, ..., e_i^*, ..., e_n^*)$ as a N.E., then for any sharing scheme $s_i(Q)$ that induces $(e_1, ..., e_i, ..., e_n)$, the following cannot hold*

$$(\forall i)[E(s_i, e_i) \geq E(\hat{s}_i, e_i^*)] \tag{12}$$
$$(\exists j)[E(s_j, e_j) > E(\hat{s}_j, e_j^*)] \tag{13}$$

Risk-Averse Team IV

- If a sharing contract does not induce $e^* = (e_1^*, ..., e_i^*, ..., e_n^*)$ as a N.E., it cannot be F.B.
- Therefore, a P.O. sharing scheme will necessarily induce $e^* = (e_1^*, ..., e_i^*, ..., e_n^*)$ as a N.E.

We know that if agents are risk neutral, i.e., if $u(x) = x$, then no BB sharing scheme can induce $e^* = (e_1^*, ..., e_i^*, ..., e_n^*)$ as a N.E.

Can a BB sharing scheme can induce $e^* = (e_1^*, ..., e_i^*, ..., e_n^*)$ as a N.E. if agents are risk-averse?

Risk-Averse Team V

Consider the following BB ‘Scapegoat’ sharing contract:

- If $Q = Q(e^*)$, then $s_i(Q) = b_i^*$, where b_i^*s are such that $\sum b_i^* = Q(e^*)$;
- If $Q > Q(e^*)$, then $s_i(Q) = b_i^* + \frac{Q - Q(e^*)}{n}$
- If $Q < Q(e^*)$, choose one agent j randomly and fix shares such that

$$\begin{aligned} s_j(Q) &= -w_j \\ (\forall i \neq j)\ s_i(Q) &= b_i^* + \frac{b_j^* + w_j + Q - Q(e^*)}{n-1} \end{aligned}$$

Risk-Averse Team VI

Remark

Note when $Q < Q(e^*)$,

$$\sum_{i=1}^{n} s_i(Q) = s_j(Q) + \sum_{i \neq j}^{n} s_i(Q) = -w_j + \sum_{i \neq j}^{n}[b_i^* + \frac{b_j^* + w_j + Q - Q(e^*)}{n-1}] = Q.$$

Therefore, the above contract meets the BB constraint.

Suppose, $e_{-i} = e_{-i}^*$, i.e., all agents apart from i have opted for FB effort. If i opts for e_i^*, his payoff is $u_i(b_i^*) - \psi_i(e_i^*)$. If he opts for some $e_i > e_i^*$, his payoff is

$$u_i(b_i^* + \frac{Q - Q(e^*)}{n}) - \psi_i(e_i).$$

Since e^* is P.O. profile,

$$u_i(b_i^* + \frac{Q - Q(e^*)}{n}) - \psi_i(e_i) > u_i(b_i^*) - \psi_i(e_i^*)$$

Risk-Averse Team VII

cannot hold.

Now, if i opts for some $e_i < e_i^*$, his share

$$s_i(Q) = \begin{cases} -w_i & \text{with probability } \frac{1}{n}; \\ b_i^* + z_i & \text{with probability } \frac{1-n}{n}, \end{cases}$$

where z_i is a random variable.

For each $j \neq i$, probability of $z_i = b_i + \frac{b_j + w_j + Q - Q(e^*)}{n-1}$ is $\frac{1}{n-1}$. Therefore, if i opts for some $e_i < e_i^*$, his payoff is

$$\frac{n-1}{n} E u_i(b_i^* + z_i) + \frac{1}{n} u(-w_i) - \psi_i(e_i) \tag{14}$$

$$\frac{n-1}{n} [\sum_{i \neq j}^{n} \frac{1}{n-1} u_i(b_i^* + z_i)] + \frac{1}{n} u(-w_i) - \psi_i(e_i) \tag{15}$$

Risk-Averse Team VIII

For $e_i < e_i^*$, agent i's payoff function is concave. Let $\hat{e}_i$ uniquely solve in region $e_i < e_i^*$. Now let

$$Y_i = u_i(b_i^*) - \psi_i(e_i^*) - [\frac{n-1}{n} E u_i(b_i^* + z_i) + \frac{1}{n} u(-w_i)] - \psi_i(\hat{e}_i) \tag{16}$$

Clearly, if $Y_i > 0$, e_i^* is a unique best response for agent i.
Now, using envelop theorem

$$\frac{dY_i}{dw_i} = \frac{1}{n} u_i' > 0 \tag{17}$$

Moreover, concavity of u_i implies

$$\frac{d^2 Y_i}{dw_i^2} = -\frac{1}{n} u''_i > 0 \tag{18}$$

Risk-Averse Team IX

That is Y_i is increasing in w_i at an increasing rate. So, there exits $\bar{w}_i$ such that for all $w_i \geq \bar{w}_i$, $Y_i > 0$. That is, for all $w_i \geq \bar{w}_i$, e_i^* is a unique best response to e_{-i}^*. Therefore,

Proposition

If $\bar{w}_i$ is sufficiently large for all i, then $e^* = (e_1^*, ..., e_i^*, ..., e_n^*)$ is a N.E.

Proposition

If r_i is sufficiently large for all i, then $e^* = (e_1^*, ..., e_i^*, ..., e_n^*)$ is a N.E.

Proof: Rewriting

$$Y_i = u_i(b_i^*) - \psi_i(e_i^*) - [\frac{n-1}{n} Eu_i(b_i^* + z_i) + \frac{1}{n} u(-w_i) - \psi_i(\hat{e}_i)]$$

as

Risk-Averse Team X

$$Y_i = u_i(b_i^*) - \psi_i(e_i^*) - [\frac{n-1}{n}[\sum_{i \neq j}^{n} \frac{1}{n-1} u_i(b_i^* + z_i)] + \frac{1}{n} u(-w_i) - \psi_i(\hat{e}_i)]$$

i.e., as

$$\begin{aligned} Y_i &= -e^{-r_i b_i^*} - \psi_i(e_i^*) \\ &+ \frac{1}{n}(\sum_{i \neq j} e^{-r_i \{b_i^* + \frac{1}{n-1}[b_j + w_j - Q(e^*) + Q(\hat{e}_i, e^*_{-i})]\}}) \\ &+ \frac{1}{n} e^{r_i w_i} + \psi_i(\hat{e}_i) \end{aligned} \tag{19}$$

Note as r_i goes up, the first and the third terms approach zero. The second term is unaffected and the fifth one is bounded by $\psi_i(0)$ and $\psi_i(e_i^*)$. But, the fourth term exploded towards infinity. Therefore, for sufficiently large r_i, $Y_i > 0$ holds. Again, $e^* = (e_1^*, ..., e_i^*, ..., e_n^*)$ is a N.E.

Scapegoats Versus Massacres

When agents are identical, the 'scapegoat' contract is:

$$s_i(Q) = \begin{cases} \frac{Q}{n} & \text{if } Q \geq Q(e^*); \\ \frac{Q+w}{n-1} \text{ with probability } \frac{n-1}{n} & \text{if } Q < Q(e^*); \\ -w \text{ with probability } \frac{1}{n} & \text{if } Q < Q(e^*). \end{cases}$$

When agents are identical, the 'massacre' contract is:

$$s_i(Q) = \begin{cases} \frac{Q}{n} & \text{if } Q \geq Q(e^*); \\ Q + (n-1)w \text{ with probability } \frac{1}{n} & \text{if } Q < Q(e^*); \\ -w \text{ with probability } \frac{n-1}{n} & \text{if } Q < Q(e^*). \end{cases}$$

Reference: Rasmusen (1984, RJE)