Moral Hazard in Teams

What happens when there are many agents? Threre are two problems:

i) If many agents produce a joint output $x$, how does one assign the output?

There is a free rider problem here as I know that the other members of the group will work, so why should I?

ii) If many agents produce similar outputs $x_i$ by undertaking action $a_i$, there is a scope for inducing effort by playing off one agent against the other. Are relative performance contracts efficient? What about the rank order tournaments?

Hőlmstrom gives gives a partial discussion of these issues in his paper.

**The Single output, Group Incentive Problem:** Assume there are n agents, agent $i$ takes action $a_i \in A_i = [0, \infty]$

Cost of taking action: $\nu_i : \ A_i \rightarrow R$, $\nu_i^{'}(\cdot) > 0$, $\nu_i^{'}(\cdot) < 0$

$a = (a_1, a_2, ..., a_n) \in A \equiv \mathrm{X}_{i=1}^{n} A_i$ and write $a_{-i} = (a_1, a_2, ..., a_{i-1}, a_{i+1}, ..., a_n)$ ; $a = (a_i, a_{-i})$

For each agent, the utility form is:

$u_i(m_i, a_i) = m_i - \nu_i(a_i)$ , over output/(consumption) $m_i$ and action $a_i$

Prodction function: $x : A \rightarrow R$, which is strictly increasing, concave and differentiable.

And the sharing rules for the $i_{th}$ agent is: $S_i(x) \geq 0$, $i = 1, 2, ...n$

If the sharing rules are budget balancing, then: $\sum_{i=1}^{n} S_i(x) = x, \quad \forall x$

Given $a_{-i}$, each agent picks

$$\underset{\{a_i\}}{\text{Max}}\, S_i(x\,(a_i, a_{-i})) - \nu_i\,(a_i) \tag{$*$}$$

$a^*$ is a Nash equilibrium if for each $i$ given $a^*_{-i}$, $a^*_i$ satisfies $(*)$

The Social Optimum: A social Pareto optimal $a^*$ is one satisfying

$$a^* = \underset{a \in A}{\arg\max}\, [x(a) - \textstyle\sum_{i=1}^{n} \upsilon_i(a_i)]$$

If sharing rules are differentiable:

$$\frac{dS_i}{dx} * \frac{\partial x(a_i, a_{-i})}{\partial a_i} = \frac{d\nu_i}{da_i} \tag{$A$}$$

Pareto optimality requires that

$$\frac{\partial x(a_i, a_{-i})}{\partial a_i} = \frac{d\nu_i}{da_i} \tag{$B$}$$

For $(A), (B)$ to be consistent, $\frac{dS_i}{dx} = 1$,

i.e. agent needs to take full output risk at the margin.

But budget balancing implies that each agent can not take full output risk as they share the output! Formally,

$$\sum_{i=1}^{n} S_i(x) = x, \quad \forall x$$

$$\Longrightarrow \sum_{i=1}^{n} \frac{dS_i}{dx} = 1$$

Contradicting the above. Of course we have only considered differentiable rules. For a complete proof including non-differentiable rules see the Appendix.

If instead of budget balancing, we imposed

$$\sum_{i=1}^{n} S_i(x) < x, \quad \forall x$$

Then exits efficient Nash equilibria now. Suppose $a^*$ is the Pareto optimum. Let $b_i$ be a set of numbers s.t.

$$\sum_{i=1}^{n} b_i = x\left(a^*\right), \text{ and}$$

$$S_i(x) = b_i, \quad \text{if } x \geq x\left(a^*\right)$$

$$S_i(x) = 0, \quad \text{if } x < x\left(a^*\right)$$

If all other agents pick $a^*_{-i}$, and agent $i$ picks $a_i > a^*_i$, he gets nothing more even though $x > x(a^*)$. But action is costly and so this is suboptimal.

If agent $i$ pick $a_i < a^*_i$, $x < x(a^*)$ and he gets nothing.

Therefore $a^*$ is Nash.

The forcing contract works because of certainty and is very special.

All the above is old wine in a new bottle. The importance of breaking a budget constraint has been noted by Groves in his classic public goods work and is quite well known. Holmstrom's paper is a clever example of how to reuse an old idea.

However, the free rider problem is important and has many applications in finance (diversified firms etc).

**Uncertainty and Forcing contracts** Even when uncertainty exists, forcing contracts of the type may work.

Let

$F(x|a)$ c.d.f.

$f(x|a)$ p.d.f.

$F_i(x|a) \equiv \frac{\partial F(x|a)}{\partial a_i}$

$f_i(x|a) \equiv \frac{\partial f(x|a)}{\partial a_i}$

exist for $i, (x, a)$.

<u>A1:</u> $F(x|a)$ is convex in $a$

<u>A2:</u> $\frac{F_i(x|a)}{F(x|a)} \to -\infty$ as $x \to -\infty$

<u>A3:</u> $\frac{F_i(x|a)}{1-F(x|a)} \to -\infty$ as $x \to +\infty$

Assumption A is closely related to what is called concavity of distribution function. Thus,

$$F(x|\lambda a_1 + (1-\lambda)a_2) \le \lambda F(x|a_1) + (1-\lambda)F(x|a_2)$$

Thus if $a_2 >> a_1$ i.e. $\forall i, a_{2i} > a_{1i}$ then

$$F(x|a_2) < F(x|a_1) \text{ any } x$$

This kind of distribution function condition is needed to make the first order approach work. But it is weird as most standard distributions do not satisfy it.

Assumption 2 and 3 imply that the lowered bound output $(-\infty)$ and upper bound $(+\infty)$ are close to perfectly revealing about the action.

The basic idea of this kind of forcing contract is due to Mirlees. Suppose for output close to $-\infty$ (i.e. less than some bound $\overline{\overline{x}}$, one punishes the agent by some amount $k$. Of course by punishing a person with $k$ for $x < \overline{\overline{x}}$, one is not acheiving the first best for $x < \overline{\overline{x}}$. But lower $\overline{\overline{x}}$ and increase $k$ proportionately. So one is imposing a very large punishment for a low output. If this limit works correctly, i.e.

$$prob(x \le \overline{\overline{x}}|a) * \text{Punishment for } x \le \overline{\overline{x}} \to 0 \text{ as } \overline{\overline{x}} \to -\infty$$

We are OK. This is where assumption 2 is used.

Theorem: If agents are <u>risk neutral</u> and A1 and A2 hold, the first best can be approximated arbitrarily closely by group penalties.

Proof: Consider

$$S_i(x) = S_i x \qquad \text{if } x \ge \overline{\overline{x}}$$

$$S_i(x) = S_i x - k_i, \quad \text{if } x < \overline{\overline{x}}$$

$$k_i > 0, \sum S_i = 1$$

Risk neutality does not imply that there is no agency problem. This arises because of the free rider problem.

With risk neutrality, the above contract has FOC:

$$S_i E[\frac{\partial x(a)}{\partial a_i}] - k_i \frac{\partial F(\overline{\overline{x}}|a)}{\partial a_i} - \nu_i^{'}(a_i) = 0$$

Thus $a^*$ the Pareto optimal contract must satisfy this. If it does, A1 ensures global optimality.

Fix $\overline{\overline{x}}$: choose $k$ so that

$$k_i = \frac{S_i E[\frac{\partial x(a^*)}{\partial a_i}] - \nu_i^{'}(a_i^*)}{\frac{\partial F(\overline{\overline{x}}|a^*)}{\partial a_i}} = \frac{A_i}{\frac{\partial F(\overline{\overline{x}}|a^*)}{\partial a_i}}$$

For $x < \overline{\overline{x}}$, first best is not achieved due to punishment:

The punishment loss is:

$$W = \sum_i k_i F(\overline{\overline{x}}|a^*) = \sum_i \frac{A_i * F(\overline{\overline{x}}|a^*)}{\frac{\partial F(\overline{\overline{x}}|a^*)}{\partial a_i}}$$

where $A_i = S_i E[\frac{\partial x(a^*)}{\partial a_i}] - \nu_i^{'}(a_i^*)$ is a constant.

Therefore, by A2,

$$\frac{\frac{\partial F(\overline{\overline{x}}|a^*)}{\partial a_i}}{F(\overline{\overline{x}}|a^*)} \rightarrow -\infty \quad \text{as } \overline{\overline{x}} \rightarrow -\infty$$

$$\Longrightarrow \frac{F(\overline{\overline{x}}|a^*)}{\frac{\partial F(\overline{\overline{x}}|a^*)}{\partial a_i}} \rightarrow 0$$

The above trick of using unbounded supports to enforce large penalties on small probabilities is due to Mirlees.

Of course if imposing very high punishments is not allowed, this strategy does not work.

The converse strategy is the bonus strategy. Here one pays very large bonuses with very small probability.

Theorem: Under A1 and A3, the first best can be inforced.

Proof:

$$S_i(x) = S_i x + k_i, \qquad \text{if } x \geq \bar{\bar{x}}$$

$$S_i(x) = S_i x \qquad \text{if } x < \bar{\bar{x}}$$

$$k_i > 0, \sum S_i = 1$$

FOC is: $S_i E[\frac{\partial x(a^*)}{\partial a_i}] - k_i \frac{\partial F(\bar{\bar{x}}|a^*)}{\partial a_i} - \nu_i'(a_i^*) = 0$

$k_i = \frac{S_i E[\frac{\partial x(a^*)}{\partial a_i}] - \nu_i'(a_i^*)}{\frac{\partial F(\bar{\bar{x}}|a^*)}{\partial a_i}} = \frac{B_i}{\frac{\partial F(\bar{\bar{x}}|a^*)}{\partial a_i}}$ and $W = \sum_{i=1}^n B_i \frac{1 - F(\bar{\bar{x}}|a^*)}{\frac{\partial F(\bar{\bar{x}}|a^*)}{\partial a_i}}$

A3 ensures this goes to zero as $\bar{\bar{x}} \to -\infty$.

This schemes depend critically on risk neutrality.

**Sufficient statistics** Risk neutral principal

Risk averse agents

$y \equiv$ vector signals

$G(y|a)$ c.d.f

$g(y|a)$ p.d.f

$g_{a_i}(y|a)$ exists for all $i$.

$$\max_{\{a, s_i(y)\}} \int \{E(x|y,a) - \sum_i S_i(y)\} dG(y|a)$$

$$\text{s.t.} \quad (i) \quad \int u_i(S_i(y))dG(y|a) - \nu_i(a_i) \geq \overline{u_i} \quad \text{for } i \in [1, n]$$

$$(ii) \quad a_i \in \arg\max_{\{a_i'\}} \int u_i(S_i(y))dG(y|a_i', a_{-i}) - \nu_i(a_i') \quad \text{for } i \in [1, n]$$

Definition: A function $T_i(y)$ is sufficient for $y$ w.r.t. $a_i$,if exist functions $h_i \geq 0, P_i \geq 0$, s.t.

$$g(y|a) = h_i(y, a_{-i})P_i(T_i(y)|a) \qquad \forall y, a \text{ in the support of } g(\cdot)$$

(In our old notation in prior class $(x, y) \equiv y, x \equiv T(y)$ )

If $T_i(y)$ is sufficient for $y$ w.r.t. $a_i$,each $i$; $T(y)$ is sufficient for $y$ w.r.t. $a$

Theorem: If $T(y)$ is sufficient for $y$ w.r.t. $a$, then given a set of schemes $\{S_i(y)\}$, exits a set of schemes $\{\widehat{S}_i(T(y))\}$, which weakly Pareto dominate it.

Proof:

Define $\widehat{S}_i(\widetilde{T})$ follows

$$u_i\left(\widehat{S}_i(\widetilde{T})\right) = \int_{\{y|T_i(y)=\widetilde{T}\}} u_i(S_i(y))\frac{g(y|a)}{P_i(\widetilde{T}|a)}dy = \int_{\{y|T_i(y)=\widetilde{T}\}} u_i(S_i(y))h_i(y,a_{-i})dy$$

$$\Longrightarrow E[u_i\left(\widehat{S}_i(T)\right)] = E[u_i(S_i(y))] \qquad \text{for any action } a.$$

So agents pick $a^*$ as a Nash equilibrium.

For the principle, from Jensen's inequality

$$u_i\left(\widehat{S}_i(\widetilde{T})\right) = E\left[u_i(S_i(y))|\widetilde{T}\right] \leq u_i(E[S_i(y)|T]) \Rightarrow \widehat{S}_i(\widetilde{T}) \leq E[S_i(y)|\widetilde{T}]$$

$$\Rightarrow E\left[\widehat{S}_i(T)\right] \leq E[S_i(y)]$$

So the principal is better off.

Trial converse of Theorem 5:

$T(y)$ is sufficient at $a$ if it is the case that for all $i$, all $T_i$

$$\frac{g_{a_i}(y_1|a)}{g(y_1|a)} = \frac{g_{a_i}(y_2|a)}{g(y_2|a)} \qquad a.s.\;\; y_1, y_2 \in \{y|T_i(y) = T_i\} \qquad (*)$$

This means that the likelihood ratio $\frac{g_{a_i}}{g}$ given $T_i$, does not move with $y$.

By integration, we know that $(*) \Rightarrow g(y|a) = h_i(y, a_{-i})P_i(T_i\,(y)|a)$ $\qquad (**)$

So, if $(*)$ holds for $\forall i, a$; $T(y)$ is globally sufficient. If $(*)$ is false at <u>some $i$ for all $a$</u>, $T\,(y)$ is globally insufficient.

Theorem 6: If $T(y)$ is globally insufficient and $\{S(T_i(y))\}$ is a collection of unique nonconstant sharing rules in equilibrium. Then exits sharing rules $\{\widehat{S}_i(y)\}$ which are Pareto improving. Moreover they can guarantee the same actions $a^*$.

Proof: It is a pain in the neck, but I will do it for one agent. i.e. suppress let $i = 1$. Exists a set $T_i$

Exists $T_1$ and sets of positive measure

$$y_{11} \subset y_1 = \{y | T(y) = T_1\}$$

$$y_{12} \subset y_1 = \{y | T(y) = T_1\}$$

s.t. $\frac{g_a(y_1|a)}{g(y_1|a)} \neq \frac{g_a(y_2|a)}{g(y_2|a)}$

$$g(y_{kl}|a) = prob\{y \in y_{kl}|a\} \qquad l = 1, 2$$

Since $S(y)$ is not constant, there exist $T_1, T_2$ s.t.

$y_2 = \{y|T(y) = T_2\}$ is of positive value measure and $S(T_1) \neq S(T_2)$

Define

$$\widehat{S}(y) = S(T(y)) + I_{11}(y)d_{s_{11}} + I_{12}(y)d_{s_{12}} + I_2\,(y)\,d_{s_2}$$

$I_{1l} = 1 \qquad y \in y_{1l} \qquad l = 1, 2$

$I_{1l} = 0 \qquad$ otherwise

$I_2 = 1 \qquad y \in y_2$

$I_2 = 0 \qquad$ otherwise

Keeping the action fixed,

$$\triangle \mathrm{P} = -[d_{s_{11}} g(y_{11}, a) + d_{s_{12}} g(y_{12}, a) + d_{s_2} g(y_2, a)] \qquad (A)$$

$$\triangle \mathrm{A} = u_1^{'} [d_{s_{11}} g(y_{11}, a) + d_{s_{12}} g(y_{12}, a)] + u_2^{'} d_{s_2} g(y_2, a) \qquad (B)$$

$$u_1^{'} = u^{'}\left(S(T_1)\right),\ u_2^{'} = u^{'}\left(S(T_2)\right)$$

Now $S(T_1) \neq S(T_2) \Rightarrow u_1^{'} \neq u_2^{'}$

Wlog, let Then,

Sgn($\triangle$A )=Sgn$\left[u_1^{'} [d_{s_{11}} g(y_{11}, a) + d_{s_{12}} g(y_{12}, a)] + u_2^{'} d_{s_2} g(y_2, a)\right]$ $\qquad (C)$

Fix $d_{s_2} > 0$,Require $\triangle \mathrm{P} = 0 \Rightarrow d_{s_{11}} g(y_{11}, a) + d_{s_{12}} g(y_{12}, a) = -d_{s_2} g(y_2, a)$

Substitute into $(C) \Rightarrow$ Sgn($\triangle$A )=Sgn$\left[u_1^{'}\left(-d_{s_2} g(y_2, a)\right) + u_2^{'} d_{s_2} g(y_2, a)\right] > 0$, as $u_2^{'} > u_1^{'}$.

**Relative Performance Evaluation** $\quad x(a, \theta) = \sum_i x_i(a_i, \theta_i) \qquad \theta = (\theta_1, \theta_2, ..., \theta_n)$

Theorem 7: Assume $x_i(a, \theta)$ is monotone in $\theta_i$. Then the optimal sharing rule of agent $i$ depends on individual $i'$s output alone if and only if outputs are independent.

Proof: If $\theta_i$ is independent

$f(x|a) = \prod_{i=1}^{n} f_i(x_i|a_i) \Rightarrow T_i(x) = x_i$ is sufficient for $x$ w.r.t. $a_i$. Therefore, by Theorem (5), $S_i$ depends on $x_i$ alone.

If $\theta_1$ and $\theta_2$ are dependent, let $a_2$ be fixed at $a_2^*$

Since $x_2 = x_2(a_2^*, \theta_2) \Rightarrow \theta_2 = x_2^{-1}(a_2^*, x_2)$

Same for $x_1$; Since $x_1 = x_1(a_1, \theta_1) \Rightarrow \theta_1 = x_1^{-1}(a_1, x_1)$

Therefore, $\widetilde{f}(\theta_1, \theta_2) = \widetilde{f}(x_1^{-1}(a_1, x_1), \theta_2)$ ; $\widetilde{f}$ is the joint distribution of $(\theta_1, \theta_2)$

$$\frac{f_{a_1}(x_1|\theta_2, a_1)}{f(x_1|\theta_2, a_1)} = \frac{\widetilde{f}_1(x_1^{-1}(a_1, x_1), \theta_2)}{\widetilde{f}(x_1^{-1}(a_1, x_1), \theta_2)} \frac{\partial x_1^{-1}(a_1, x_1)}{\partial a_1}$$

Since $\theta_1, \theta_2$ are dependent $\frac{\widetilde{f}_1}{\widetilde{f}}$depends on $\theta_2$. Sufficiency of $x_1$ does not hold.

Implications: Contracts on other agents output is optimal only if correlations between outcomes $x_i$ exist. If they not, i.e., correlations are independent, such contracts do not pay.

Two examples:

$$\text{(I)} \quad x_i(a_i, \theta_i) = a_i + \widetilde{\eta} + \widetilde{\varepsilon}_i \qquad i = 1, 2...n$$

$$\text{(II)} \quad x_i(a_i, \theta_i) = a_i(\widetilde{\eta} + \widetilde{\varepsilon}_i) \qquad i = 1, 2...n$$

$\varepsilon_i$ are independent idiosyncratic shocks, normally distributed with precision $\tau_i$

$\eta$ is a common shock normally distributed independent of $\varepsilon_i$

Theorem:

For (I) (II) above, let, $\overline{\overline{x}} = \sum \alpha_i x_i$

For (I), let $\alpha_i = \frac{\tau_i}{\overline{\overline{\tau}}}$, $\quad \overline{\overline{\tau}} = \sum \tau_i$

For (II), let $\alpha_i = \frac{\tau_i}{\overline{\overline{\tau}} a_i}$, $\quad \overline{\overline{\tau}} = \sum \tau_i$

For both (I) and (II), the optimal contract is a function of $(x_i, \overline{\overline{x}})$ alone.

Proof: For (I)

$f(x|a) = k\int \exp\{ -\frac{1}{2}[\sum_j \tau_j(x_j - a_j - \mu_j - \eta)^2 + \tau_0(\eta - \mu_0)^2]\}d\eta \qquad (E)$

We get this by setting

$$f(x|a) = \int f(x|a,\eta)g\,(\eta)\,d\eta$$

$\mu_0 =$ mean $\eta$ , $\tau_0 =$ precision $\eta$

$\tau_j, \mu_j$ are mean precision and mean of $\varepsilon_j$

Let

$$\overline{Z_{-i}} = \sum_{k\neq i}(\frac{\tau_k}{\overline{\overline{\tau}}_{-i}})(x_k - a_k - \mu_k)$$

$$\overline{\tau}_{-i} = \sum_{k\neq i}\tau_k$$

Then

$$\sum_j \tau_j(x_j - a_j - \mu_j - \eta)^2 = \sum_{j\neq i}\tau_j(x_j - a_j - \mu_j - \overline{Z_{-i}} + \overline{Z_{-i}} - \eta)^2 + \tau_i(x_i - a_i - \mu_i - \eta)^2$$

$$= \sum_{j\neq i}\tau_j(x_j - a_j - \mu_j - \overline{Z_{-i}})^2 + (n-1)(\overline{Z_{-i}} - \eta)^2 + \tau_i(x_i - a_i - \mu_i - \eta)^2$$

Take this mess , substitute in $(E)$ . Integrate over $\eta$. Then;

$$f(x|a) = h_i(x|a_{-i})\widehat{P}_i(\overline{Z_{-i}}, x_i|a)$$

But $\overline{Z_{-i}} = (\overline{\tau}\overline{x} - \tau_i x_i)/\tau_i - \sum_{k\neq i}(\frac{\tau_k}{\overline{\overline{\tau}}_{-i}})(a_k + \mu_k) \Rightarrow \widehat{P}_i(\overline{Z_{-i}}, x_i|a) = P_i(\overline{x}, x_i|a)$ , Sufficiency follows.

$\overline{x} = \sum \frac{\tau_i}{\overline{\overline{\tau}}} x_i$, is a "scale" weighted average. This measures the information in output $i$.

If $\tau_i \to \infty$,the $\eta$ effect is small (noise dominates). Therefore, searching for correlation is meaningless. i.e. we do not contract on $\overline{x}$.

What if we have a large number of agents? Does the Central Limit Theorem reveal $\eta$ ? Yes.

Theorem 9: Suppose $\eta, \varepsilon_1, \varepsilon_2, ..., \varepsilon_n$ are independent with uniformly bounded variance.

Suppose when $\eta = 0$, the single agent solution is a unique $a_i^*$. then we achieve this as the number of agents $n \to \infty$.

Proof:

Let $S_i^*(x_i)$ be the optimal contract when $\eta = 0$

$a_i^*$be the optimal action when $\eta = 0$

Let

$$q_j = \eta + \varepsilon_j$$

$$\overline{q_{-i}} = \frac{1}{n-1}\sum_{j\neq i} q_j$$

As $n \to \infty, \overline{q_{-i}} \to \eta$. Thus

$$\int u_i(S_i^*(a_i + \eta + \varepsilon_i - \overline{q_{-i}}))dP(\eta, \varepsilon_1, \varepsilon_2, ..., \varepsilon_n)$$

$$\xrightarrow{uniformly} \quad \int u_i(S_i^*(a_i+\varepsilon_i))dP(\varepsilon_i)$$

Since $a_i^*$ is a unique solution to

$$\max_{\{a_i\}} \int u_i(S_i^*(a_i+\varepsilon_i))dP(\varepsilon_i) - \nu_i(a_i)$$

We are done.

Now $\overline{q_{-i}}$ can be inferred by calculating $x_i - a_i = \eta + \varepsilon_i$

■ QED

Thus, a large number of agents can eliminate the common uncertainty $\eta$ from contract!