EC951, Lecture 1: Moral Hazard and the Principal-Agent model

"Hardly a competent worker can be found who does not devote a considerable amount of time to studying just how slowly he can work and still convince his employer that he is going at a good pace." –Fredrick Taylor, *The Principles of Scientific Management*, 1929.

Quotes

“...If you cut every corner, it’s really not so bad, everybody does, even mom and dad. If nobody sees it, then nobody gets mad.” – Shari Bobbins, The Simpsons.

“Some of the activities we undertook contributed to the prevailing mood at the time. We didn’t know it then or even today when it actually crossed over into bubble territory. But we lent money out too cheaply ... without the traditional safeguards.” –Lloyd Blankfein, Chairman & CEO, Goldman Sachs, at Financial Crisis Inquiry.

Moral hazard I: Two actors, two actions, two outcomes

Readings:

Moralhazardlecture1.pdf, Laffont and Martimort, Ch. 4 (4.1– 4.4 for this lecture)

Alt: Salanie, ch 5

A good supplement with many motivating examples: Milgrom and Roberts, Chapter 6 (and part of 7)

(Optional: L&M Intro, ch 1)

Note: Lecture notation follows L&M as much as possible

Exercises: Moral hazard practice questions.pdf (1.1 and 1.2; limited liability parts if you are ready)

Supplementary questions:

M&R "food for thought" (p. 197, q. 3), "quantitative problems 4 and 5"

Moral Hazard: Motivation and background

- Principal(s) (e.g., employer) enlists Agent(s) to perform a task
- Agent (e.g., worker) chooses level of effort/action, performs a task which can yield various outcomes
- Conflict between P and A over preferred action.
- The outcome of the task is stochastic (has a random component), but the probability of each outcome may depend on the agent's action. I.e., output a noisy signal of action.
- Principal observes *outcome* but not *effort/action*, thus Principal can pay Agent based on *outcome* only.
- Principal wants motivate Agent to take the "correct" action, thus might pay more for "good" outcome,
  - Could "sell the farm"; but there may be barriers to "efficiency": Agent may be risk-averse or liability-constrained

Working from home?

Key results and issues: An intuitive preview

- Uncertainty is key; if the outcome were perfectly correlated to action, the action would be, in effect, observable
- Could make agent "residual claimant" $\Longrightarrow$correct incentives (in simple model), but with risk aversion or limited liability (LL) this is costly
- To minimize this cost, the second-best effort may differ from the first best (*) With multiple types of effort, effort can be distorted downwards slightly. $\Longrightarrow$possible "allocative inefficiency"
- Insurance/efficiency trade-off
  - If we pass any risk to the risk-averse agent this is itself inefficient – Risk aversion can lead to a joint loss; agent gets no 'rent'
- LL may lead to the agent getting a 'rent'

(*) With two types of effort, first best effort e=1 occurs "less often" (for a narrower set of parameters).}

See flowchart diagram: agencyresultsA.png

2 x 2 x 2 model

- Players: Principal (P)— employer, Agent(A) — worker
- Agent chooses effort $e \in \{0, 1\}$
- Agent receives transfer t
- Cost or disutility of effort is $\psi(e)$ (Note: this assumes separability)(Essentially, no income effects.) For the models in L&M ch. 4, this assumption is irrelevant., i.e., $U = u(t) - \psi(e)$
- Output can either be high $(\bar{q})$ or low $(\underline{q})$. The probability of high output is higher if effort is high $(e = 1)$than if it is low$(e = 0)$.
- Principal's expected payoff is just the expected output minus wage: $E(q - t)$
  - L&M let revenue be a concave function $S(q)$ and thus let P be risk-neutral over "net revenue" $S(q) - t$. In these notes we make the simpler assumption that revenue is linear in q and thus P can be risk-neutral over $q - t$. The difference is inconsequential here; you could just think of the agent directly producing 'benefits' q for the principal.

Players: P, A

Actions:

1. P sets "take-it-or-leave-it" (TIOLI) contract $t(e, q).(*)$ Here this is just a pair of transfers $\{(\bar{t}, \underline{t})\}$
2. Agent chooses to work or not, and chooses effort $e \in \{0, 1\}$
3. P observes output (*and observes – or does not observe – effort*)

$$\begin{aligned} \text{Where output} &\in \{\bar{q}, \underline{q}\} \\ \Pr(\bar{q}|e &= 1) = \pi_1 \\ \Pr(\bar{q}|e &= 0) = \pi_0 \end{aligned}$$

(*) Can only condition on effort if it is observable, of course.

4. Agent gets transfer $t(e, q)$

5. Payoffs:

$$\begin{aligned} EU_P &= E(q-t) \\ EU_A \equiv v &= \begin{cases} E(u(t)) - \psi(e) \text{ if accept job} \\ \quad = \bar{U} = 0 \text{ otherwise} \end{cases} \\ \psi(0) &= 0, \psi(1) = \Psi \end{aligned} \tag{1}$$

Note: $\bar{U}$ represents reservation utility. t_r represents the wage that provides this (with zero effort).

I.e., $u(t_r) = \bar{U} = 0$, where the latter equality is a normalization.

Further assumptions

Assumption

$$1 > \pi_1 > \pi_0 > 0 \tag{2}$$

This ensures that the problem is non-trivial – if this didn't hold, then effort is something neither party wants, so there is no conflict.

More generally, the models will assume "first-order stochastic dominance":

$\Pr(\tilde{q} \leq q^* | e)$is decreasing in e for any level of production q^*

Assumption

"High effort is efficient"

$$\pi_1\bar{q} + (1-\pi_1)\underline{q} - u^{-1}(\psi(1)) > \pi_0\bar{q} + (1-\pi_0)\underline{q} - u^{-1}(\psi(0)) \tag{3}$$

$$I.e., (\pi_1 - \pi_0)(\bar{q} - \underline{q}) > u^{-1}(\Psi)$$

This also ensures the problem is non-trivial. Otherwise it will surely be too costly to induce high effort (even under complete information). *[class question: why?]*

Note: this does *not* mean that the principal will necessarily decide to induce high effort under *incomplete* information.

Assumption

Expected production (revenue) with low effort is at least as large as the reservation wage, i.e.,

$$\pi_0 \bar{q} + (1 - \pi_0)\underline{q} \geq t_r. \tag{4}$$

Assumption

The agent is risk-averse, i.e., $u(t)$ is strictly concave. This will imply a risk-sharing/incentive tradeoff.

Assumption

If an agent is indifferent between two choices, she will choose the one the principal prefers.

This is a conventional assumption.

It is considered "without loss of generality" (WLOG) because we assume the principal could always 'sweeten the deal' or increase the incentive a tiny amount, inducing cooperation but yielding nearly identical results. This assumption must hold for an equilibrium to exist, because there is no "smallest real number." Even without this as assumption, this is equilibrium behaviour.

Note: In all of this, we assume that contracts are costlessly enforceable, and all that is mutually observable is verifiable to an outside party (who enforces the contract).

Definition

A contract satisfies the **Participation Constraint (PC)** if it ensures that the Agent(s) prefer(s) to take part in the contract rather than opt out.[a]

[a] Other texts call this the "individual rationality" (IR) constraint.

Definition

A contract satisfies the **Incentive Compatibility Constraint (IC)** if it ensures that the Agent(s) prefer(s) to take the action that the Principal wants them to take, rather than any other action.

Definition

If both the PC and IC constraints are satisfied then the contract is **Incentive Feasible** (aka 'feasible' or 'implementable').

We can assume WLOG that the principal will set a contract that is incentive feasible. The proof is fairly trivial.

Optimal Contract Under **Observable and verifiable** Effort ("first-best")

$$t^O(e) = \mathbf{1}[e=1] \times t^* + \mathbf{1}[e=0] \times Z \quad (5)$$

where $u(t^*) - \Psi = \bar{U} = 0 \Longrightarrow u(t^*) = \boldsymbol{\Psi}$

where Z is some small or negative number (punishment)

Note: $1[]$is the indicator function

$$\Longrightarrow t^* = u^{-1}(\Psi)$$
$$\Longrightarrow t(e,q) = t(e) = \mathbf{1}[e=1] \times u^{-1}(\Psi) + \mathbf{1}[e=0] \times Z$$

Assuming the utility function is invertible. Note: the L&M text defines $h(\cdot) = u^{-1}(\cdot)$

This is the optimal contract for the principal because:

1. It induces $e = 1$

Agent accepts & works because this contract meets her PC and IC.

$$E[u(t^O(1)) - \boldsymbol{\Psi}(1)] \\ = u(u^{-1}(\Psi)) - \Psi = \Psi - \Psi \geq 0 = \bar{U} \tag{PC}$$

$$E[u(t^O(1))] - \boldsymbol{\Psi}(1) = u(t^*) - \Psi = 0 \geq E[u(Z)] - \Psi(0) = u(Z) \tag{IC}$$

This inequality must hold for some sufficiently small or negative Z. If we assume zero utility occurs at zero salary then $Z = 0$ will suffice; more generally we require $Z < t_r$.

2. It minimizes t s.t. $e = 1$

Since this contract meets the PC (and IC) with equality, the principal can offer no less. {class: consider what would change if t were lowered a tiny bit}

Since the payment is a certainty, the agent bears no risk, thus there is no 'net loss'; total surplus is maximized.

Note: Since there is no net loss, and we know (we assumed) high effort is efficient, and P captures the entire surplus, we know that P prefers to induce high effort here.

Surplus:

P gets

$$EU_P = V_1 = \pi_1 \bar{q} + (1 - \pi_1)\underline{q} - u^{-1}(\Psi) \tag{6}$$

A gets no surplus

A bears no risk: $EU_A = u_A = 0$

P bears all risk.

A Pareto-optimal outcome is obtained: efficient effort, risk optimally distributed.

The optimal contract in such a case is a *"wage contract."*

Optimal Contract With **Unobservable** Effort

Contract: $(\bar{t}, \underline{t} \equiv \vec{C}) \in \mathbf{R}^2$

To get $e = 1$, need $\vec{C}$ to satisfy PC and IC

... but P can condition payment on output only. Let $\bar{t}$ be the pay following high output and $\underline{t}$ the pay following low output.

PC:

The PC is:(*)

$$\pi_1 u(\bar{t}) + (1 - \pi_1) u(\underline{t}) - \Psi \geq \bar{U} = 0 \tag{PC}$$

(*)The PC implicitly 'assumes' e=1, i.e., assumes the IC holds. Basically, each constraint 'assumes' the other holds.

IC

$$(EU_A|e=1) \geq (EU_A|e=0)$$

$$\begin{aligned} \pi_1 u(\bar{t}) + (1-\pi_1)u(\underline{t}) - \Psi &\geq \pi_0 u(\bar{t}) + (1-\pi_0)u(\underline{t}) \\ \text{i.e.,}(\pi_1 - \pi_0)(u(\bar{t}) - u(\underline{t})) &\geq \Psi \end{aligned}$$

$\Longrightarrow$ $\bar{t} > \underline{t}$, is a necessary but not sufficient condition for IC.

Principal's problem

The principal's problem is:

$$\min E(t) = \pi_1 \bar{t} + (1 - \pi_1)\underline{t} \quad (7)$$
$$s.t. \text{ IC and PC}$$

Assuming this will be superior to inducing 0 effort; which can and should be checked!

(We already assumed conditions such that satisfying PC only and hiring the agent and inducing no effort is superior to not hiring; but in general this should also be checked.)

Proof of binding constraints

(Note: Binding constraints will allow a "plug in and solve" solution)

Conjecture

IC holds with equality:

$$(EU_A|e = 1) = (EU_A|e = 0) \tag{8}$$

Sketch:[a] *Proof by contradiction. If IC not binding, i.e., if $(EU_A|e = 1) > (EU_A|e = 0)$, can "push them together", i.e., raise $\underline{t}$ and reduce $\overline{t}$ at least marginally, (I) preserving IC and PC while (II) reducing $E(t)$ and thus increasing P's payoff.*

(a) This proof is similar to that of Salanie, p. 123. You may prefer the computational 'Lagrangian' proof in L&M

Intuition: The greater the gap between the payments for high and low output, the more P must compensate A for sustaining risk. Thus, P tries to minimize this gap while maintaining IC.

IC binding: proof (by contradiction) in 2 steps

Proof.

I. If IC (or any constraint) holds with inequality, *we can relax it while preserving it.* Imagine 'small' changes in the transfers, $d\bar{t}$ and $d\underline{t}$. We preserve PC iff

$$\pi_1 u(\bar{t}) + (1-\pi_1)u(\underline{t}) - \Psi \geq 0 \tag{9}$$

continues to hold after the deviations $d\bar{t}$ and $d\underline{t}$. To first order, the change in the left hand side of the above inequality is:

$$\pi_1 u'(\bar{t})d\bar{t} + (1-\pi_1)u'(\underline{t})d\underline{t} = 0 \tag{10}$$

Thus, if

$$d\bar{t} = \frac{-(1-\pi_1)}{\pi_1}\frac{u'(\underline{t})}{u'(\bar{t})}d\underline{t}$$

...there is only a 'second order' effect on the utility of agent who puts in high effort (the lhs of the PC) from these changes. For small enough dt this will be outweighed by any first-order effects □

Proof.

II. This PC (& IC)-preserving adjustment will decrease $E(t)$:[a]

$$dE(t) = \pi_1 d\bar{t} + (1-\pi_1)d\underline{t} \tag{11}$$

$$= \pi_1 \frac{-(1-\pi_1)}{\pi_1}\frac{u'(\underline{t})}{u'(\bar{t})}d\underline{t} + (1-\pi_1)d\underline{t}$$

$$= (1-\pi_1)(1-\frac{u'(\underline{t})}{u'(\bar{t})})d\underline{t} < 0$$

$0 < u'(\bar{t}) < u'(\underline{t})$ by concavity (risk aversion). Thus, $dE(t) < 0$, lowering costs. Note this effect on costs is a *first-order effect.* □

[a] "Plugging in" the dt terms from the previous slide

...

Since the cost savings are a first-order effect,

and the agent's loss of utility is a second order effect,

for some small increase in $d\underline{t}$ and decrease in $d\bar{t}$ (in the above proportions)

the cost savings are enough to compensate the agent for her loss of utility (if any)

while still saving money.

...

I.e, decreasing $\bar{t}$ while increasing $\underline{t}$ enough to maintain PC, the principal's net change in costs is:

$$= (1 - \pi_1)(1 - \frac{u'(\underline{t})}{u'(\bar{t})})d\underline{t} + \text{'second-order term'} + \text{'third-order term'} + ...$$

By Taylor's theorum (Cauchy's estimate) we know that for small enough $d\underline{t}$ the first term, which is negative, will be larger in magnitude than the other terms (the 'remainder'), hence the net change in costs is negative.

Conjecture

PC holds with equality *(which implies the agent gets no rent)*

Proof.

Sketch: If $EU_A > \bar{U}$, P can reduce $\underline{t}$, **I.** preserving PC and IC, and **II.** reducing $E(t)$. →contradicts optimality.

I. PC is naturally preserved because it held with inequality.

IC continues to hold because, taking the derivative of the left hand side of the IC "$(\pi_1 - \pi_0)(u(\bar{t}) - u(\underline{t})) > \psi$":

$$\frac{dIC}{d\underline{t}} = \frac{d}{d\underline{t}}[(\pi_1 - \pi_0)(u(\bar{t}) - u(\underline{t})) - \psi] = -(\pi_1 - \pi_0)u'(\underline{t}) < 0 \quad (12)$$

So if we *reduce* $\underline{t}$ ($d\underline{t} < 0$) we *increase* the relative benefit of hard work, making the constraint *less* binding. □

Conjecture

The cost of inducing high effort ($e = 1$) is greater under asymmetric information then under observable effort.

Thus, inducing a high effort occurs 'less often' (for a smaller parameter space) under asymmetric information.

[Draw diagram: Fig 4.4 from L&M]

Sketch of proof: Any (nondegenerate) 'probablistic pay' that has the same expected value (hence cost to the principal) as a certain pay will be less valuable to the agent, since she is risk averse. Hence, when wage is stochastic, as it must be it must be to induce effort with asymmetric information, it must be higher on average, hence more costly.

Proof: Agency costs positive (with a risk-averse agent)

Proof.

PC holds with equality $\Longrightarrow \pi_1 u(\bar{t}) + (1-\pi_1)u(\underline{t}) = \psi + \bar{U}$

From this we can prove that the expectation of the (minimum) transfer that induces high effort is higher under asymmetric information then under observable effort, i.e.,

$$\pi_1 \bar{t} + (1-\pi_1)\underline{t} > t^* = u^{-1}(\psi) + \bar{U} = u^{-1}(\psi)$$

...This comes from the concavity (risk aversion) assumption, which implies:

whenever $' \neq \underline{t}'$and probabilities are nondegenerate

Including where $\pi_1 \bar{t}' + (1-\pi_1)\underline{t}' = t^*$

So a stochastic contract with the same expected cost as the first-best contract will not meet the PC constraint, hence meeting the PC constraint must be more costly □

..

Proof.

Thus, transfers with the same expected value as the one that just satisfies PC in the first-best (conditioned on observable effort case) do not satisfy PC when conditioned on output, i.e.,

$$\pi_1 u(\bar{t}') + (1-\pi_1)u(\underline{t}') < \bar{U} \tag{13}$$

$$\text{...if } \bar{t}' \text{ , } \underline{t}' \text{ satisfy } \pi_1\bar{t}' + (1-\pi_1)\underline{t}' = t^O(1)$$
$$= t^* = u^{-1}(\Psi) \tag{14}$$

$$\Longrightarrow EU_A(\bar{t}', \underline{t}' | e = 1) < \Psi \tag{15}$$
$$\Longrightarrow E(t) > t^* \text{ must hold for PC to hold}$$

□

Result: for e=1

The participation and incentive constraints hold with equality. Hence, we have 2 (non-redundant) equations in 2 unknowns; we can solve this.

PC: $E(u(t)|e=1) - \mathbf{\Psi}(1) = \bar{U} = 0$

$\implies \pi_1 u(\bar{t}) + (1-\pi_1)u(\underline{t}) = \Psi + \bar{U} = \Psi$

$\implies u(\bar{t}) = \frac{\Psi - (1-\pi_1)u(\underline{t})}{\pi_1}$

IC: $\pi_1 u(\bar{t}) + (1-\pi_1)u(\underline{t}) - \Psi = \pi_0 u(\bar{t}) + (1-\pi_0)u(\underline{t})$

$\Longrightarrow (\pi_1 - \pi_0)u(\bar{t}) - (\pi_1 - \pi_0)u(\underline{t}) = \Psi$

$\Longrightarrow (\pi_1 - \pi_0)(u(\bar{t}) - u(\underline{t})) \equiv \Delta\pi\Delta u = \Psi$

$\Longrightarrow \Delta u = \frac{\Psi}{\Delta\pi}$

Note that the utility (and thus the payment) to the agent for a high and low outcome are different; this implies there is inefficient risk sharing.

$\Longrightarrow u(\bar{t}) = \frac{\Psi}{\Delta\pi} + u(\underline{t})$

Plugging PC in to IC: (*)

$(\pi_1 - \pi_0)(\frac{\Psi-(1-\pi_1)\underline{u}}{\pi_1} - \underline{u}) = \Psi$

$\implies \underline{u} = \Psi\frac{\pi_0}{\pi_0-\pi_1} = \Psi - \pi_1\frac{\Psi}{\Delta\pi}$

Using $\frac{\pi_0}{\pi_0-\pi_1} = \frac{\pi_1}{\pi_0-\pi_1} + 1 = 1 - \frac{\pi_1}{\Delta\pi}$ by partial fractions

Note $\underline{U} \equiv \underline{u} - \Psi < 0$ since $\pi_0 - \pi_1 < 0$. Thus $\underline{u} < \bar{U}$

Plugging in to IC: $\bar{u} = \frac{\Psi}{\Delta\pi} + \underline{u} = \frac{\Psi}{\Delta\pi} + \Psi - \pi_1\frac{\Psi}{\Delta\pi} = \Psi + (1-\pi_1)\frac{\Psi}{\Delta\pi}$

Note $\bar{U} \equiv \bar{u} - \Psi > 0$, thus $\bar{u} > \bar{U}$

$EU = 0$ of course, from PC binding.

(*) note $\bar{u} = u(\bar{t})$ and $\underline{u} = u(\underline{t})$ notation corresponds to (sub)utility of high and low *transfers*

Properties of Optimal Contact for e=0

I.e., what if P chooses not to induce high effort?

No incentives are needed, only PC. Thus the problem is:

$$\min E(t) = \min_{\underline{t},\bar{t}}\{\pi_0 \underline{t} + (1-\pi_0)\bar{t}\} \tag{16}$$
$$s.t.\ \pi_0 u(\bar{t}) + (1-\pi_0)u(\underline{t}) \geq \bar{U} = 0$$

Thus the optimal contract if there is to be zero effort is $\bar{t} = \underline{t} = t_r$.

This minimizes the risk bourne by the agent, thus satisfying the PC at the minimum cost to the principal. See previous argument under "agency costs positive"

This is just intuition. See text for formal statement.

Risk Neutral agent, unlimited liability, unobserved effort (LM 4.2)

Theorem

With a risk neutral agent and unlimited liability, there is no cost (relative to the perfect-information case) to providing incentives. I.e., the unobservability of effort is not costly.

Proof.

Choose $\underline{t}, \bar{t}$ so that PC is satisfied:

$$\begin{aligned} EU_A &= \pi_1 u(\bar{t}) + (1 - \pi_1)u(\underline{t}) \quad (17) \\ \text{...from risk neutrality...} &= u(\pi_1 \bar{t} + (1 - \pi_1)\underline{t}) \\ &= \Psi + \bar{U} \end{aligned}$$

□

Lemma

There is no cost to increasing the difference between $\bar{t}$ and $\underline{t}$.

Proof.

If P raises $\bar{t}$ by $\frac{X}{\pi_1}$, he can lower $\underline{t}$ by $\frac{X}{1-\pi_1}$ and preserve the same utility level and hence the PC. These two terms can be moved as far apart as needed!

This will cost the principal (in expectation) $\pi_1 \frac{X}{\pi_1} - (1-\pi_1)\frac{X}{1-\pi_1} = X - X = 0$, i.e., will be costless (to both P and A since the expected value is the same).

Thus the IC can be satisfied as in the 'first-best' case at no additional cost, by making the agent bear all of the risk (or even more).

□

Variety of first-best implementations with risk neutrality

Since the agent is risk-neutral, the principal can induce high effort by meeting or exceeding the IC constraint, as much as he likes, at no additional cost, and still capture all the surplus.

One way to do this is to make the agent the "residual claimant," e.g., "sell him the store" for the value of the expected profit (see L&M pp 153-154).

References/Suggestions

- IMPORTANT: We will go over limited liability to some extent in lecture/class, but you must go over it in detail on your own, especially for the 2-type case (L&M, section 4.3).

- For a parametric example, see L&M pp. 161
- For a simpler numerical example, see Milgrom and Roberts, pp. 200-203
- M&R ch 6-7 for a more applied treatment.
- More mathematical options: Tirole; Mas-Collel et al.