Labor Economics, 14.661. Lectures 6 and 7: Moral Hazard and Applications

Daron Acemoglu

MIT

November 15 and 17, 2011.

Introduction

- Labor economics typically dealing with supply, demand and allocations in the market.
- Much of labor is transacted within firms.
- Potential new frontier of labor economics: understand what is happening within firms.
- Two aspects:
  1. Incentives within firms
  2. Allocation of workers to firms
- We start with incentives within firms.

Basic Moral Hazard Framework

- Imagine a single worker (agent) is contracting with a single employer (principal).
- The agent's utility function is

$$H(w,a) = U(w) - c(a)$$

- w =wage,
- $a \in \mathbb{R}_+$ = action/effort,
- $U(\cdot)$ = concave utility function
- $c(\cdot)$ =convex cost of effort/action.
- $\bar{H}$ = outside option of the agent.
- x = output/performance.

Basic Moral Hazard Framework (continued)

- Output a function of effort a and random variable $\theta \in \mathbb{R}$

$$x(a, \theta).$$

- Greater effort→higher output, so

$$x_a \equiv \frac{\partial x}{\partial a} > 0$$

- Typically, x is publicly observed, but a and θ *private information* of the worker.
- The principal cares about output minus costs:

$$V(x - w)$$

- V typically increasing concave utility function.
- Special case: V linear (risk neutral principal).

Contracts

- Let Ω be the set of observable and contractible events, so when only x is observable, $\Omega = \mathbb{R}$.
  - what is the difference between observable in contractible events?
- When any two of x, a, and θ are observable, then $\Omega = \mathbb{R}_+ \times \mathbb{R}$ (why only two?).
- A contract is a mapping

$$s : \Omega \to \mathbb{R}$$

  specifying how much the agent will be paid as a function of contractible variables.
- When there is limited liability, then

$$s : \Omega \to \mathbb{R}_+$$

Timing of Events

- Again a dynamic game of asymmetric or incomplete information (though incompleteness of information not so important here, why?)
- *Timing:*

1. The principal offers a contract $s : \Omega \to \mathbb{R}$ to the agent.
2. The agent accepts or rejects the contract. If he rejects the contract, he receives his outside utility $\overline{H}$.
3. If the agent accepts the contract $s : \Omega \to \mathbb{R}$, then he chooses effort a.
4. Nature draws θ, determining $x(a, \theta)$.
5. Agent receives the payment specified by contract s.

- Look for a Perfect Bayesian Equilibrium.

Contract Design without Asymmetric Information

- Suppose a is observed.
- Then the principal chooses both the contract $s(x, a)$ (why is it a function of both x and a?), and the agents chooses a.
- Given what types of contracts will be accepted by the agent and what the corresponding effort level will be, at step 1 the principal chooses the contract that maximizes her utility.
- With analogy to oligopoly games$\approx$ Stackleberg leader.
  - i.e., the principal anticipates the action that the agent will choose.
  - equivalently, the principal chooses the effort level as well taking the optimizing behavior of the agent as a constraint$\longrightarrow$ *incentive compatibility constraint (IC)*.

Solution without Asymmetric Information

- Maximization problem:

$$\begin{array}{lll} & \max_{s(x,a),a} \mathbb{E}\left[V(x-s(x,a))\right] & \\ \text{s.t.} & \mathbb{E}\left[H(s(x,a),a)\right] \geq \overline{H} & \text{Participation Constraint (PC)} \\ \text{and} & a \in \arg\max_{a'} \mathbb{E}\left[H(s(x,a'),a')\right] & \text{Incentive Constraint (IC)} \end{array}$$

where expectations are taken over the distribution of θ.

- Much simpler than the canonical moral hazard problem, because the principal is choosing $s(x,a)$.
  - In particular, she can choose s such that $s(x,a) = -\infty$ for all $a \neq a^*$, thus effectively implementing a^*.
  - because there is no moral hazard problem here given that there is no *hidden action.*

Solution without Asymmetric Information (continued)

- Therefore, when optimal effort is a^*, the maximization problem becomes:

$$\max_{s(x)} \mathbb{E}\left[V(x - s(x)\right]$$

subject to

$$\mathbb{E}\left[U\left(s\left(x\right)\right)\right] \geq \overline{H} + c\left(a^*\right).$$

Solution without Asymmetric Information (continued)

- This shows that given a^*, we have a simple risk-sharing problem.
- This is a solution to the following Lagrangean:

$$\min_{\lambda} \max_{s(x)} \mathcal{L} = \mathbb{E}\left[V(x - s(x)\right] - \lambda\left[\overline{H} + c\left(a^*\right) - \mathbb{E}\left[U\left(s\left(x\right)\right)\right]\right]$$

- Nonstandard feature: choosing a function $s\left(x\right)$.
- But no constraint on the form of the function, so the maximization can be carried out pointwise
  - intuition: suppose x takes a finite number of values.

Incorrect Solution

- One might be tempted to write:

$$\mathbb{E}\left[V'(x - s(x)\right] = \lambda \mathbb{E}\left[U'\left(s\left(x\right)\right)\right].$$

- But this would be incorrect.
- Recall that $x = x\left(a, \theta\right)$, so once we fix $a = a^*$ and condition on x, there is no more uncertainty, so there should not be any expectations.

Solution without Asymmetric Information

- Hence, the right first-order conditions are:

$$\frac{V'(x - s(x))}{U'(s(x))} = \lambda \text{ for all } x, \tag{1}$$

- This implies *perfect risk sharing* (why is this perfect risk sharing?)
- To see this more clearly, we can reason as follows:
  - given level of a, the variation in θ induces a distribution of x, denote this by

$$F(x \mid a)$$

  - since $a = a^*$ and we can choose $s(x)$ separately for each x, there is no more uncertainty conditional on x.
  - thus (1) follows.

Incentives-Insurance Tradeoff

- Now suppose that Ω only includes the output performance, x
- Therefore, *incentive-insurance trade-off* (why?)
- The problem becomes

$$\begin{array}{lll} & \max_{s(x),a} \mathbb{E}\left[V(x - s(x)\right] & \\ \text{s.t.} & \mathbb{E}\left[H(s(x), a)\right] \geq \overline{H} & \text{Participation Constraint (PC)} \\ \text{and} & a \in \arg\max_{a'} \mathbb{E}\left[H(s(x), a')\right] & \text{Incentive Constraint (IC)} \end{array}$$

- Major difference: $s(x)$ instead of $s(x, a)$.

Incentives-Insurance (continued)

- Again suppressing θ, we work directly with $F(x \mid a)$.
- Natural assumption:

$$F_a(x \mid a) < 0,$$

  (implied by $x_a > 0$)
  $\rightarrow$an increase in a leads to a *first-order stochastic-dominant* shift in F.
- Recall that F first-order stochastically dominates another G, if

$$F\left(z\right) \leq G\left(z\right)$$

  for all z.

Basic Moral Hazard Problem

- Canonical problem:

$$\begin{aligned} \max_{s(x),a} \quad & \int V(x - s(x)) dF(x \mid a) \\ \text{s.t.} \quad & \int \left[U(s(x) - c(a)) \right] dF(x \mid a) \geq \overline{H} \\ a \in \arg\max_{a'} \quad & \int \left[U(s(x)) - c(a') \right] dF(x \mid a') \end{aligned}$$

- Considerably more difficult, because the *incentive compatibility*, IC, constraint is no longer an inequality constraint, but an abstract constraint requiring the value of a function,

$$\int \left[U(s(x)) - c(a') \right] dF(x \mid a'),$$

  to be highest when evaluated at $a' = a$.
- Difficult to make progress on this unless we take some shortcuts.

The First-Order Approach

- The standard shortcut is the "*first-order approach*,".
- It involves replacing the IC constraint with the first-order conditions of the agent, that is, with

$$\int U(s(x)) f_a(x \mid a) dx = c'(a).$$

- Why is this a big assumption?
  - *Incorrect argument*: suppose that

  $$\max_{a'} \int \left[U(s(x)) - c(a') \right] dF(x \mid a')$$

  is strictly concave
  - Why is this argument in correct?
- The first-order approach is a very strong assumption and often invalid.
- **Special care necessary.**

Solution to the Basic Moral Hazard Problem

- Now using the first-order approach the principal's problem becomes

$$\min_{\lambda,\mu} \max_{s(x),a} \mathcal{L} = \int \{V(x - s(x)) \quad + \lambda \left[U(s(x)) - c(a) - \overline{H}\right] \qquad + \mu \left[U(s(x)) \frac{f_a(x \mid a)}{f(x \mid a)} - c'(a)\right]\Big\} f(x \mid a) dx$$

- Again carrying out point-wise maximization with respect to $s(x)$:

$$\begin{aligned} 0 &= \frac{\partial \mathcal{L}}{\partial s(x)} \\ &= -V'(x - s(x)) + \lambda U'(s(x)) + \mu U'(s(x)) \frac{f_a(x \mid a)}{f(x \mid a)} \text{ for all } x \end{aligned}$$

Solution to the Basic Moral Hazard Problem (continued)

- Therefore:

$$\frac{V'(x - s(x))}{U'(s(x))} = \lambda + \mu \frac{f_a(x \mid a)}{f(x \mid a)}. \tag{2}$$

- Identical to the full information case ((1) above) when $\mu = 0$. Why?
- Also, we must have $\lambda > 0$. Why?

Incentive-Insurance Tradeoff Again

- Can we have perfect risk sharing?
- This would require

$$V'(x - s(x))/U'(s(x)) = \text{constant}.$$

- Since V' is constant, this is only possible if U' is constant.
- Since the agent is risk-averse, so that U is strictly concave, this is only possible if $s(x)$ is constant.
- But if $s(x)$ is constant and effort is costly, the incentive compatibility constraint will be violated (unless the optimal contract asks for $a = 0$).

Incentive-Insurance (continued)

- Another extreme case

$$s(x) = x - s_0.$$

- Why is this an extreme case?
- What are the costs?

A Little Paradox

- Where is the uncertainty?
- Since the IC constraint is satisfied, once the agent signs to contract $s(x)$, there is *no uncertainty* about action choice a.
- But remuneration of the agent depends on x, and in particular it is lower when x is lower (more on this below).
- Why is this?
- Because *threat of punishment ex post* necessary for ex ante effort choices.

More on the Paradox

- Therefore, no need for the principal to draw inferences about the effort choice a from the realizations of x.
- But the optimal way of incentivizing the agent has many similarities to an optimal signal extraction problem.
- Consider the following maximum likelihood estimation problem:
  - we know the distribution of x conditional on a,
  - we observe x,
  - we want to estimate a.
- This is a solution to the maximization problem (for given x)

$$\max_{a'} \ln f(x \mid a').$$

- First-order condition of this problem$\longrightarrow$

$$\frac{f_a(x \mid a(x))}{f(x \mid a(x))} = 0$$

where $a(x)$ is the solution to this first-order condition.

Back to the Optimal Contract

- Let the level of effort that the principal wants to implement be $\bar{a}$.
- Then the optimal contract solves:

$$\frac{V'(x - s(x))}{U'(s(x))} = \lambda + \mu \frac{f_a(x \mid \bar{a})}{f(x \mid \bar{a})}.$$

- If $a(x) > \bar{a}$, then

$$f_a(x \mid \bar{a}) / f(x \mid \bar{a}) > 0$$

  so V' / U' has to be greater, which means that U' has to be lower.
- Therefore $s(x)$ must be increasing in x.
- Intuitively, when the realization of output is *good news* relative to what was expected, the agent is rewarded, when it is *bad news*, he is punished.

Form of Performance Contracts

- Intuitively $s(x)$ should be increasing.
- Is it? **Example:**

$$a \in \{a_H, a_L\}$$

$$F(x \mid a_H) = \begin{cases} 4 & \text{with probability } \frac{1}{2} \\ 2 & \text{with probability } \frac{1}{2} \end{cases}$$

$$F(x \mid a_L) = \begin{cases} 3 & \text{with probability } \frac{1}{2} \\ 1 & \text{with probability } \frac{1}{2} \end{cases}$$

- The agent has an arbitrary strictly concave utility function.

Form of Performance Contracts (continued)

- An optimal contract that achieves full risk sharing:

$$\begin{aligned} s(2) &= s(4) = \overline{H} + c(a_H) \\ s(1) &= s(3) = -K \end{aligned}$$

  for K sufficiently large number.

- Therefore

$$s(3) < s(2).$$

Form of Performance Contracts (continued)

- Is this special because of the discrete distribution of x.
- No; consider a continuous distribution with peaks at $\{2, 4\}$ for $a = a_H$ and $\{1, 3\}$ for $a = a_L$.
- So how can we ensure that $s(x)$ is increasing in x?
- A sufficient condition for $s(x)$ to be increasing is that higher values of x are "good news" about a

$$\text{i.e., } \frac{f_a(x \mid a)}{f(x \mid a)} \text{ is increasing in } x$$

$$\Leftrightarrow \frac{f(x \mid a_1)}{f(x \mid a_2)} \text{ is increasing in } x \text{ for } a_1 > a_2$$

This is referred to as the *Monotone Likelihood Ratio Property* (MLRP).

Form of Performance Contracts (continued)

- MLRP implies that

$$\int xf(x \mid a_1)dx > \int xf(x \mid a_2)dx \text{ for } a_1 > a_2.$$

- Therefore, the expected value of x to increase with the level of effort is necessary but not sufficient for performance contracts to be increasing.
- Intuition: MLRP ensures that higher output levels correspond to better news about the level of effort that the agent must have exerted.
- The counter example above violated this principle.

Sufficient Statistics

- Another implication of the informational perspective: the *sufficient statistic* result.
- Imagine that in addition to x, the principal observes another signal of the agent's effort, y, in the sense that y is a random variable with distribution

$$G\left(y \mid a\right).$$

- The principal does not care about y per se, and still wants to maximize $\mathbb{E}(V(x-s))$.
- Question: should the principal should offer a contract $s(x, y)$ which depends (non-trivially) on the signal y as well as the output x?
- Answer:
    - *yes*, if y helps reduce noise or yields extra information on a, and
    - *no* if x is a sufficient statistic for (x, y) in the estimation of a.

Sufficient Statistics (continued)

- Recall that a statistic T is a sufficient statistic for some family of random variables $\mathcal{F}$ in estimating a parameter $\theta \in \Theta$ if and only if the marginal distribution of θ conditional on T and $\mathcal{F}$ coincide, that is,

$$f(\theta \mid T) = f(\theta \mid \mathcal{F}) \text{ for all } \theta \in \Theta.$$

- Now consider the first-order conditions for choosing $s(x, y)$.
- As before,

$$\frac{V'(x - s(x, y))}{U'(s(x, y))} = \lambda + \mu \frac{f_a(x, y \mid a)}{f(x, y \mid a)}$$

- Question: can we express $s(x, y) = S(x)$ for all x and y?
- This is equivalent to

$$\frac{f_a(x, y \mid a)}{f(x, y \mid a)} = k(x \mid a) \text{ for all } x \text{ and } y$$

for some function k.

Sufficient Statistics (continued)

- What does this condition mean?
- The condition

$$\frac{f_a(x, y \mid a)}{f(x, y \mid a)} = k(x \mid a) \; \forall \, x, y$$

  is equivalent to

$$f(x, y \mid a) = g(x, y) h(x \mid a)$$

  (why?)
- This condition, in turn, means that *conditional on* x, y has no additional information on a, or using Bayes' rule

$$f(a \mid x, y) = f(a \mid x)$$

  that is, x is a *sufficient statistic* for (x, y) with respect to inferences about a.
- The implication is the important suggested result:
  - the optimal contract conditional on x and y, $s(x, y)$, will not use y if and only if x is a sufficient statistic for (x, y) with respect to a.

Moral Hazard and Limited Liability

- *Limited liability constraint*, so that negative salaries are not allowed, i.e.,

$$s(x) \geq 0.$$

- Canonical moral hazard problem becomes

$$\max_{s(x),a} \int V(x - s(x)) dF(x \mid a)$$

subject to

$$\begin{aligned} & \int [U(s(x) - c(a))] \, dF(x \mid a) \geq \overline{H} \\ a \quad \in \quad & \arg\max_{a'} \int [U(s(x)) - c(a')] \, dF(x \mid a') \\ & s(x) \geq 0 \text{ for all } x \end{aligned}$$

Moral Hazard and Limited Liability (continued)

- Let us again use the first-order approach and also assign multiplier $\eta(x)$ to the last set of constraints.
- Then

$$V'(x - s(x)) = \left[\lambda + \mu \frac{f_a(x \mid a)}{f(x \mid a)}\right] U'(s(x)) + \eta(x).$$

- If $s(x)$ was going to be positive for all x in any case, the multiplier for the last set of constraints, $\eta(x)$, would be equal to zero, and the problem would have an identical solution to before.
- However, if, previously, $s(x) < 0$ for some x, the structure of the solution has to change.
- One can, for example, increase all of the $s(x)$ so that need none of it is negative (though this will not be optimal in general).

Moral Hazard and Limited Liability (continued)

- Implication: in this case the participation constraint will no longer be binding, thus $\lambda = 0$.
- General intuition that without limited liability constraints, there are no rents ($\lambda > 0$)
- But with limited liability there will be typically be rents, making the participation constraint slack.

Example with Limited Liability

- Suppose that effort takes two values $a \in \{a_L, a_H\}$.
- And output also takes only two values: $x \in \{0, 1\}$, with

$$F(x \mid a_H) = 1 \quad \text{with probability } 1$$

$$F(x \mid a_L) = \begin{cases} 1 & \text{with probability } q \\ 0 & \text{with probability } 1-q \end{cases}$$

- Normalize $\bar{H}$ and $c(a_L)$ to zero, and assume

$$c(a_H) = c_H < 1 - q.$$

- Also assume that both the agent and the principal are risk neutral.

Example with Limited Liability (continued)

- Without the limited liability constraint, the problem is trivial.
- Sincet $c(a_H) = c_H < 1 - q$, high effort is optimal and will be implemented by solving the problem

$$\min_{s(0),s(1)} s(1)$$

subject to

$$\begin{aligned} s(1) - c_H &\geq qs(1) + (1-q)s(0) \\ s(1) - c_H &\geq 0 \end{aligned}$$

where $s(0)$ and $s(1)$ are the payments to the agent conditional on the outcome.

- Solution:

$$s(1) = c_H \text{ and } s(0) \leq -\frac{q}{1-q}c_H.$$

- Participation constraint is binding.

Example with Limited Liability (continued)

- Now impose limited our ability and suppose high effort will be implemented.
- Then

$$\min_{s(0),s(1)} s(1)$$

subject to

$$\begin{aligned} s(1) - c_H &\geq qs(1) + (1-q)s(0) \\ s(1) - c_H &\geq 0 \\ s(0) &\geq 0 \end{aligned}$$

Example with Limited Liability (continued)

- Solution:

$$\begin{aligned} s(0) &= 0 \\ s(1) &= \frac{c_H}{1-q} \end{aligned}$$

- Participation constraint is slack.
- The agent is receiving *a rent* given by

$$\text{rent} = \frac{q}{1-q} c_H$$

Example with Limited Liability (continued)

- The presence of rents may also distort the choice of effort.
- Return to high effort for the principal:

$$\text{Return}_H = 1 - \frac{q}{1-q} c_H$$

- Return to low effort (implemented by paying $s(0) = s(1) = 0$):

$$\text{Return}_L = q$$

- This can be greater than Return_H.
- In order to reduce the rents that the agent receives the principal may *distort* the structure of effort.

Robustness of Contracts

- Basic moral hazard problem leads to contracts that are potentially very complex than nonlinear.
- Recall example of nonmonotonic contracts.
- Are these "robust"?
- What does "robust" mean?

Robustness (continued)

- Holmstrom and Milgrom: no manipulation in *dynamic principal-agent* problems
- Consider a model in continuous time.
- The interaction between the principal and the agent take place over an interval normalized to $[0, 1]$.
- The agent chooses an effort level $a_t \in A$ at each instant after observing the relaxation of output up to that instant.
- The output process is given by the continuous time random walk, that is, the following *Brownian motion process*:

$$dx_t = a_t dt + \sigma dW_t$$

where W is a standard Brownian motion (Wiener process).

- This implies that its increments are independent and normally distributed, that is, $W_{t+\tau} - W_t$ for any t and τ is distributed normally with variance equal to τ.

Robustness (continued)

- Let

$$X^t = (x_\tau ; 0 \leq \tau < t)$$

  be the entire history of the realization of the increments of output x up until time t (or alternatively a "sample path" of the random variable x).
- Effort choice

$$a_t : X^t \longrightarrow A.$$

- Similarly, the principal also observes the realizations of the increments (though obviously not the effort levels and the realizations of W_t).
- Therefore, contract

$$s_t : X^t \longrightarrow \mathbb{R}.$$

Robustness (continued)

- Holmstrom and Milgrom assume that utility of the agent is

$$u\left(C_1 - \int_0^1 a_t d_t\right)$$

- C_1 is consumption at time $t = 1$.
- Two special assumptions:
  1. the individual only derives utility from consumption at the end (at time $t = 1$) and
  2. the concave utility function applies to consumption minus the total cost of effort between 0 and 1.
- A further special assumption constant absolute risk aversion (CARA) utility:

$$u(z) = -\exp(-rz) \tag{3}$$

Robustness: Key Result

- In this case, optimal contracts are only a function of (cumulative) output x_1 and are *linear*.
- Independent of the exact sample path leading to the cumulative output.
- Moreover, in response to this contract the optimal behavior of the agent is to choose a constant level of effort, which is also independent of the history of past realizations of the stochastic shock.
- Loose intuition: with any nonlinear contract there will exist an event, i.e., a sample path, after which the incentives of the agent will be distorted, whereas the linear contract achieves a degree of "robustness".

Linear Contracts

- Motivated by this result, many applied papers look at the following *static* problem:
  1. The principal chooses a linear contract, of the form $s = \alpha + \beta x$.
  2. The agents chooses $a \in A \equiv [0, \infty]$.
  3. $x = a + \varepsilon$ where $\varepsilon \sim \mathcal{N}\left(0, \sigma^2\right)$
- The principal is risk neutral
- The utility function of the agent is

$$U(s, a) = -\exp\left(-r\left(s - c\left(a\right)\right)\right)$$

with

$$c(a) = ca^2/2$$

Linear Contracts (continued)

- Loose argument: a linear contract is approximately optimal here.
- Is this true?

Linear Contracts (continued)

- Even if linear contracts are not optimal in the static model, they are attractive for their simplicity and can be justified as thinking of the dynamic model.
- They are also easy to characterize.
- The first-order approach works in this case.
- The maximization problem of the agent is

$$\begin{aligned} & \max_a \mathbb{E}\left\{-\exp\left(-r\left(s\left(a\right)-c\left(a\right)\right)\right)\right\} \\ = \quad & \max_a \left\{-\exp\left(-r\mathbb{E}s\left(a\right)+\frac{r^2}{2}\mathrm{Var}\left(s\left(a\right)\right)-rc\left(a\right)\right)\right\} \end{aligned}$$

- Where is the second line coming from?

Linear Contracts (continued)

- Therefore, the agent's problem is

$$\max_a \left\{ \mathbb{E}s\left(a\right) - \frac{r}{2}\text{Var}\left(s\left(a\right)\right) - \frac{c}{2}a^2 \right\}$$

- Substituting for the contract:

$$\max_a \beta a - \frac{c}{2}a^2 - \frac{r}{2}\beta^2\sigma^2$$

- The first-order condition for the agent's optimal effort choice is:

$$a = \frac{\beta}{c}$$

Linear Contracts (continued)

- The principal will then maximize

$$\max_{a,\alpha,\beta} \mathbb{E}\left(\left(1-\beta\right)\left(a+\varepsilon\right)-\alpha\right)$$

subject to

$$\begin{aligned} a &= \frac{\beta}{c} \\ \alpha + \frac{\beta^2}{2}\left(\frac{1}{c} - r\sigma^2\right) &\geq \bar{h} \end{aligned}$$

- First equation is the incentive compatibility constraint in the second is the participation constraint (with $\bar{h} = -\ln\left(-\bar{H}\right)$).

Linear Contracts: Solution

- Solution:

$$\beta^* = \frac{1}{1 + rc\sigma^2} \tag{4}$$

and

$$\alpha^* = \bar{h} - \frac{1 - rc\sigma^2}{2c^2 \left(1 + rc\sigma^2\right)^2},$$

- Because negative salaries are allowed, the participation constraint is binding.
- The equilibrium level of effort is

$$a^* = \frac{1}{c \left(1 + rc\sigma^2\right)}$$

- Always lower than the first-best level of effort which is $a^{fb} = 1/c$.

Linear Contracts: Comparative Statics

- Incentives are *lower powered*—i.e., β^* is lower, when
  - the agent is more risk-averse is the agent, i.e., the greater is r,
  - effort is more costly, i.e., the greater is c,
  - there is greater uncertainty, i.e., the greater is σ^2.

Linear Contracts: Sufficient Statistics

- Suppose that there is another signal of the effort

$$z = a + \eta,$$

- η is $\mathcal{N}\left(0, \sigma_\eta^2\right)$ and is independent of ε.
- Let us restrict attention to linear contracts of the form

$$s = \alpha + \beta_x x + \beta_z z.$$

- Note that this contract can also be interpreted alternatively as

$$s = \alpha + \mu w$$

where

$$w = w_1 x + w_2 z$$

is a sufficient statistic derived from the two random variables x and z.

Linear Contracts: Sufficient Statistics (continued)

- Now with this type of contract, the first-order condition of the agent is
$$a = \frac{\beta_x + \beta_z}{c}$$
- Therefore, the optimal contract gives:
$$\beta_x = \frac{\sigma_\eta^2}{\sigma^2 + \sigma_\eta^2 + rc\left(\sigma^2\sigma_\eta^2\right)}$$
and
$$\beta_z = \frac{\sigma^2}{\sigma^2 + \sigma_\eta^2 + rc\left(\sigma^2\sigma_\eta^2\right)}$$
- These expressions show that generally x is not a sufficient statistic for (x, z), and the principal will use information about z as well to determine the compensation of the agent.

Linear Contracts: Sufficient Statistics (continued)

- The exception is when $\sigma_\eta^2 \to \infty$ so that there is almost no information in z regarding the effort chosen by the agent.
- In this case, $\beta_z \to 0$ and $\beta_x \to \beta^*$ as given by (4), so in this case x becomes a sufficient statistic.

Evidence

- The evidence on the basic principal-agent model is mixed.
- Evidence in favor of the view that *incentives matter.*
- Lazear: data from a large auto glass installer, high incentives lead to more effort.
    - For example, Lazear's evidence shows that when this particular company went from fixed salaries to piece rates productivity rose by 35% because of greater effort by the employees (the increase in average wages was 12%), but part of this response might be due to selection, as the composition of employees might have changed.

Evidence (continued)

- Similar evidence is reported in other papers.
  - For example, Kahn and Sherer, using the personnel files of a large company, show that employees (white-collar office workers) whose pay depends more on the subjective evaluations obtain better evaluations and are more productive.
- Incentives also matter in extreme situations.
  - John McMillan on the responsibility system in Chinese agriculture
  - Ted Groves similar effects from the Chinese industry.

Evidence (continued)

- But various pieces of evidence that high-powered incentives might backfire.
  - Ernst Fehr and Simon Gachter: that incentive contracts might destroy voluntary cooperation.
- More standard examples are situations in which high-powered incentives lead to distortions that were not anticipated by the principals.
  - e.g., consequences of Soviet incentive schemes specifying "performance" by number of nails or the weight of the materials used, leading to totally unusable products.
- Similar results from the negative effects of nonlinear high-powered performance contracts; *gaming of the system*
  - Paul Oyer,
  - Pascal Courty and Gerard Marschke

Evidence (continued)

- The most negative evidence against the standard moral hazard models is that they do not predict the form of performance contracts.
- Prendergast: there is little association between riskiness and noisiness of tasks and the types of contracts when we look at a cross section of jobs.
- In many professions performance contracts are largely absent.
- Why could this be?
    - Again robustness.
    - Multitask issues.
    - Career concerns.

Multitask Models

- Let us now modify the above linear model so that there are two efforts that the individual chooses, a_1 and a_2, with a cost function

$$c(a_1, a_2)$$

which is increasing and convex as usual.
- These efforts lead to two outcomes:

$$x_1 = a_1 + \varepsilon_1$$

and

$$x_2 = a_2 + \varepsilon_2,$$

where ε_1 and ε_2 could be correlated.
- The principal cares about both of these inputs with potentially different weights, so her return is

$$\phi_1 x_1 + \phi_2 x_2 - s$$

where s is the salary paid to the agent.

Multitask Models (continued)

- Main difference: only x_1 is observed, while x_2 is unobserved.
- Example: the agent is a home contractor where x_1 is an inverse measure of how long it takes to finish the contracted work, while x_2 is the quality of the job, which is not observed until much later, and consequently, payments cannot be conditioned on this.
- Another example: the behavior of employees in the public sector, where quality of the service provided to citizens is often difficult to contract on.
- High-powered incentives may distort the composition of effort.

Multitask Models: Solution

- Let us focus on linear contracts of the form

$$s(x_1) = \alpha + \beta x_1$$

since x_1 is the only observable output.

- The first-order condition of the agent now gives:

$$\begin{aligned} \beta &= \frac{\partial c(a_1, a_2)}{\partial a_1} \\ 0 &= \frac{\partial c(a_1, a_2)}{\partial a_2} \end{aligned} \tag{5}$$

- So if

$$\frac{\partial c(a_1, a_2)}{\partial a_2} > 0$$

whenever $a_2 > 0$, then the agent will choose $a_2 = 0$, and there is no way of inducing him to choose $a_2 > 0$.

Multitask Models: Solution (continued)

- However, suppose that

$$\frac{\partial x\left(a_1, a_2 = 0\right)}{\partial a_2} < 0.$$

- Then, without "incentives" the agent will exert some positive effort in the second task.
- Now providing stronger incentives in task 1 can *undermine* the incentives in task 2;
    - this will be the case when the two efforts are substitutes, i.e.,

$$\partial^2 c\left(a_1, a_2\right) / \partial a_1 \partial a_2 > 0.$$

Multitask Models: Solution (continued)

- More formally, imagine that the first-order conditions in (5) have an interior solution (why is an interior solution important?).
- Then differentiate these two first-order conditions with respect to β.
- Using the fact that these two first-order conditions correspond to a maximum (i.e., the second order conditions are satisfied), we obtain

$$\frac{\partial a_1}{\partial \beta} > 0.$$

- This has the natural interpretation that high-powered incentives lead to stronger incentives as the evidence discussed above suggests.

Multitask Models: Solution (continued)

- However, in addition provided that $\partial^2 c(a_1, a_2)/\partial a_1 \partial a_2 > 0$, we also have
$$\frac{\partial a_2}{\partial \beta} < 0,$$
- Therefore, high-powered incentives in one task adversely affect the other task.
- What are the implications for interpreting empirical evidence?

Multitask Models: Solution (continued)

- What about the optimal contract?
- If the second task is sufficiently important for the principal, then she will "shy away" from high-powered incentives; if you are afraid that the contractor will sacrifice quality for speed, you are unlikely to offer a contract that puts a high reward on speed.
- In particular, the optimal contract will have a slope coefficient of

$$\beta^{**} = \frac{\phi_1 - \phi_2 \left(\partial^2 c(a_1, a_2) / \partial a_1 \partial a_2\right) / \left(\partial^2 c(a_1, a_2) / \partial a_2^2\right)}{1 + r\sigma_1^2 \left(\partial^2 c(a_1, a_2) / \partial a_1^2 - \left(\partial^2 c(a_1, a_2) / \partial a_1 \partial a_2\right)^2 / \partial^2 c(a_1, a_2)\right.}$$

- As expected β^{**} is declining in ϕ_2 (the importance of the second task) and in $-\partial^2 c(a_1, a_2) / \partial a_1 \partial a_2$ (degree of substitutability between the efforts of the two tasks).

Career Concerns

- "Career concerns" $\approx$ reasons to exert effort unrelated to current compensation.
- These could be social effects.
- Or more standard: anticipation of future compensation
- Question: is competition in market for managers sufficient to give them sufficient incentives without agency contracts?

The Basic Model of Career Concerns

- Basic model due to Holmstrom.
- The original Holmstrom model is infinite horizon, but useful to start with a 2-period model.
- Output produced is equal to

$$x_t = \underbrace{\eta}_{\text{ability}} + \underbrace{a_t}_{\text{effort}} + \underbrace{\varepsilon_t}_{\text{noise}} \quad t = 1, 2$$

- Since the purpose is to understand the role of career concerns, let us go to the extreme case where there are no performance contracts.
- As before $a_t \in [0, \infty)$.

Career Concerns (continued)

- Also assume that

$$\varepsilon_t \sim \mathcal{N}(0, 1/h_\varepsilon)$$

where h is referred to as "precision" (inverse of the variance)

- Also, the prior on η has a normal distribution with mean m_0, i.e.,

$$\eta \sim \mathcal{N}(m_0, 1/h_0)$$

and $\eta, \varepsilon_1, \varepsilon_2$ are independent.

- What does it mean for the prior to have distribution?

Career Concerns (continued)

- Differently from the basic moral hazard model this is an *equilibrium* model, in the sense that there are other firms out there who can hire this agent. This is the source of the career concerns.
- Loosely speaking, a higher perception of the market about the ability of the agent, η, will translate into higher wages.
- This class of models are also referred to as "signal jamming" models, since the agent might have an interest in working harder in order to improve the perception of the market about his ability.

Career Concerns: Timing and Information Structure

- Information structure:
  - the firm, the worker, and the market all share prior belief about η (thus there is no asymmetric information and adverse selection; is this important?).
  - they all observe x_t each period.
  - only worker sees a_t (moral hazard/hidden action).
- In equilibrium firm and market correctly conjecture a_t (Why?)
  - $\rightarrow$along-the-equilibrium path despite the fact that there is hidden action, information will stay symmetric.
- The labor market is competitive, and all workers are paid their expected output.
- Recall: no contracts contingent on output (and wages are paid at the beginning of each period).

Career Concerns: Wage Structure

- Competition in the labor market$\longrightarrow$ the wage of the worker at a time t is equal to the mathematical expectation of the output he will produce given the history of its outputs

$$w_t(x^{t-1}) = \mathbb{E}(x_t \mid x^{t-1})$$

where $x^{t-1} = \{x_1, ..., x_{t-1}\}$ is the history of his output realizations.

- Alternatively,

$$\begin{aligned} w_t(x^{t-1}) &= \mathbb{E}(x_t \mid x^{t-1}) \\ &= \mathbb{E}(\eta \mid x^{t-1}) + a_t(x^{t-1}) \end{aligned}$$

where $a_t(x^{t-1})$ is the effort that the agent will exert given history x^{t-1}

- Important: $a_t(x^{t-1})$ is perfectly anticipated by the market along the equilibrium path.

Career Concerns: Preferences

- Instantaneous utility function of the agent is

$$u(w_t, a_t) = w_t - c(a_t)$$

- With horizon equal to T, preferences are

$$U(w, a) = \sum_{t=1}^{T} \beta^{t-1} [w_t - c(a_t)]$$

- For now $T = 2$.
- Finally,

$$c'(\cdot) > 0, \quad c''(\cdot) > 0$$
$$c'(0) = 0$$

- First best level of effort a^{fb} again solves

$$c'(a^{fb}) = 1.$$

Career Concerns: Summary

- Recall that all players, including the agent himself, have prior on $\eta \sim \mathcal{N}(m_0, 1/h_0)$
- So the world can be summarized as:

$$\text{period 1: } \begin{cases} \text{wage } w_1 \\ \text{effort } a_1 \text{ chosen by the agent (unobserved)} \\ \text{output is realized } x_1 = \eta + a_1 + \varepsilon_1 \end{cases}$$

$$\text{period 2: } \begin{cases} \text{wage } w_2(x_1) \\ \text{effort } a_2 \text{ chosen} \\ \text{output is realized } x_2 = \eta + a_2 + \varepsilon_2 \end{cases}$$

- Appropriate equilibrium concept: Perfect Bayesian Equilibrium.

Career Concerns: Equilibrium

- Backward induction immediately implies

$$a_2^* = 0$$

- Why?
- Therefore:

$$\begin{aligned} w_2(x_1) &= \mathbb{E}(\eta \mid x_1) + a_2(x_1) \\ &= \mathbb{E}(\eta \mid x_1) \end{aligned}$$

- The problem of the market is the estimation of η given information $x_1 = \eta + a_1 + \varepsilon_1$.
- The only difficulty is that x_1 depends on first period effort.
- In equilibrium, the market will anticipate the correct level of effort a_1.

Career Concerns: Equilibrium (continued)

- Let the conjecture of the market be $\bar{a}_1$.
- Define

$$z_1 \equiv x_1 - \bar{a}_1 = \eta + \varepsilon_1$$

  as the deviation of observed output from this conjecture.
- Once we have z_1, standard normal updating formula implies that

$$\eta \mid z_1 \sim \mathcal{N}\left(\frac{h_0 m_0 + h_\varepsilon z_1}{h_0 + h_\varepsilon}, h_0 + h_\varepsilon\right)$$

- Interpretation: we start with prior m_0, and update η according to the information contained in z_1. How much weight we give to this new information depends on its precision relative to the precision of the prior. The greater its h_ε relative to h_0, the more the new information matters.
- Also important: the variance of this posterior will be less than the variance of both the prior and the new information (Why?).

The Basic Model of Career Concerns: Equilibrium (continued)

- Combining these observations

$$\mathbb{E}(\eta \mid z_1) = \frac{h_0 m_0 + h_\varepsilon z_1}{h_0 + h_\varepsilon}$$

- Or equivalently:

$$\mathbb{E}(\eta \mid x_1) = \frac{h_0 m_0 + h_\varepsilon (x_1 - \bar{a}_1)}{h_0 + h_\varepsilon}$$

- Therefore, equilibrium wages satisfy

$$w_2(x_1) = \frac{h_0 m_0 + h_\varepsilon (x_1 - \bar{a}_1)}{h_0 + h_\varepsilon}$$

- To complete the characterization of equilibrium we have to find the level of a_1 that the agent will choose as a function of $\bar{a}_1$, and make sure that this is indeed equal to $\bar{a}_1$, that is, this will ensure a *fixed point*.

Career Concerns: Equilibrium (continued)

- Let us first write the optimization problem of the agent:

$$\max_{a_1} [w_1 - c(a_1)] + \beta[\mathbb{E}\{w_2(x_1) \mid \bar{a}_1\}]$$

where we have used the fact that $a_2 = 0$.

- Substituting from above and dropping w_1:

$$\max_{a_1} \beta \, \mathbb{E} \left\{ \frac{h_0 m_0 + h_\varepsilon (x_1 - a_1)}{h_0 + h_\varepsilon} \middle| \bar{a}_1 \right\} - c(a_1)$$

- Important: both η and ε_1 are uncertain to the agent as well as to the market.

Career Concerns: Equilibrium (continued)

- Therefore

$$\max_{a_1} \beta \mathbb{E} \left\{ \frac{h_0 m_0 + h_\varepsilon(\eta + \varepsilon_1 + a_1 - \bar{a}_1)}{h_0 + h_\varepsilon} \middle| a_1 \right\} - c(a_1)$$

- And since a_1 is not stochastic (the agent is choosing it), we have

$$\max_{a_1} \beta \frac{h_\varepsilon}{h_0 + h_\varepsilon} a_1 - c(a_1) + \beta \mathbb{E} \left\{ \frac{h_0 m_0 + h_\varepsilon(\eta + \varepsilon_1 - \bar{a}_1)}{h_0 + h_\varepsilon} \right\}$$

- The first-order condition is:

$$c'(a_1^*) = \beta \frac{h_\varepsilon}{h_0 + h_\varepsilon} < 1 = c'(a_{fb})$$

- This does not depend on $\bar{a}_1$, so the fixed point problem is solved immediately.
- First result: equilibrium effort is always less than first this.

Career Concerns: Equilibrium (continued)

- Why?: because there are two "leakages" (increases in output that the agent does not capture): the payoff from higher effort only occurs next period, therefore its value is discounted to β, and the agent only gets credit for a fraction $h_\varepsilon / (h_0 + h_\varepsilon)$ of her effort, the part that is attributed to ability.
- The characterization of the equilibrium is completed by imposing $\bar{a}_1 = a_1^*$
- This was not necessary for computing a_1^*, but is needed for computing the equilibrium wage w_1.
- Recall that

$$\begin{aligned} w_1 &= \mathbb{E}(y_1 \mid \text{prior}) \\ &= \mathbb{E}(\eta) + \bar{a}_1 \\ &= m_0 + a_1^* \end{aligned}$$

Career Concerns: Comparative Statics

- We immediately obtain:

$$\frac{\partial a_1^*}{\partial \beta} > 0$$

$$\frac{\partial a_1^*}{\partial h_\varepsilon} > 0$$

$$\frac{\partial a_1^*}{\partial h_0} < 0$$

- Greater β means that the agent discounts the future less, so exerts more effort because the first source of leakage is reduced.
- Greater h_ε implies that there is less variability in the random component of performance. This, from the normal updating formula, implies that any given increase in performance is more likely to be attributed to ability, so the agent is more tempted to jam the signal by exerting more effort.
- The intuition for the negative effect of h_0 is similar.

Multiperiod Career Concerns

- Considered the same model with three periods.
- This model can be summarized by the following matrix

$$\begin{array}{ll} w_1 & a_1^* \\ w_2(x_1) & a_2^* \\ w_3(x_1, x_2) & a_3^* \end{array}$$

- With similar analysis to before, the first-order conditions for the agent are

$$c'(a_1^*) = \beta \frac{h_\varepsilon}{h_0 + h_\varepsilon} + \beta^2 \frac{h_\varepsilon}{h_0 + 2h_\varepsilon}$$

$$c'(a_2^*) = \beta \frac{h_\varepsilon}{h_0 + 2h_\varepsilon}.$$

Multiperiod Career Concerns (continued)

- First result:
$$a_1^* > a_2^* > a_3^* = 0.$$
- Why?
- More generally, in the T period model, the relevant first-order condition is
$$c'(a_t^*) = \sum_{\tau=t}^{T-1} \beta^{\tau-t+1} \frac{h_\varepsilon}{h_0 + \tau h_\varepsilon}.$$

Multiperiod Career Concerns: Overeffort

- With T sufficiently large, it can be shown that there exists a period $\overline{\tau}$ such that

$$a^*_{t<\overline{\tau}} \geq a_{fb} \geq a^*_{t>\overline{\tau}}.$$

- In other words, workers work too hard when young and not hard enough when old—
  - compare assistant professors to tenured faculty.
  - important: these effort levels depend on the horizon (time periods), but not on past realizations.

Multiperiod Career Concerns: Generalizations

- Similar results hold when ability is not constant, but evolves over time (as long as it follows a normal process).
- For example, we could have

$$\eta_t = \eta_{t-1} + \delta_t$$

with

$$\begin{array}{rcl} \eta_0 & \sim & \mathcal{N}(m_0, 1/h_0) \\ \delta_t & \sim & \mathcal{N}(0, 1/h_\delta) \ \forall \ t \end{array}$$

- In this case, it can be shown that the updating process is stable, so that the process and therefore the effort level converge, and in particular as $t \to \infty$, we have

$$a_t \longrightarrow \bar{a}$$

but as long as $\beta < 1$, $\bar{a} < a_{fb}$.
- Also, the same results apply when the agent knows his ability.
    - Why is this? In what ways it special?