Some Notes on Moral Hazard

John Morgan
University of California at Berkeley

1 Preliminaries

Up until this point, we have been concerned mainly with the problem of private information on the part of the agent, information that was economically valuable for the principal in undertaking some decision.

In this section, we will be worried about hidden action on the part of the agent. Problems here center on setting the right incentive scheme for the agent to choose an action consistent with the principal’s objectives. It will turn out that general conclusions are difficult to draw for this class of problems. These notes will highlight some simple and prominent special models as well as applications.

2 Two Performance Outcomes

Suppose that performance, which may be thought of as output $q$, can take on only 2 outcomes: $q \in \{0, 1\}$. When $q = 1$, the performance can be thought of as a success and when $q = 0$, it can be thought of as a failure.

When $A$ exerts effort $e$, the probability of success is $p(e)$. We make the following standard assumptions (which help guarantee interior solutions):

1. $p(\cdot)$ is strictly increasing and strictly concave.
2. $p(0) = 0$, $p(\infty) = 1$
3. $p'(0) > 1$

$P$’s utility function is given by $V(q - w)$ where $w$ is the wage paid by $P$ to $A$. We assume $V' > 0$ and $V'' \leq 0$.

$A$’s utility function is (conveniently) separable between wage and effort. That is

$$U(w, e) = u(w) - c(e)$$

where $u' > 0$, $u'' \leq 0$ and $c' > 0$, $c'' \geq 0$.

We’ll make life even simpler and assume that $c(e) = e$. 
2.1 First-Best Contracts

Suppose that $e$ is observable by $P$. Then we can condition contracts on outcomes and effort level. Thus, $P$’s problem is to choose $e, w$ to maximize

$$p(e) V(1 - w) + (1 - p(e)) V(-w)$$

subject to

$$p(e) u(w) + (1 - p(e)) u(w) - e \geq 0$$

The contract $w_0, w_1$ specifies the payout in the event effort $a$ is undertaken and failure or success (resp.) is achieved. $P$ can force this contract on $A$ by specifying punitive transfers in the event that any effort other than $a$ is undertaken.

Normalize $\bar{U} = 0$. Then, the Lagrangian for $P$ is

$$\max_{e, w_i} L = p(e) V(1 - w) + (1 - p(e)) V(-w) + \lambda (p(e) u(w) + (1 - p(e)) u(w) - e)$$

This yields

$$p'(e) (V(1 - w) - V(-w)) + \lambda p'(e) (u(w) - u(w)) - \lambda = 0$$

$$-p(e) V'(1 - w) + \lambda p(e) u'(w) = 0$$

$$-(1 - p(e)) V'(-w) + \lambda (1 - p(e)) u'(w) = 0$$

Put these equations together and we have:

Optimal co-insurance:

$$\frac{V'(1 - w)}{u'(w)} = \frac{V'(-w)}{u'(w)}$$

and optimal effort:

$$p'(e) (V(1 - w) - V(-w)) + \lambda p'(e) (u(w) - u(w)) - \lambda = 0$$  \hspace{1cm} (1)

**Easy case: Risk-neutral principal**

In this case, $V(x) = x$, so

1. Equalize marginal utility of wealth across states: i.e., $u'(w_0) = u'(w_1)$ (also implies equalizing wealth across states)
2. Hit the participation constraint for $A$: $u(w_1) - e \geq 0$

Hence

$$u(w^*) = e^*$$

Finally, use equation (1) to compute optimal effort.

$$p'(e) = \lambda$$

and

$$\lambda = \frac{1}{u'(w^*)}.$$
**Proposition 1** When the principal is risk-neutral, optimal contracting entails paying the agent a fixed wage that just compensates the agent’s cost of effort.

### 2.2 Second Best Contracts

Now suppose that wages cannot be conditioned on $e$. Then, P’s problem is to choose $e, w_0, w_1$ to maximize

$$p(e) V(1 - w_1) + (1 - p(e)) V(-w_0)$$

subject to

$$p(e) u(w_1) + (1 - p(e)) u(w_0) - e \geq 0$$

and

$$e \in \arg\max p(\hat{e}) u(w_1) + (1 - p(\hat{e})) u(w_0) - \hat{e}$$

The first-order condition for the A’s choice of $e$ is

$$p'(e) (u(w_1) - u(w_0)) = 1$$

There is a unique solution for $e$ and we can replace the FOC into P’s optimization. (Note that this replacement, the so-called “first-order approach”, is not kosher in general.)

Thus, P’s new optimization is to choose $e, w_0, w_1$ to maximize

$$p(e) V(1 - w_1) + (1 - p(e)) V(-w_0)$$

subject to

$$p(e) u(w_1) + (1 - p(e)) u(w_0) - e \geq 0$$

and

$$p'(e) (u(w_1) - u(w_0)) = 1$$

**Risk-neutral agent, limited liability**

Since the agent is risk-neutral, then the first-best effort is given by

$$p'(e^*) = 1$$

and the FOC becomes

$$p'(e) (w_1 - w_0) = 1.$$ 

Under first-best effort, this reduces to

$$w_1 - w_0 = 1$$

and, because of limited liability, the cheapest way to do this is to set $w_0 = 0$ and $w_1 = 1$. 

3
Under this contract, the agent’s expected payoff from choosing \( e^* \) is simply

\[
p(e^*) - e^*
\]

or equivalently

\[
\int_0^{e^*} (p'(t) - t) \, dt > 0
\]

since \( p \) is strictly concave.

Thus, inducing the first-best effort level is a feasible contract (although the wage packet is no longer fixed as it was).

Is it optimal?

Fact 1: In any optimal contract, \( w_0 = 0 \).

Reasoning: Only the difference between \( w_1 \) and \( w_0 \) matters for incentive compatibility and only the average matters for IR.

Fact 2: In any optimal contract, \( w_1 = \frac{1}{p'(e)} \)

Reasoning: Follows from satisfying IC along with Fact 1.

Temporarily ignoring IR, the optimal contract reduces to choosing \( e \) to maximize

\[
p(e) \left( 1 - \frac{1}{p'(e)} \right)
\]

This yields

\[
p'(e) \left( 1 - \frac{1}{p'(e)} \right) + p(e) \frac{p''(e)}{p'(e)^2} = 0
\]

Factoring out \( \frac{1}{p'(e)} \) and rewriting this expression one obtains

\[
p'(e)^2 = p'(e) - p(e) \frac{p''(e)}{p'(e)}
\]

Let \( e^{**} \) be the solution to this optimization. Then we have

\[
p'(e^{**}) = 1 - p(e^{**}) \frac{p''(e^{**})}{p'(e^{**})^2}
\]

and since the \( RHS > 1 \), it then follows that \( e^{**} < e^* \).

The optimal effort induced is less than first-best.

**Comments**

Why is the first-best effort still feasible but no longer optimal? The reason can be readily seen from the IR constraint. Recall that the agent’s expected utility when effort \( e \) is induced is

\[
U(e) = p(e) w_1 - e = \frac{p(e)}{p'(e)} - e
\]
Notice that the derivative WRT \(e\) is:

\[
\frac{dU(e)}{de} = \frac{p'(e)^2 - p''(e) p(e)}{(p'(e))^2} - 1
\]

\[
= \frac{-p''(e) p(e)}{(p'(e))^2}
\]

\[
> 0
\]

Hence, the higher the effort, the greater the rents transferred to the agent. Thus, there is an additional cost associated with providing effort incentives in the second best contract that is absent in the first-best contract.

**Applications**

Think of the agent as an owner-manager of the firm and the principal as outside shareholders. Then \(w_1\) is the fraction of inside equity and \(1 - w_1\) the fraction of outside equity. The result is the familiar one—the sale of outside equity dilutes the incentives of the owner manager relative to the case where 100% of the firm is closely held. (Jensen-Meckling, 1976)

Similar idea: This of \(1 - w_1\) as the outstanding debt of the firm (Myers, 1977) then one can have the perverse result that the expected value of the outside debt exhibits a laffer curve. Raising the amount of the outside debt (equivalent to lowering \(e\)) actually raises the default risk and reduces the expected value of the debt held by creditors.

Another variation: The agent can be thought of as a sharecropper and the principal an owner. \(w_1\) is the fraction of the crop yield that the sharecropper gets to retain. The upshot is that land leased out for sharecropping is less productive than that which is closely held. (Stiglitz, 1974).

### 2.3 Costly Monitoring

Now consider a cousin to the standard moral hazard problem. Suppose we cannot contract on outcomes at all (perhaps because they are ill-measured or occur too far into the future. Instead, suppose that, at some cost, \(P\) can observe the effort undertaken by \(A\). \(P\) has two choices: Either pay \(M\) to monitor perfectly or pay zero to monitor with probability (say) equal to 1/2.

Under perfect monitoring, the problem is identical to the first-best up above. \(w^* = e^*\) and \(p'(e^*) = 1\). \(P\) earns

\[
V(e^*) = p(e^*) - w^* - M
\]

Otherwise, the compensation depends on whether the \(P\) detects shirking or not. If he does, clearly \(w_s = 0\) (where \(s\) denotes shirking). Thus, to induce effort level \(e\) requires

\[
w - e \geq \frac{1}{2}w
\]
or
\[ w = 2e \]
It costs twice the effort level to induce working rather than shirking.

What does optimal effort then look like?
\[
\max_e p(e) - 2e
\]

or
\[
p'(e^{**}) = 2
\]
so that less than first-best effort is induced.

Whether to monitor or not then comes down to the cost of \( M \) relative to the loss from the distortion in work effort from not monitoring.

**Comments**
One interpretation of the above result is to view it as a form of efficiency wages (Shapiro-Stiglitz, 1984). \( P \) overpays for effort and cuts back on monitoring in an attempt to induce effort.

### 2.4 Risk Aversion by the Agent

Now consider the case where there is no limited liability and the agent is risk averse. \( P \)'s problem is
\[
\max_{e,w_0,w_1} p(e) (1 - w_1) - (1 - p(e)) w_0
\]
subject to
\[
p(e) u(w_1) + (1 - p(e)) u(w_0) - e \geq 0
\]
and
\[
p'(e) (u(w_1) - u(w_0)) = 1
\]
Letting \( \lambda \) be the Lagrange multiplier associated with IR and \( \mu \) the multiplier associated with IC, one obtains
\[
p'(e) (1 - w_1 + w_0) + \lambda (p'(e) (u(w_1) - u(w_0)) - 1) + \mu (p''(e) (u(w_1) - u(w_0))) = 0
\]
as well as
\[
-p(e) + \lambda p(e) u'(w_1) + \mu p'(e) u'(w_1) = 0
\]
\[
-(1 - p(e)) + \lambda ((1 - p(e)) u'(w_0)) - \mu (p'(e) u'(w_0)) = 0
\]
Arranging the latter two conditions in the usual way yields
\[
\frac{1}{u'(w_1)} = \lambda + \mu \frac{p'(e)}{p(e)}
\]
and
\[
\frac{1}{u'(w_0)} = \lambda - \mu \frac{p'(e)}{1 - p(e)}
\]
If $\mu = 0$, then this gives back first-best risk sharing. However, $\mu > 0$. Hence we have

$$\frac{1}{w'(w_1)} > \frac{1}{w'(w_0)}$$

which has the natural interpretation: $P$ punishes $A$ for the bad outcome and rewards for the good outcome.

**Comments**

Unlike the case of adverse selection, $P$ perfectly predicts $A$’s effort and knows the outcome is pure luck. Nonetheless, $P$ rewards and punishes on the basis of the luck part of the outcome to create effort incentives.

### 3 Multitask Moral Hazard

Next, we study the two state model under multiple tasks. Suppose that there are two tasks (labeled 1 and 2) each of which either leads to a success (worth 1) or a failure (worth zero to the Principal). The probability of success depends on the agent’s effort in each task. Here, we’ll assume that effort is binary. The agent either works or shirks in task $i$; that is $e_i \in \{0, 1\}$. Let $c_i$ denote the cost of effort when the agent works at $i$ of the tasks and suppose that $0 = c_0 < c_1 < c_2$. Suppose that if an agent works at a task, the chance of success is $p_1$ otherwise it is $p_0$. Assume that $0 < p_0 < p_1 < 1$.

The “production technology” associated with effort may be usefully divided into two cases:

1. Effort has positive returns to scope: $c_2 < 2c_1$
2. Effort has negative returns to scope: $c_2 > 2c_1$

The interesting case is where effort has negative returns to scope (sometimes called the “substitutes” case).

As usual, we’ll assume that the agent is protected by limited liability, so only non-negative payoff schedules are allowed. Further, all parties are risk neutral.

#### 3.1 First-Best

Suppose that the Principal can perfectly monitor the agent’s effort. In that case, $P$ needs to determine whether to incent effort for both tasks, one task, or neither tasks.

If he incents effort for both tasks, his expected payoff is

$$V_2^{FB} = 2p_1 - c_2$$

If he incents effort for one task (say task 1), his expected payoff is

$$V_1^{FB} = p_1 + p_0 - c_1$$
If he incent effort for neither taks, his expected payoff is

\[ V_{0}^{FB} = 2p_{0} \]

Under what conditions are each of these regimes optimal?

2 vs 1:

\[ V_{2}^{FB} - V_{1}^{FB} = \Delta p - (c_{2} - c_{1}) \]

2 vs 0:

\[ V_{2}^{FB} - V_{0}^{FB} = 2\Delta p - c_{2} \]

1 vs 0:

\[ V_{1}^{FB} - V_{0}^{FB} = \Delta p - c_{1} \]

Result 1: If \( \Delta p < c_{1} \), incent neither task.

Reasoning: With diseconomies of scope, it is clear that since \( c_{1} < c_{2}/2 \); thus, \( V_{1}^{FB} - V_{0}^{FB} < 0 \) implies \( V_{2}^{FB} - V_{0}^{FB} < 0 \) and \( V_{2}^{FB} - V_{1}^{FB} < 0 \).

Result 2: If \( c_{1} \leq \Delta p < c_{2} - c_{1} \), incent 1 task

Reasoning: Since \( \Delta p \geq c_{1} \) then \( V_{1}^{FB} - V_{0}^{FB} \geq 0 \), so one is better than zero. Since \( \Delta p < (c_{2} - c_{1}) \) then \( V_{2}^{FB} - V_{1}^{FB} < 0 \), so one is better than 2.

Result 3: If \( \Delta p \geq c_{2} - c_{1} \), incent both tasks.

Reasoning: \( \Delta p \geq c_{2} - c_{1} \) implies \( V_{2}^{FB} - V_{1}^{FB} \geq 0 \). Further, \( \Delta p \geq c_{2} - c_{1} \) implies \( \Delta p \geq \frac{c_{2}}{2} \) (since \( \frac{c_{2}}{2} > c_{1} \)) hence \( V_{2}^{FB} - V_{0}^{FB} > 0 \).

### 3.2 Second Best

Next, suppose that effort is unobservable.

**Incent both tasks**

Suppose that P wants to incent both tasks. Risk-neutrality and limited liability imply that the optimal contract is to make a transfer \( t_{2} \) in the event of two successes and transfer nothing otherwise. As usual, we need to check incentive compatibility.

2 vs 1:

\[ (p_{1})^{2} t_{2} - c_{2} \geq p_{1}p_{0}t_{2} - c_{1} \]

Rewriting

\[ t_{2} \geq \frac{1}{\Delta p} \frac{c_{2} - c_{1}}{p_{1}} \]

We’ll refer to this inequality as the local incentive compatibility condition.

2 vs 0:

\[ (p_{1})^{2} t_{2} - c_{2} \geq (p_{0})^{2} t_{2} \]

Rewriting

\[ t_{2} \geq \frac{1}{\Delta p p_{0} + p_{1}} \]

We’ll refer to this as the global incentive compatibility condition.
Thus, P’s expected profit when incenting two tasks is

\[ V_2^{SB} = 2p_1 - (p_1)^2 \frac{1}{\Delta p} \max \left( \frac{c_2 - c_1}{p_1}, \frac{c_2}{p_0 + p_1} \right) \]

**An aside: Which constraint is binding?**

Under what conditions is the local incentive constraint binding? This occurs when

\[ \frac{c_2 - c_1}{p_1} \geq \frac{c_2}{p_0 + p_1} \]

Or

\[ c_1 \leq \frac{p_0}{p_0 + p_1} c_2 \]

Claim: If \( c_1 \leq \frac{p_0}{p_0 + p_1} c_2 \) then there are diseconomies of scope.

Proof: Rewriting this condition, we have

\[ c_2 \geq \left( 1 + \frac{p_1}{p_0} \right) c_1 \]

> \( 2c_1 \)

since \( p_1 > p_0 \).

Implication 1: When \( c_1 \leq \frac{p_0}{p_0 + p_1} c_2 \) we’ll say that the task technology exhibits **strong substitutes**. That is, the diseconomies of scope are extreme.

Implication 2: When \( c_2 > c_1 > \frac{p_0}{p_0 + p_1} c_2 \) we’ll say that the task technology exhibits **weak substitutes**. The diseconomies of scope are not so severe.

**Incenting one task:**

Suppose P chooses to incent only one task. In that case, we’re back to the earlier model and we know that

\[ t_1 = \frac{c_1}{\Delta p} \]

Hence, P earns

\[ V_1^{SB} = p_1 + p_0 - p_1 \frac{c_1}{\Delta p} \]

Incenting neither task yields the same expected payoff to P as under the first-best.

**How many tasks to incent: Strong substitutes**

Of primary interest are the conditions on \( \Delta p \) where we’d choose to incent zero, one or both tasks when tasks are strong substitutes.

2 vs 1

\[ V_2^{SB} - V_1^{SB} = \Delta p - p_1 \frac{1}{\Delta p} (c_2 - 2c_1) \]

2 vs 0

\[ V_2^{SB} - V_0^{FB} = 2\Delta p - (p_1)^2 \frac{1}{\Delta p} (c_2 - c_1) \]
\[ V_1^{SB} - V_0^{FB} = \Delta p - p_1 \frac{c_1}{\Delta p} \]

Result 1: The costs of incenting a single task have gone up. 
Reasoning: The costs of incenting one task under moral hazard are \( p_1 \frac{c_1}{\Delta p} > c_1 \), which are the costs under first-best.

Result 2: Under strong substitutes, the cost of incenting two tasks versus 1 task have gone up.
Reasoning: The costs of incenting the second task under strong substitutes and moral hazard are 
\[ p_1 \frac{1}{\Delta p} (c_2 - 2c_1) \]
versus \( c_2 - c_1 \) under first-best.
Claim: \( p_1 \frac{1}{\Delta p} (c_2 - 2c_1) > c_2 - c_1 \).
Rewriting 
\[ c_2 \left( \frac{p_1 - \Delta p}{\Delta p} \right) > c_1 \left( \frac{2p_1 - \Delta p}{\Delta p} \right) \]
Simplifying 
\[ p_0 c_2 > (p_0 + p_1) c_1 \]
or 
\[ c_1 < \frac{p_0}{p_0 + p_1} c_2 \]
which is what we assumed under strong substitutes.

Combining results 1 and 2 we have:

**Proposition 2** Under strong substitutes, P incent both a single task and both tasks less often under moral hazard

The interpretation is that, under strong substitutes, the incentive costs magnify the diseconomies of scope.

**How many tasks to incent: Weak substitutes**
Result 1: Continues to hold under this assumption.
Result 2: Under weak substitutes, the cost of incenting two tasks versus 1 task have gone down.

The incremental cost of incenting both tasks under weak substitutes and moral hazard is 
\[ (p_1)^2 \frac{1}{\Delta p} \frac{c_2}{p_0 + p_1} - p_1 \frac{1}{\Delta p} c_1 \]
versus a cost of \( c_2 - c_1 \) under first-best.
We claim \( (p_1)^2 \frac{1}{\Delta p} \frac{c_2}{p_0 + p_1} - p_1 \frac{1}{\Delta p} c_1 < c_2 - c_1 \)
Rewriting 
\[ c_2 \left( \frac{(p_1)^2 - \Delta p (p_0 + p_1)}{\Delta p (p_0 + p_1)} \right) < c_1 \left( \frac{p_1 - \Delta p}{\Delta p} \right) \]
Simplifying
\[ c_2 \left( \frac{p_0^2}{\Delta p (p_0 + p_1)} \right) < c_1 \left( \frac{p_0}{\Delta p} \right) \]
or
\[ c_1 > c_2 \frac{p_0}{(p_0 + p_1)} \]
which is exactly the condition for weak substitutes

Combining results 1 and 2 we have:

**Proposition 3** Under weak substitutes, \( P \) incents a single task less often and both tasks more often under moral hazard

Intuitively, when the global incentive constraint binds, then the marginal cost of incenting the second task is lower than under first-best. Recall that under first-best, substitutes alone implied that the local incentive was the critical factor in determining whether to pay for a second task.

## 4 Standard Model

Now we’ll generalize a bit. The agent gets to choose from among \( n \) possible effort levels: \( e_1, e_2, ..., e_n \). These efforts produce one of \( m \) outcomes: \( x_1, x_2, ..., x_m \). Suppose that when the agent chooses effort \( e_i \), \( P \) observes outcome \( x_j \) with probability \( p_{ij} > 0 \).

We shall assume separability between wealth and effort on the part of the agent i.e. the agent’s utility function is
\[ U(w, e) = u(w) - e \]
where \( u \) is increasing and concave and \( w \) is the agent’s final wealth state. We’ll assume that \( P \) is risk-neutral and that her profits are equal to the outcome less the wage paid to the agent. That is,
\[ V = x - w \]

### 4.1 The Agent’s Problem

\( P \) offers a contract \( w_j \) as a function of the outcomes. Then the agent chooses an effort \( e_i \) to maximize
\[ \sum_{j=1}^{m} p_{ij} u(w_j) - e_i \]
This implies \( n - 1 \) incentive constraints:
\[ \sum_{j=1}^{m} p_{ij} u(w_j) - e_i \geq \sum_{j=1}^{m} p_{kj} u(w_j) - e_k \]
for all $k \neq i$.

Finally, we assume an ex ante participation constraint

$$\sum_{j=1}^{m} p_{ij} u(w_j) - e_i \geq \bar{U}$$

where $\bar{U}$ is the outside option of the agent.

### 4.2 The Principal’s Problem

P needs to choose a contract $\{w_j\}$ to maximize

$$\sum_{j=1}^{m} p_{ij} (x_j - w_j)$$

subject to the $(n - 1)$ IC constraints and the IR constraint. Notice that P also controls $e_i$ indirectly by his choice of contract.

Fixing $e_i$, the Lagrangian of P’s problem is

$$L = \sum_{j=1}^{m} p_{ij} (x_j - w_j) + \sum_{k \neq i} \lambda_k \left( \sum_{j=1}^{m} p_{ij} u(w_j) - e_i - \left( \sum_{j=1}^{m} p_{kj} u(w_j) - e_k \right) \right) + \mu \left( \sum_{j=1}^{m} p_{ij} u(w_j) - e_i - \bar{U} \right)$$

Differentiating wrt $w_j$ we obtain

$$-p_{ij} + \sum_{k \neq i} \lambda_k \left( p_{ij} u'(w_j) - p_{kj} u'(w_j) \right) + \mu p_{ij} u'(w_j) = 0$$

Or

$$-\frac{1}{u'(w_j)} + \sum_{k \neq i} \lambda_k \left( 1 - \frac{p_{kj}}{p_{ij}} \right) + \mu = 0$$

or

$$\frac{1}{u'(w_j)} = \sum_{k \neq i} \lambda_k \left( 1 - \frac{p_{kj}}{p_{ij}} \right) + \mu \quad \text{(4)}$$

At the first-best, there would be efficient risk-sharing and the MU of the agent would be constant, i.e. $\frac{1}{u'(w_j)} = \mu_0$ but notice that the IC constraints create distortions.

Key to P’s problem is trying to infer the effort $e_i$ from the outcomes. This is like a maximum likelihood problem.

Given an outcome $x_j$ the maximum likelihood estimator $\hat{e} = e_i$ satisfies

$$\frac{p_{kj}}{p_{ij}} \leq 1$$

for all $k \neq i$.

Intuitively, the wage is going to tend to be higher when the likelihood of A choosing the correct effort level is higher conditional on the outcome.
4.3 Properties of the Optimal Contract

Suppose that we order the efforts and outcomes in an increasing fashion. Then, we might expect that the optimal contract would entail higher wages for more profitable outcomes. It turns out that, without additional structure on the problem, this is not a general property of the optimal contract. The best one can say (Grossman-Hart 1985) is

- \( w_j \) cannot be uniformly decreasing in \( j \).

Returning to equation (4), let’s see what properties we might need to get “sensible” contracts.

Clearly as wages increase, the LHS of equation (4) increases. Thus, we need conditions where the RHS also increases in \( j \).

First, assume that the monotone likelihood ratio property holds. This says that for all \( k < i \) and all \( l < j \)

\[
\frac{p_{ij}}{p_{ii}} \geq \frac{p_{kj}}{p_{kl}}
\]

that is a more profitable outcome is relatively more likely with a higher effort.

This then implies that

\[
\lambda_k \left( 1 - \frac{p_{kj}}{p_{ij}} \right)
\]

is increasing in \( j \) if \( k < i \) and decreasing in \( j \) when \( k > i \). So the point is that if the only incentive constraints that mattered were those for efforts below the effort we wish to induce (i.e, \( \lambda_k > 0 \) iff \( k < i \)) then we’d be done. Unfortunately, we need more assumptions to get this.

One assumption that works is called the convexity of the distribution function (CDFC) condition, proposed by Grossman-Hart. This says that the cumulative distribution function of the outcome is convex in effort. Formally, let \( i < j < k \) and choose \( t \in [0,1] \) such that

\[
e_j = te_i + (1-t)e_k
\]

then for outcome \( l \)

\[
P_{jl} \leq tP_{il} + (1-t)P_{kl}
\]

where \( P_{ij} \) is the cdf of outcomes under effort \( i \).

To see that this is enough, first solve the problem where \( e_i \) is desired by \( P \) and the agent can only choose efforts at or below this level. Clearly, this solution satisfies the desired property. Now we need only show that under this solution, \( A \) has no incentive to choose higher actions.

Suppose not. Then there exists some effort \( e_k > e_i \) such that

\[
\sum_{j=1}^{m} p_{ij} u(w_j) - e_i < \sum_{j=1}^{m} p_{kj} u(w_j) - e_k.
\]
Now, let \( l < i \) be an effort where IC holds with equality (i.e., \( \lambda_i > 0 \)) then

\[
\sum_{j=1}^{m} p_{ij}u(w_j) - e_i = \sum_{j=1}^{m} p_{ij}u(w_j) - e_l
\]

Therefore, there exists a \( t \in [0, 1] \) such that

\[
e_i = te_l + (1 - t)e_k
\]

and hence

\[
P_{ij} \leq tP_{lj} + (1 - t)P_{kj}
\]

Thus, we know that

\[
\sum_{j=1}^{m} p_{ij}u(w_j) - e_i = \sum_{j=1}^{m-1} P_{ij} (u(w_j) - u(w_{j+1})) + u(w_m) - e_i
\]

\[
\leq \left( \sum_{j=1}^{m-1} P_{ij} (u(w_j) - u(w_{j+1})) + u(w_m) - e_l \right)
\]

\[
+ (1 - t) \left( \sum_{j=1}^{m-1} P_{kj} (u(w_j) - u(w_{j+1})) + u(w_m) - e_k \right)
\]

\[
= t \left( \sum_{j=1}^{m} p_{ij}u(w_j) - e_l \right) + (1 - t) \left( \sum_{j=1}^{m} p_{kj}u(w_j) - e_k \right)
\]

and this is a contradiction.

5 Application-The optimality of debt financing

Stolen from Innes (1990)

A risk neutral entrepreneur has to exert effort \( e \) to generate profits \( \pi \). Profits are
drawn from the atomless cts cdf \( G(\pi|e) \) with associated density \( g \)

We assume that \( G \) satisfies MLRP.

The entrepreneur’s utility function is

\[
V(w, e) = w - c(e)
\]

with \( c' > 0 \) \( c'' > 0 \).

It costs the entrepreneur and amount \( I \) to set up the firm.

The entrepreneur has no money, so he turns to a risk-neutral investor to give him
\( I \) in exchange for a contract repayment \( B(\pi) \).

1. There is limited liability for \( 0 \leq B(\pi) \leq \pi \)
2. Contracts must be monotonic and not increase too steeply so $0 \leq B' \leq 1$.

The idea behind 2 is that the entrepreneur can freely burn money without the knowledge of the investor. Were that the case, the entrepreneur strictly gains by burning money in any region where $B' > 1$. For $B' > 0$, suppose that, while the entrepreneur does not have the cash to fund $I$, he does have a little money. He could then use his personal cash to boost performance marginally in regions where $B' < 0$ and this would be profitable.

Main result: If $G$ satisfies MLRP then the optimal contract is a debt contract. Further, there is insufficient effort put forth under the optimal contract than if the entrepreneur had the money to finance the start-up himself.

We use the first-order approach (which turns out to be okay here because of MLRP) to solve for the optimal contract.

To obtain the result that effort under the optimal contract is less than first-best, notice that increases in effort lead to increases in the expected payment (via FOSD). Thus, the marginal return to effort for the entrepreneur is less under such a contract than the social return. Hence, there is too little effort.

Formally

$$E(B(\pi)) = \int_{0}^{\infty} \min\{\pi, z\} g(\pi|e) d\pi \geq 0$$

where $z$ is the face value of the debt.

Now differentiate

$$\frac{dEB}{d\varepsilon} = \int_{0}^{z} \pi g_e(\pi|e) d\pi - z G_e(z|e)$$

Integrate by parts

$$z G_e(z|e) - \int_{0}^{z} G_e(z|e) d\pi - z G_e(z|e) = \int_{0}^{z} G_e(z|e) d\pi \geq 0$$

by FOSD.

Next, we show the optimality of debt.

Suppose not. Then there is some other contract $\hat{B}$ that is optimal. Suppose that $\hat{B}$ induces effort $\hat{e}$. That is

$$\int_{0}^{z} \left( \pi - \hat{B}(\pi) \right) g_e(\pi|e) d\pi = \psi'(e)$$
Then
\[ \hat{e} \in \arg \max_\pi \int_0^\infty \left( \pi - \hat{B}(\pi) \right) g(\pi|e) \, d\pi - \psi(e) \]

Now choose a debt contract with face value \( z \) that yields the same expected repayment under effect \( \hat{e} \). That is
\[ E\left( \hat{B}(\pi) \mid \hat{e} \right) = E(\min(z, \pi) \mid \hat{e}) \]

By the regularity properties of \( \hat{B} \), there is a unique solution \( z \).

Next \( \hat{B} \) must differ for a positive measure of \( \pi \) realizations from the debt contract. There is a single profit level \( \pi_B > 0 \) where the two contracts intersect. For \( \pi \leq \pi_B \), \( \hat{B}(\pi) - \min(\pi, z) \leq 0 \) and for \( \pi > \pi_B \), \( \hat{B}(\pi) - \min(\pi, z) > 0 \). Further, for \( \pi > \pi_B \), the gap is increasing in \( \pi \).

Define the gap to be
\[ \phi(\pi) = \min\{z, \pi\} - \hat{B}(\pi) \]

We seek to show that this contract provides incentives to work harder that \( \hat{e} \). That is we want to compare the marginal benefit to effort under the two contracts and show that the MB from debt is higher. Thus we aim to show
\[ \int_0^\infty \left( \pi - \min\{z, \pi\} - \left( \pi - \hat{B}(\pi) \right) \right) g_e(\pi|e) \, d\pi = \int_0^\infty -\phi(\pi) g_e(\pi|e) \, d\pi \]

First, notice that, by construction
\[ \int_0^\infty \hat{B}(\pi) g(\pi|\hat{e}) \, d\pi = \int_0^\infty \min\{\pi, z\} g(\pi|\hat{e}) \, d\pi \]

Second, for all \( \pi < \pi_B \), define
\[ \delta(\pi) = \frac{\phi(\pi) g(\pi|\hat{e})}{\int_0^{\pi_B} \phi(\pi) g(\pi|\hat{e}) \, d\pi} = \frac{\phi(\pi) g(\pi|\hat{e})}{\int_0^{\infty} \phi(\pi) g(\pi|\hat{e}) \, d\pi} \]

or
\[ \phi(\pi) = \frac{\delta(\pi)}{g(\pi|\hat{e})} \left[ -\int_{\pi_B}^{\infty} \phi(\pi) g(\pi|\hat{e}) \, d\pi \right] \]
By construction
\[ \int_0^{\pi_B} \delta (\pi) \, d\pi = 1 \]

Third, rewrite \( \int_0^\infty \phi (\pi) g_e (\pi|e) \, d\pi \) as:
\[
\int_0^\infty \phi (\pi) g_e (\pi|e) \, d\pi \\
= \int_0^{\pi_B} g_e (\pi|e) \times \frac{\delta (\pi)}{g (\pi|e)} \left[ -\int_{\pi_B}^\infty \phi (t) g (t|e) \, dt \right] \, d\pi \\
+ \int_{\pi_B}^\infty \phi (t) g_e (t|e) \, dt \times \int_0^{\pi_B} \delta (\pi) \, d\pi
\]

Now organize by order of integration
\[
\int_0^{\pi_B} \left( g_e (\pi|e) \times \frac{\delta (\pi)}{g (\pi|e)} \left[ -\int_{\pi_B}^\infty \phi (t) g (t|e) \, dt \right] + \delta (\pi) \int_{\pi_B}^\infty \phi (t) g_e (t|e) \frac{g (t|e)}{g (\pi|e)} \, dt \right) \, d\pi \\
= \int_0^{\pi_B} \left\{ \int_{\pi_B}^\infty \delta (\pi) \phi (t) g (t|e) \left[ \frac{g_e (t|e)}{g (t|e)} - \frac{g_e (\pi|e)}{g (\pi|e)} \right] \, dt \right\} \, d\pi
\]

And notice that, by MLRP, the bracketed term is positive. \( \delta (\pi) > 0 \) and \( \phi (t) < 0 \); hence the whole expression is negative. Thus, we have proved the claim.