Pure Moral Hazard

Robert A. Miller

Advanced Economic Analysis

February 2020
In an Arrow Debreu world with a Walrasian equilibrium, it doesn’t matter whether an employee is paid the value of his marginal product less the amenity value with a certain wage or a piece rate.

Both the employer and the employee can adjust their portfolio of financial assets at the competitive equilibrium rate to achieve the same resource allocation.

For example if the uncertainty is idiosyncratic, both the employee or the employer could full insure at actuarially fair rates.

The next several lectures analyze compensation and labor supply when the contract form matters.

This arises naturally in environments with asymmetric information.
A Pure Moral Hazard Model
Framework

- A risk neutral principal proposes a compensation plan to a risk averse agent, an explicit contract or an implicit agreement, which depends on the future realization of gross revenue to the principal.
- The agent accepts or rejects the principal’s (implicit) offer.
- If he rejects the offer he receives a fixed utility from an outside option.
- If he accepts the offer, the agent chooses between pursuing the principal’s objectives of value maximization (working), versus following objectives he would pursue if he was paid a fixed wage (shirking).
- The principal observes whether the offer is accepted, but not the agent’s work routine.
- After revenue is realized, the agent receives compensation according to the explicit contract or implicit agreement, and the principal pockets the remainder as profit.
Denote the workplace employment decision of the agent by an indicator $l_0 \in \{0, 1\}$, where $l_0 = 1$ means the agent rejects the principal’s offer.

Denote the effort level choices by $l_j \in \{0, 1\}$ for $j \in \{1, 2\}$, where diligence work is defined by setting $l_2 = 1$, and shirking is defined by setting $l_1 = 1$.

Since taking the outside option, working diligently and shirking are mutually exclusive activities, $l_0 + l_1 + l_2 = 1$. 

A Pure Moral Hazard Model
Revenue and profits of the principal

- Gross revenue to the principal is denoted by $x$, a random variable drawn from a probability distribution that is determined by the agent’s work routine.
- After $x$ is revealed the both the principal and the agent at the end of the period, the agent receives compensation according to the contract or implicit agreement.
- To reflect its potential dependence on (or measurability with respect to) $x$, we denote compensation by $w(x)$.
- The principal’s profit is revenue less compensation, $x - w(x)$. 
A Pure Moral Hazard Model
Marginal product of the agent

- Denote by \( f(x) \) the probability density function for revenue conditional on the agent working, and let \( f(x)g(x) \) denote the probability density function for revenue when the agent shirks.

- We assume:

\[
E[xg(x)] \equiv \int xf(x)g(x) \, dx < \int xf(x) \, dx \equiv E[x]
\]

- The inequality reflects the preference of principal for working over shirking.

- Since \( f(x) \) and \( f(x)g(x) \) are densities, \( g(x) \), the ratio of the two densities, is a likelihood ratio.

- That is \( g(x) \) is nonnegative for all \( x \) and:

\[
E[g(x)] \equiv \int g(x)f(x) \, dx = 1
\]
We assume there is an upper range of revenue that might be achieved with diligence, but is extremely unlikely to occur if the agent shirks. Formally:

$$\lim_{x \to \infty} [g(x)] = 0$$

Intuitively this assumption states that a truly extraordinary performance can only be attained if the agent works.

We assume that $g(x)$ is bounded, an assumption that rules out the possibility of setting a contract that is arbitrarily close to the first best resource allocation, first noted by Mirrlees (1975), by severely punishing the agent when $g(x)$ takes an extremely high value.
A Pure Moral Hazard Model

Preferences of the agent

- We assume the agent is an expected utility maximizer and utility is exponential in compensation, taking the form:

\[-l_0 - l_1 \alpha_1 E \left[ e^{-\gamma \omega(x)} g(x) \right] - l_2 \alpha_2 E \left[ e^{-\gamma \omega(x)} \right]\]

where without further loss of generality we normalize the utility of the outside option to negative one.

- Thus \(\gamma\) is the coefficient of absolute risk aversion, and \(\alpha_j\) is a utility parameter with consumption equivalent \(-\gamma^{-1} \log (\alpha_j)\) that measures the distaste from effort level \(j \in \{1, 2\}\).

- We assume \(\alpha_2 > \alpha_1\) meaning that shirking gives more utility to the agent, than working.

- A conflict of interest arises between the principal and the agent because he prefers shirking, meaning \(\alpha_1 < \alpha_2\), yet the principal prefers working since \(E [xg(x)] < E [x]\).
To induce the agent to accept the principal’s offer and engage in his preferred activity, shirking, it suffices to propose a contract that gives the agent an expected utility of at least minus one.

In this case we require $w(x)$ to satisfy the inequality:

$$\alpha_1 E \left[ e^{-\gamma w(x)} g(x) \right] \leq 1$$
Solving the Pure Moral Hazard Model
Participation and incentive compatibility constraints

- To elicit work from the agent, the principal must offer a contract that gives the agent a higher expected utility than the outside option, and a higher expected utility than shirking.
- In this case we require:

\[ \alpha_2 E \left[ e^{-\gamma w(x)} \right] \leq 1 \]

and:

\[ \alpha_2 E \left[ e^{-\gamma w(x)} \right] \leq \alpha_1 E \left[ e^{-\gamma w(x)} g(x) \right] \]
Solving the Pure Moral Hazard Model

Cost minimization inducing work

- Defining $v(x) \equiv \exp[-\gamma w(x)]$ note that:

$$-E[w(x)] = \gamma^{-1}E\{\log[v(x)]\}$$

the participation constraint can be expressed as:

$$\alpha_2 E[v(x)] \leq 1$$

and the incentive compatibility constraint becomes:

$$\alpha_2 E[v(x)] \leq \alpha_1 E[v(x)g(x)]$$

- In the transformed problem we maximize a strictly concave objective function with linear constraints. Applying the Kuhn Tucker theorem applies, we choose $v(x)$ for each $x$ to maximize:

$$E\{\log[v(x)]\} + \eta_0 E[1 - \alpha_2 v(x)] + \eta_1 E[\alpha_1 g(x) v(x) - \alpha_2 v(x)]$$
Lemma (Margiotta and Miller, 2000)

To minimize the cost of inducing the agent to accept employment and work diligently the board offers the contract:

\[ w^o(x) = \gamma^{-1} \ln \alpha_2 + \gamma^{-1} \ln \left[ 1 + \eta \left( \frac{\alpha_2}{\alpha_1} \right) - \eta g(x) \right] \]

where \( \eta \) is the unique positive solution to the equation:

\[
E \left[ \frac{g(x)}{\alpha_2 + \eta[(\alpha_2/\alpha_1) - g(x)]]} \right] = E \left[ \frac{(\alpha_2/\alpha_1)}{\alpha_2 + \eta[(\alpha_2/\alpha_1) - g(x)]} \right]
\]

- Differentiate the Lagrangian with respect \( v(x) \) to obtain:

\[
v(x)^{-1} = \eta_0 \alpha_2 + \eta_1 \alpha_2 - \eta_1 \alpha_1 g(x)
\]

- We can show both constraints are met with equality, establishing the formula for \( \eta \), and showing \( \eta_0 = 1 \), to yield:

\[
v(x)^{-1} = \alpha_2 + \eta_1 \alpha_2 - \eta_1 \alpha_1 g(x)
\]
There is no point exposing the manager to uncertainty in a shirking contract by tying compensation to revenue.

Hence a agent paid to shirk is offered a fixed wage that just offsets his nonpecuniary benefits, $\gamma^{-1} \ln \alpha_1$.

The certainty equivalent of the cost minimizing contract that induces diligent work is $\gamma^{-1} \ln \alpha_2$, higher than the optimal shirking contract to compensate for the lower nonpecuniary benefits because $\alpha_2 > \alpha_1$.

Moreover the agent is paid a positive risk premium of $E \left[ w^o (x) \right] - \gamma^{-1} \ln \alpha_2$.

In this model of pure moral hazard these two factors, that working is less enjoyable than shirking, and more certainty in compensation is preferable, explains why compensating an agent to align his interests with the principal is more expensive than merely paying them enough to accept employment.
Profit maximization by the principal determines which cost minimizing contract the principal should offer the agent.

The profits from inducing the agent to work are \( x - w^o(x) \), while the profits from employing the agent to shirk are \( xg(x) - \gamma^{-1} \log(\alpha_1) \).

Thus work is preferred by the principal if and only if:

\[
\max \{0, \gamma E [xg(x)] - \log(\alpha_1)\} \leq \gamma E [x - w^o(x)]
\]

while a shirking contract is offered if and only if:

\[
\max \{0, \gamma E [x - w^o(x)]\} \leq \gamma E [xg(x)] - \log(\alpha_1)
\]

Otherwise no contract is offered.
Measuring the Importance of Moral Hazard

Three measures

- Recall the optimal compensation with moral hazard is $w^o(x)$ and to meet the participation constraint, shareholders must pay $\gamma^{-1}\ln\alpha_2$.

- Therefore the maximal amount shareholders would pay to rid the firm of the moral hazard problem is:

$$\tau_1 \equiv \mathbb{E}_t \left[ w^o(x) - \gamma^{-1}\ln\alpha_2 \right] = \gamma^{-1}\mathbb{E}\left\{ \ln \left[ 1 + \eta \left( \frac{\alpha_2}{\alpha_1} \right) - \eta g(x) \right] \right\}$$

- A second measure of moral hazard is the nonpecuniary benefits the manager obtains from shirking.

- This is the monetized utility loss from working versus shirking:

$$\tau_2 \equiv \gamma^{-1}\ln\alpha_1 - \gamma^{-1}\ln\alpha_2 = -\gamma^{-1}\ln \left( \frac{\alpha_2}{\alpha_1} \right)$$

- Third is the gross loss a firm incurs from the manager shirking instead of working:

$$\tau_3 \equiv \mathbb{E} \left[ x - xg(x) \right]$$
A Fully Parametric Specification
Truncated Normal distribution and Absolute Risk Aversion (CARA)

- Assume \( x \) is distributed truncated normal with lower truncation point \( \psi \) (representing bankruptcy or limited liability) with mean \( \mu_w (\mu_s) \) and variance \( \sigma^2 \) for parent normal if agent works (shirks):

\[
f(x) = \frac{1}{\sigma_w \sqrt{2\pi}} \Phi \left( \frac{\mu_w - \psi}{\sigma} \right)^{-1} \exp \left[ -\frac{(x - \mu_w)^2}{2\sigma^2} \right]
\]

\[
\ln g(x) = \ln \Phi \left[ (\mu_s - \psi) / \sigma \right] - \ln \Phi \left[ (\mu_w - \psi) / \sigma \right]
+ \frac{\mu^2_w - \mu^2_s}{2\sigma^2} + \frac{(\mu_s - \mu_w)}{\sigma^2} x
\]

- Thus the model is parameterized by \((\psi, \mu_w, \sigma, \mu_s, \gamma, \alpha_1, \alpha_2)\).

- Suppose there are \( N \) observations on \((\tilde{w}_n, x_n)\) where:

\[
\tilde{w}_n \equiv w_n + \epsilon_n \text{ and } E[\epsilon_n | x_n] = 0.
\]
Margiotta and Miller (2000) estimate:

1. \( \psi \) with \( \hat{\psi} \equiv \min \{x_1, \ldots, x_N\} \). (Note \( \hat{\psi} \) converges to \( \psi \) at rate faster than \( \sqrt{N} \) but is sensitive to measurement error.)

2. \((\mu_w, \sigma)\) with LIML by forming likelihood for \( f(x) \) with \( \{x_1, \ldots, x_N\} \) under the assumption that \( \hat{\psi} = \psi \). (No first stage correction is necessary.)

3. \((\mu_s, \gamma, \alpha_1, \alpha_2)\) with NLS based on

\[
\tilde{w}_n = \gamma^{-1} \ln \alpha_2 + \gamma^{-1} \ln \left[ 1 + \eta \left( \frac{\alpha_2}{\alpha_1} \right) - \eta g(x) \right] + \epsilon_n
\]

using an inner loop at each iteration to solve for \( \eta \) as a mapping of \((\alpha_2, \alpha_2, \mu_s)\) given \((\hat{\psi}, \hat{\mu}_w, \hat{\sigma})\).

4. Correct the standard errors for \((\mu_s, \gamma, \alpha_1, \alpha_2)\) in the third step induced by \((\hat{\mu}_w, \hat{\sigma})\) obtained from the second step.
A Fully Parametric Specification

Estimating the importance of moral hazard (Tables 1 and 8, Margiotta and Miller 2000)

- We used the Masson-Antle-Smith (MAS) data set (37 firms in aerospace, electronics, chemicals from 1944 - 1977).
- The annual cost of moral hazard pales in comparison to losses shareholders would make if managers were paid a fixed wage.
50 Years of Managerial Compensation

Changes in managerial compensation (Table 3, Gayle and Miller, 2009)

- We compare MAS data with data from:
  - a subset formed from those firms in the three MAS sectors
50 Years of Managerial Compensation
Changes in components of managerial compensation (Table 4, Gayle and Miller, 2010)
50 Years of Managerial Compensation
Changes in sample composition of firms (Table 2, Gayle and Miller, 2010)
50 Years of Managerial Compensation

What were the driving forces behind these changes?

- If managers in the COMPUSTAT population ran firms the same size as managers in MAS, their compensation would have increased by a factor of 2.3, the increase in national income per capita.
- After adjusting for the general increase in living standards over these years, the model attributes:
  - hardly any of the increased managerial compensation to changes in \( \gamma^{-1} \ln \frac{\alpha_2}{\alpha_0} \), or the certainty equivalent wage
  - practically all the increase to changes the risk premium \( \tau_1 \)
- The factors driving the change in \( \tau_1 \) were:
  - not risk preferences: managers in the MAS (COMPUSTAT) population were willing to $240,670 ($248,620) to avoid a gamble of winning or losing $1 million.
  - not \( \Delta f(x) \): the biggest \( \Delta \tau_1 \) in aerospace where the abnormal returns became less dispersed, which reduces the risk premium
  - the sharp increase in \( \frac{\alpha_2}{\alpha_1} \) mainly due to increased firm assets, which provides managers with more opportunities to shirk.