# Contents

# Chapter 12: Principal-Agent Models

In screening models, the asymmetric information disappears in separating equilibrium after the informed party chooses a contract from the menu. Now we consider situations where the asymmetric information persists even after the contract is signed. Such situations are referred to as "hidden action" or "moral hazard."

The leading example is the Principal-Agent model. The first player, the Principal, wants to write a contract such that the second player, the Agent, will choose to work hard, even when the Principal cannot observe his effort. The Principal is subject to two constraints. First, she must write a sufficiently attractive contract so that the Agent will want to sign up (the participation constraint). Second, it must give the Agent incentive to work hard (or hard enough: the incentive constraint). We will see that the standard equilibrium contract includes base pay plus a bonus based on an observable metric that is correlated with the unobservable effort level.

We proceed in three steps. First, as a benchmark, we find the optimal contract if effort were observable. In this case a risk-neutral Principal will pay a fixed wage and completely insure the risk-averse agent against any wage risk. Second, we isolate the role of risk-aversion by looking at the optimal contract if the Agent were risk-neutral when effort is not observable. The optimal contract in this case involves the Principal "selling the store" to the agent for a fixed payment. The agent fully internalizes all the expected profit realizations from their effort and there is no role for risk-sharing since both the Principal and the Agent are risk-neutral. Finally, we show the optimal contract when the Agent is risk-averse and effort is not observable. In this case there is incomplete risk-sharing and the bonus serves to incentivize the Agent to work hard. We provide an example to illustrate the optimal contact in these three cases.

We then briefly discuss Holmstrom and Milgrom 1987 and the Revelation Principle. The focus in on contracts (or mechanisms) where truthfully revealing your type is a dominant strategy.

# 1 Contract Theory

This chapter is a first foray into contract theory.

- The basic question: How can an uninformed party design a contract that motivates the informed party to behave in the uninformed party's interest?
- Signaling and screening: the asymmetric information exists at the time of signing the contract, and then goes away (in separating PBE).
- Now, we are looking at informational asymmetries that occur *after* the contract is signed.
- Example: worker effort is impossible/difficult/costly to observe by the firm.
- The question is: how do you write a contract such that someone accepts it (signs it) and then subsequently is motivated to choose actions that are in your best-interest?
- Model: Hidden action, also called moral hazard. In contrast to hidden information, as in signalling and screening models.
- Agent: informed player, chooses an effort level from the set of effort levels.
  - We will let this be high or low effort for simplicity, but effort could also be continuous.
    $$e \in E = \{e_H, e_L\} \tag{1}$$
  - Assume that effort $e$ is not observed by the Principal, but profit $\pi$ is observed.
    * Profit is stochastically determined by effort. (If it were deterministic and 1:1, then the profit realization would exactly reveal the effort choice.)
    * Here profit is an imperfect signal regarding effort and writing the best contract is the goal.

## 2 Principal-agent problem

- Outcomes observed by the Principal (and Agent) typically include profit ($\pi$), but could be multi-dimensional as in Holmstrom and Milgrom (1987).
  - Other possible observables include revenue, costs, sales volume, inventories, selling expenses and cost-of-goods.
- We will consider only profit as the observable outcome.

$$\pi \in \Pi = [\underline{\pi}, \bar{\pi}]$$

  - Where profit is a continuous distribution.
- Assume a conditional density function:

$$\begin{aligned} &f(\pi|e) \text{ with full support} \\ &f(\pi|e_H) \text{ first-order stochastically dominates } f(\pi|e_L) \end{aligned} \tag{2}$$

  - That is, $f(\pi|e_H)$ yields unambiguously yields higher expected returns.
- Look at the cumulative density function (CDF) $F$:

$$F(\pi|e) = \int_{\underline{\pi}}^{\bar{\pi}} f(\pi|e) d\pi \tag{3}$$

  - First-order stochastic dominance states that (i) the CDF is always larger for low effort, and (ii) the expected profit from high effort is strictly greater:

$$\begin{aligned} (i) F(\pi|e_L) &\geq F(\pi|e_H) \text{ at all } \pi \in [\underline{\pi}, \bar{\pi}], \\ & \quad \text{with strict equality on some set } \Pi \subset [\underline{\pi}, \bar{\pi}] \\ (ii) E(\pi|e_H) &= \int_{\underline{\pi}}^{\bar{\pi}} \pi f(\pi|e_H) d\pi > E(\pi|e_L) = \int_{\underline{\pi}}^{\bar{\pi}} \pi f(\pi|e_L) d\pi \end{aligned} \tag{4}$$

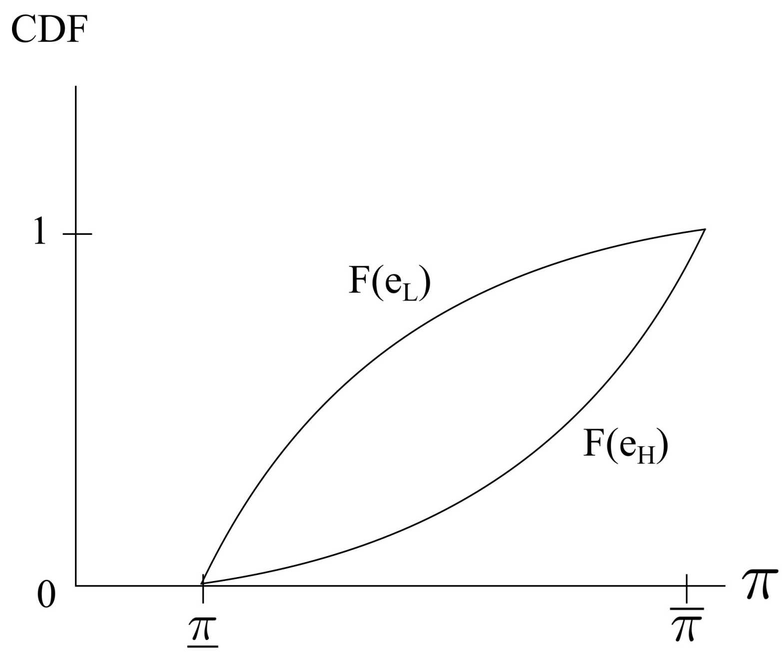


Figure 1: The cdf $F(\cdot|e_H)$ first order stochastically dominates $F(\cdot|e_L)$.

- The contract:
  - Principal will offer Agent a wage $w$ that may depend on profit $\pi$ (observable) but not on effort $e$ (unobservable).
- Payoffs:
  - Principal is risk-neutral and seeks to maximize the expected value of $\pi - w$.
  - Thus, $\pi$ is better thought of as gross profit.
  - Could have a risk-averse principal, but this complicates matters without providing much additional insight.
- Agent gains expected utility of wage but has a cost of effort.
- Specify preferences as additively separable to obtain a single crossing.
- Agent's payoff is $U_A(w, e) = u(w) - g(e)$.

(i) $g(e_H) > g(e_L)$,

(ii) $u'(w) > 0, \quad u''(w) \leq 0$,

(iii) agent has reservation utility $\bar{u}$.

- Thus, high effort has a greater cost, utility is concave and the participation constraint must satisfy the agent's reservation utility $\bar{u}$.

## 2.1 Case 1: effort observable

To isolate the role of asymmetric information, first look at what the outcome would be if effort were observable.

- The Principal has a two part problem:
  - (i) Find the cheapest contract to induce each effort level $e \in E$,
  - (ii) Find the most profitable $e$ to induce from (i).
- First consider problem (i).
- Principal chooses $w(\pi, e_H)$ and $w(\pi, e_L)$ taking into account the participation constraint.
  - That is, the expected utility of the agent must be greater than or equal to the reservation utility level: $E(U_A) \geq \bar{u}$.
  - The participation constraint [PC] is:

$$E(U_A) = \int_{\underline{\pi}}^{\bar{\pi}} u(w(\pi)) f(\pi|e) d\pi - g(e) \geq \bar{u} \tag{5}$$

    * Tie breaker: the agent participates if [PC] in equation (5) holds with equality (no hermit condition).

- Let $\gamma$ denote the Lagrange multiplier on the participation constraint, so $\gamma > 0$ if the constraint is binding.

- The Principal solves:

$$\max_{\{w(\pi)\}_{\pi \in \Pi}} \int_{\underline{\pi}}^{\bar{\pi}} [\pi - w(\pi)]\, f(\pi|e) d\pi \quad \text{s.t. [PC]} \tag{6}$$

- For this case, effort is observable, thus the contract specifies the effort level, and this problem is equivalent to:

$$\min_{\{w(\pi)\}_{\pi \in \Pi}} \int_{\underline{\pi}}^{\bar{\pi}} w(\pi) f(\pi|e) d\pi \quad \text{s.t. [PC]} \tag{7}$$

  * Thus, the Principal wishes to minimize the wage cost for a given level of effort.

- The Principal's Lagrangian is:

$$\max_{\{w(\pi)\}_{\pi \in \Pi}} \int_{\underline{\pi}}^{\bar{\pi}} [\pi - w(\pi)]\, f(\pi|e) d\pi + \gamma \left[ \int_{\underline{\pi}}^{\bar{\pi}} u(w(\pi)) f(\pi|e) d\pi - g(e) - \bar{u} \right] \tag{8}$$

- The first-order condition (or Euler) condition, it turns out, for this sort of problem is a continuum:

$$-f(\pi|e) + \gamma u'(w(\pi)) f(\pi|e) = 0 \quad \forall \pi \in [\underline{\pi}, \bar{\pi}] \tag{9}$$

- Or,

$$\gamma = \frac{1}{u'(w(\pi))} > 0 \text{ since } u'(w(\pi)) > 0 \tag{10}$$

  * Since $u'$ in general is not constant, this can only hold if $w$ is constant, independent of $\pi$.
  * Thus, the optimal compensation scheme is a fixed wage.
  * The Principal completely insures the risk-averse Agent against any income risk.

- The optimal contract $w_e^*$ solves:

$$u(w_e^*) - g(e) = \bar{u} \tag{11}$$

  - The Agent receives exactly the reservation utility level, the participation constraint is binding, since $\gamma > 0$.
  - The solution is a flat wage that is independent of the profit realization.
  - To solve for $w_e^*$, we just invert the utility function:

$$w_e^* = u^{-1}(\bar{u} + g(e)) \tag{12}$$

- This is the optimal contract to induce a particular effort level $e$.

- In addition, note that (11) implies that the wage premium to induce high effort is such that the additional utility from the higher wage is exactly equal to the additional effort cost.

- The next step is to determine what level of effort the Principal should induce to maximize profit.

$$\max_{e \in \{e_H, e_L\}} \int_{\underline{\pi}}^{\bar{\pi}} \pi f(\pi|e) d\pi - u^{-1}(\bar{u} + g(e)) \tag{13}$$

  - The first term is expected gross profit, the second is $w_e^*$.
  - Note that (12) specifies a wage for both effort levels and (13) implies that only one of these will be offered, the one with the highest expected net profit.

- The joint surplus is maximized here: Agent requires no risk premium, and the payoff sum is the largest possible expected gross product net of effort cost.

- That joint surplus is split it the the way most favorable to the Principal: the Agent just gets her reservation utility, while the Principal gets all the rest.

## 2.2 Case 2: effort not observable, agent risk-neutral

- This means $u''(w) = 0$. Without further loss of generality, we can set $u(w) = w$.
- The optimal contract turns out to be what can be called "sell the store:"

$$w^* = \pi - \alpha \tag{14}$$

  - The flat payment $\alpha$ goes to the Principal and the agent bears all the risk.
  - The idea is that you have the risk-neutral agent bear all the risk.
  - They receive all the profit realization after making the flat payment, so they internalize all of the benefits from their effort choice.

- We use the method of guess-and-check to establish that conclusion. The guess is equation (14).
- Here is the check. If Agent is risk neutral in the previous case, then equation (13) tells us that the Principal gets

$$\max_{e \in \{e_H, e_L\}} \int_{\underline{\pi}}^{\bar{\pi}} \pi f(\pi|e) d\pi - (\bar{u} + g(e)) \tag{15}$$

- The Agent's expected utility from the contract (14) looks almost the same:

$$EU_A(w^*) = \max_{e \in \{e_H, e_L\}} \int_{\underline{\pi}}^{\bar{\pi}} \pi f(\pi|e) d\pi - (\alpha + g(e)) \tag{16}$$

- The joint surplus is therefore the same as in Case 1 (when Agent is risk neutral).
- The Principal will now choose $\alpha$ to keep the Agent at reservation utility.
- Thus the payoffs to P and A are the same as in Case 1.
- The Principal in Case 2 has a constraint (non-observability of $e$) that wasn't present in Case 1. Imposing this (or any) constraint can't increase her payoff.

- Therefore the proposed contract is indeed optimal in Case 2.
- Specifically, comparing equations (15 - 16), we see that $EU_A = \bar{u}$ and

$$\alpha = \int_{\underline{\pi}}^{\bar{\pi}} \pi f(\pi|e^*) d\pi - (\bar{u} + g(e^*)), \tag{17}$$

where $e^*$ is the efficient (joint surplus maximizing) effort level.

## 2.3 Case 3: effort not observable, agent risk-averse.

This is the situation of interest.

- Now we have an additional component, the incentive compatibility constraint.
- The Principal's inner problem (finding a minimal wage schedule to induce a given effort level $e$) is

$$\max_{\{w(\pi)\}_{\pi \in \Pi}} \int_{\underline{\pi}}^{\bar{\pi}} [\pi - w(\pi)] f(\pi|e) d\pi \text{ s.t.}$$

(i) participation constraint ($\gamma$)

(ii) incentive compatibility constraint ($\mu$).

- Step 1: For each $e \in \{e_H, e_L\}$

$$\min E(w) = \int_{\underline{\pi}}^{\bar{\pi}} w(\pi) f(\pi|e) d\pi \tag{18}$$

  - subject to the participation constraint (5) as before, and the incentive constraint (IC)

$$EU(w|e, e) \geq EU(w|\tilde{e}, \tilde{e}) \quad \forall \tilde{e} \in E \tag{19}$$

  - Thus, the agent must earn a higher expected utility from choosing the Principal's desired effort level than all other effort levels.
  - Using the conditional pdfs we can write the IC constraint as

$$\int_{\underline{\pi}}^{\bar{\pi}} u(w(\pi)) f(\pi|e) d\pi - g(e) \geq \int_{\underline{\pi}}^{\bar{\pi}} u(w(\pi)) f(\pi|\tilde{e}) d\pi - g(\tilde{e}). \tag{20}$$

- For $e = e_L$, use the case 1 wage, where the reservation utility is obtained.
- For $e = e_H$, the IC constraint binds and the participation constraint binds, otherwise you could subtract a fixed payment.
  * Now the FOC has two multipliers.

- Inducing high effort $e_H$.
  - The Lagrangian is

$$\max_{\{w(\pi)\}_{\pi\in\Pi}} \int_{\underline{\pi}}^{\bar{\pi}} [\pi - w(\pi)]\, f(\pi|e_H)d\pi + \gamma \left[\int_{\underline{\pi}}^{\bar{\pi}} u(w(\pi))f(\pi|e_H)d\pi - g(e_H) - \bar{u}\right] \tag{21}$$
$$+\mu\left[\int_{\underline{\pi}}^{\bar{\pi}} u(w(\pi))f(\pi|e_H)d\pi - \int_{\underline{\pi}}^{\bar{\pi}} u(w(\pi))f(\pi|e_L)d\pi - g(e_H) + g(e_L)\right] .$$

  - The first-order condition is

$$-f(\pi|e_H) + \gamma u'(w(\pi))f(\pi|e_H) + \mu u'(w(\pi))\,[f(\pi|e_H) - f(\pi|e_L)] = 0 \tag{22}$$

    or adding $f(\pi|e_H)$ to both sides and then dividing by $f(\pi|e_H)$ and $u'(w(\pi))$

$$\frac{1}{u'(w(\pi))} = \gamma + \mu\left[1 - \frac{f(\pi|e_L)}{f(\pi|e_H)}\right]. \tag{23}$$

- Thus, the optimal contract contains a "base pay" related to $\gamma$, which satisfies the participation constraint, plus a bonus, which depends on a likelihood ratio.
  - If $\frac{f(\pi|e_L)}{f(\pi|e_H)} > 1$ then the bonus is negative.
  - Note that the bonus is linear in the likelihood ratio, but not linear in profitability.
  - Indeed, it is even possible that the optimal wage contract is non-monotonic in $\pi$, depending on the conditional density functions.

- Both $\gamma$ and $\mu$ are positive.

- The proof relies on $\frac{f(\pi|e_L)}{f(\pi|e_H)} > 1$ for an open set of profit levels $\tilde{\pi} \in [\underline{\pi}, \bar{\pi}]$ (this will occur at low profit levels) since $f(\pi|e_H)$ first-order stochastically dominates $f(\pi|e_L)$.
- Also, strictly concave utility implies $u'$ is decreasing, and the LHS of (23) is positive since $u$ is a Bernoulli function.

# 3 Principal-Agent example

- Suppose the Principal is risk-neutral and seeks to maximize

$$E(\pi - w). \tag{24}$$

- The agent is risk-averse and seeks to maximize

$$u(w) - g(e) = \sqrt{w} - e. \tag{25}$$

where effort $e \in \{e_L, e_H\} = \{0, 1\}$.

- Suppose the following conditional density function

$$(\pi|e) \sim N(2e, 1). \tag{26}$$

  - That is, profit is normally distributed with mean $2e$ and variance 1.
  - This implies $\bar{\pi} = \infty$ and $\underline{\pi} = -\infty$.

- Let the reservation utility level be normalized to zero: $\bar{u} = 0$.
  - Since utility is ordinal this is not a problem.
  - This means we can interpret the utility outcome from working for the firm as the utility gain relative to the status-quo of self-employment.

- Participation constraint (PC) is

$$\sqrt{w} - e \geq \bar{u} = 0. \tag{27}$$

  - This means that the Principal will have to pay at least 1 for high effort.
  - For $e = e_H = 1$

$$w \geq 1^2 = 1. \tag{28}$$

- Begin with the effort observable case.
  - This means there is no incentive constraint required to induce a given effort level.
- The expected values of profit are

$$\begin{aligned} E(\pi | e_H = 1) &= 2 \\ E(\pi | e_L = 0) &= 0. \end{aligned} \tag{29}$$

  - The joint surplus is $0 - 0 = 0$ with low effort and is $2 - 1 = 1$ with high effort, so the efficient contract here induces high effort.
- The optimal contract is

$$w^* = \left\{ \begin{array}{c} 1 \text{ if } e = e_H = 1 \\ 0 \text{ otherwise} \end{array} \right\}. \tag{30}$$

- In the effort observable case, the PC binds and the multiplier $\gamma = 1$.
  - This is the shadow value of breaking the constraint by one unit.
  - The constraint is that expected utility minus effort cost must be greater than or equal to the reservation utility level.
  - For the two action case $e \in \{e_H, e_L\}$ the PC always binds.

Now, effort unobservable.

- We have a two-step problem.
  - Step 1) Find the minimum cost contracts to induce $e_L = 0$ and $e_H = 1$,
  - step 2) pick the most profitable contract from step 1).
- First, the $e_L = 0$ case.

$$\min E(w) = \min_{\{w(\pi)\}} \int_{-\infty}^{\infty} w(\pi) f(\pi|e) d\pi \tag{31}$$

subject to:

  - (i) Participation constraint, $E(u) \geq 0 = \bar{u}$.

$$\int_{-\infty}^{\infty} \sqrt{w(\pi)} f(\pi|e_L = 0) d\pi \geq 0 \tag{32}$$

    * Note that the effort cost is zero for low effort.

  (ii) Incentive constraint:

$$\int_{-\infty}^{\infty} \sqrt{w(\pi)} f(\pi|e_L = 0) d\pi \geq \int_{-\infty}^{\infty} \sqrt{w(\pi)} f(\pi|e_H = 1) d\pi - 1 \tag{33}$$

    * Thus, expected utility minus effort cost is higher from choosing low effort.

  - Then the Principal finds net profit $E(\pi - w)$ from the low effort case.
    * Optimal contract is

$$w^* = 0. \tag{34}$$

    * The optimal contract is a fixed wage, independent of $\pi$, equal to zero that induces $e_L = 0$. The expected payoff to the Principal is

$$E(\pi|e_L = 0) - E(w) = 0. \tag{35}$$

* The solution is

$$w^* = u^{-1}(\bar{u} + g(0)) = u^{-1}(0) = 0^2 = 0. \tag{36}$$

– So, you pretend to pay them nothing and they pretend to work.

* Note, the participation constraint binds and the incentive constraint holds $(0 > -1)$, but is not binding.

- Next, the $e_H = 1$ case.

  – (i) Participation constraint, with multiplier $\gamma$.

$$\int_{-\infty}^{\infty} \sqrt{w(\pi)} f(\pi | e_H = 1) d\pi - 1 \geq 0 \tag{37}$$

* Note that the effort cost is 1 for high effort.

– (ii) Incentive compatibility constraint, with multiplier $\mu$.

$$\int_{-\infty}^{\infty} \sqrt{w(\pi)} f(\pi | e_H = 1) d\pi - 1 \geq \int_{-\infty}^{\infty} \sqrt{w(\pi)} f(\pi | e_L = 0) d\pi \tag{38}$$

* Thus, the contract induces high effort in this case.

– The Principal's problem is to minimize the cost needed to induce high effort, subject to the two constraints.

$$\min_{\{w(\pi)\}} \int_{-\infty}^{\infty} w(\pi) f(\pi | e_H = 1) d\pi \text{ subject to (i) PC and (ii) IC} \tag{39}$$

– Writing out the Lagrangian, noting that we have a minimization problem so we are subtracting the non-negative constraints.

$$\begin{aligned} \min_{\{w(\pi)\}} \int_{-\infty}^{\infty} w(\pi) f(\pi | e_H = 1) d\pi \quad &- \quad \gamma \left[ \int_{-\infty}^{\infty} \sqrt{w(\pi)} f(\pi | e_H = 1) d\pi - 1 \right] \\ &- \quad \mu \left[ \begin{array}{c} \int_{-\infty}^{\infty} \sqrt{w(\pi)} f(\pi | e_H = 1) d\pi - 1 \\ - \int_{-\infty}^{\infty} \sqrt{w(\pi)} f(\pi | e_L = 0) d\pi \end{array} \right] \end{aligned} \tag{40}$$

- Taking the partial with respect to the wage contract yields a continuum of FOC, $\frac{\partial \mathcal{L}}{\partial w(\pi)} = 0$.

$$0 = f(\pi|e_H) - \gamma\left(\frac{1}{2\sqrt{w(\pi)}} f(\pi|e_H)\right) - \mu\left(\frac{1}{2\sqrt{w(\pi)}}\left[f(\pi|e_H) - f(\pi|e_L)\right]\right) \tag{41}$$

- Since the profit distribution has full support we can divide by the positive constant $f(\pi|e_H)$ and multiply by $2\sqrt{w(\pi)}$ to obtain

$$2\sqrt{w(\pi)} = \gamma + \mu\left(1 - \frac{f(\pi|e_L)}{f(\pi|e_H)}\right). \tag{42}$$

  * Note that this is the same form as in equation (23), namely $\frac{1}{u'}$ = base pay of $\gamma$ plus a bonus $\mu$ determined by the likelihood ratio $\frac{f(\pi|e_L)}{f(\pi|e_H)}$.

- Next, use the normal distribution assumption $(\pi|e) \sim N(2e, 1)$ to find the likelihood ratio.

  * With a slight double-use abuse of notation, for a normal distribution, $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$, where $\sigma$ is the variance, $\mu$ is the mean, and $\pi \simeq 3.14$, and $x$ is profit.

$$\begin{aligned}
\frac{f(\pi|e_L)}{f(\pi|e_H)} &= \frac{\frac{1}{\sqrt{2\pi}} e^{\frac{-(x-0)^2}{2}}}{\frac{1}{\sqrt{2\pi}} e^{\frac{-(x-2)^2}{2}}} \\
\frac{f(\pi|e_L)}{f(\pi|e_H)} &= \frac{e^{\frac{-x^2}{2}}}{e^{\frac{-(x^2-4x+4)}{2}}} \\
\frac{f(\pi|e_L)}{f(\pi|e_H)} &= e^{\frac{-x^2+x^2-4x+4}{2}} \\
\frac{f(\pi|e_L)}{f(\pi|e_H)} &= e^{-2x+2}
\end{aligned} \tag{43}$$

  * Note that the bonus is equal to zero when profit $x = 1$ since the likelihood ratio is equal to one.
  * That is, the bonus is zero since $e = e_H$ and $e = e_L$ are equally likely.

- Now, put the simplified likelihood ratio back into (42) to obtain

$$\begin{aligned} 2\sqrt{w(\pi)} &= \gamma+\mu\left(1-e^{-2x+2}\right) \\ w^{*}(\pi) &= \left[\frac{\gamma+\mu\left(1-e^{-2x+2}\right)}{2}\right]^{2}. \end{aligned} \tag{44}$$

  - Can take this wage and put it back into the participation constraint and it will hold with equality. (show this later)
  - Or, put it into the incentive constraint and integrate it out using the normal distribution and show that both constraints will be binding.
  - The conditional density functions look like:

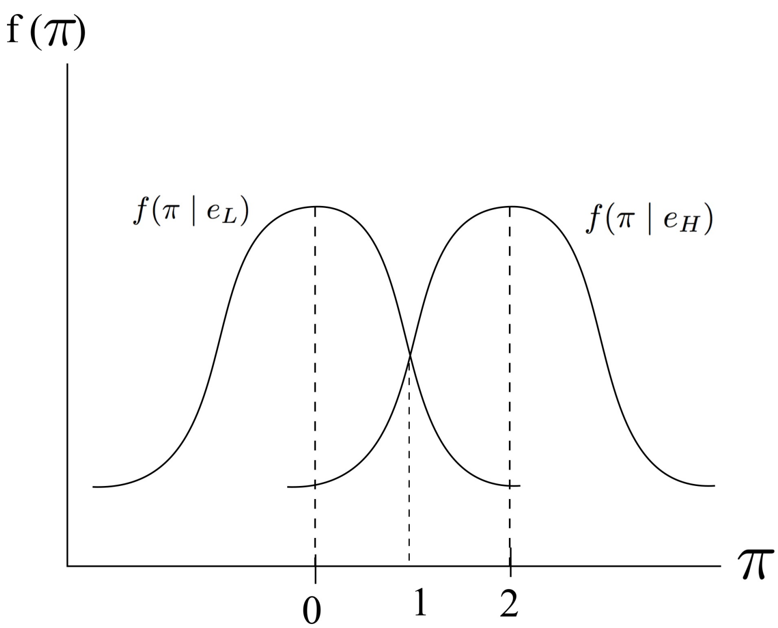


Figure 2: Two conditional pdfs diagram

- Thus, the bonus is zero for $\pi = 1$, positive for $\pi > 1$ and negative for $\pi < 1$.

# 4 Holmstrom and Milgrom 1987

- The previous analysis looked at inducing high effort, by comparing the relative cost in $w$ with the net effort cost.
- A more general problem looks at a continuum of effort levels.
- Holmstrom and Milgrom (1987) *Econometrica* 55, 303-28.
  - The bonus depends on the observable outcomes.
  - Then the problem is to find the regression coefficients:

$$w(\pi) = \alpha + \beta\pi + \gamma Y \tag{45}$$

  - Where $\pi$ is profit and $Y$ is another observable, such as output. This requires that $Y$ is correlated with effort $e$, and $e \in [\bar{e}, \underline{e}]$. This is more similar to a real-world contract, where there is a vector of observable outcomes.

# 5 Revelation Principle

- Mechanism design.
  - Principal wants to set up a contract (game) to achieve some goal.
  - What kind of games can be constructed to obtain the desired outcome.
- Create a "general game" where any BNE of this general game can be reformulated to obtain a dominant strategy equilibrium.
  - Want "truthtelling" to be a unique dominant strategy equilibrium in a direct revelation game.
  - That is, truthfully revealing private information is the unique equilibrium.

- Ex: Public goods provision.
  - Individual's have an incentive to understate preferences for a public good when they are taxed based on their stated preferences.
- Mechanisms:
  1. First-price sealed bid auction (pg. 868, MCWG)
  2. Groves-Clarke mechanism:
     - Suppose quasi-linear preferences:

$$u_i(x, \theta_i) = v(k, \theta_i) + (\bar{m}_i + t_i) \tag{46}$$

where $t_i$ is a monetary transfer, $\bar{m}_i$ is an endowment of the numeraire, $\theta_i$ is a type, $x = (k, t_1, ..., t_I)$ is a "project choice" and $k = \{0, 1\}$ is an indicator function.

       * That is $k = 1$ implies the project is undertaken, $k = 0$ implies it is not.
- Set up a mechanism where individuals have an incentive to truthfully reveal their true preferences for the public good $\theta_i$ since this is a dominant strategy, given the mechanism.

# 6 Behavioral Considerations

to be added in the next edition

- How social preferences can affect the Principal-Agent problem in practice.
- The role of corporate culture.

- When incentive bonuses are counterproductive: multiple goals, some more easily incentivized than others.
- Milgrom and Roberts (1992) elaborates in an accessible fashion.

# 7 Further Reading

Holmstrom and Milgrom (1987) *Econometrica,* 55(2), 303-328.

Milgrom, Paul R., and John Donald Roberts. "Economics, organization and management." (1992).

From other chapters, not necessarily 12.

- Section on cooperative game theory and Shapley value should include a reference to Albizuri, Arin and Rubio (2005)
  - International Game Theory Review that has a nice characterization of the Shapley value extended with a 5th axiom for games in partition function form, which is more general than the characteristic function form.
- Note: The number of partitions of a set is called Bell's number.
- There is also an article by Jean Lemarc ?? (2004?) that is a very nice explanation of the basics and sequence of thoughts, ect.
  - It talks about Aumann and XX's extension of NBS to include a 5th axiom of monotonicity and shows an interesting counter-intuitive problem with just the 4 axioms.
- Barrett (1999) Journal of theoretical politics has a very nice explanation of Grim trigger (folk theorem Fudenberg and Maskin (1996) Econometrica, Tit-for-tat (Axelrod The evolution of cooperation: 176, getting even for two players (Myerson (1991)

Game theory: analysis of conflict), getting even for more than two players (Barrett 1991, JTP) and coalition-proof Nash equilibrium Berheim et al (1987) Journal of economic theory. For reneogitation-proof equilibria see Farrell and Maskin (1989) Games and Economic Behavior. The idea is the punishment phase of a dynamic game must be credible. In other words the punishment must be immune to potential ”re-negotiation” such that punishers could earn a higher payoff if they did not carry out the punishment. For example, Barrett (1999) shows that Grim is not re-negotitation proof and thus not collectively rational.