

道德风险: 激励合同

委托代理问题

考虑如下委托代理情形:

- 委托人: 公司
- 代理人: 打工人张三
- 公司希望张三能努力干活, 因为努力干活能提高项目的成功概率, 进而提高公司的期望利润.
 - 但是, 公司不能直接观察到张三是否努力干活 (**道德风险**)

委托代理问题

考虑如下委托代理情形:

- 委托人: 公司
- 代理人: 打工人张三
- 公司希望张三能努力干活, 因为努力干活能提高项目的成功概率, 进而提高公司的期望利润.
 - 但是, 公司不能直接观察到张三是否努力干活 (**道德风险**)

问: 如果你是公司老板, 你会如何激励张三努力干活?

公司可以为项目成功提供额外的**激励**:

- 项目不成功时, 支付工资 w_0
- 项目成功时, 支付工资 $w_1 > w_0$

上面这个形式的合同也可以等价地表述为:

- 公司支付基本工资 w_0
- 如果项目成功了, 公司会额外再支付**绩效** (bonus): $b = w_1 - w_0$

无道德风险情形

如果**不存在道德风险问题**, 公司支付的工资可以直接由张三的努力程度决定, 而不是由最后张三负责的项目是否成功决定.

- 此时, 我们称公司可以**直接监督**张三的努力程度. 这意味着:
 - 公司可以直接观测到张三的努力程度.
 - 如果双方发生劳资纠纷, 公司可以向法院证明张三确实没有努力工作.

无道德风险情形

如果**不存在道德风险问题**, 公司支付的工资可以直接由张三的努力程度决定, 而不是由最后张三负责的项目是否成功决定.

- 此时, 我们称公司可以**直接监督**张三的努力程度. 这意味着:
 - 公司可以直接观测到张三的努力程度.
 - 如果双方发生劳资纠纷, 公司可以向法院证明张三确实没有努力工作.

当不存在道德风险问题时, 公司应该直接为“张三干活”提供绩效.

- **问:** 为什么从公司的角度出发, “给张三干活提供绩效”比“给张三负责的项目成功提供绩效”更为高效?

激励合同模型

考虑一个最简单的关于激励合同的道德风险模型.

这个模型只包括:

- 两个可能的代理人行动 (“努力干活”, “消极怠工”)
- 两种可能的产出水平 (“项目成功”, “项目失败”)

类似前面关于保险公司的模型, 打工人张三可以通过提高努力程度来提高项目的成功率.

张三的可能行动

- 努力工作: $a = 1$
- 消极怠工: $a = 0$

努力工作会提高项目的成功率, 但会给张三带来负效用.

张三的可能行动

- 努力工作: $a = 1$
- 消极怠工: $a = 0$

努力工作会提高项目的成功率, 但会给张三带来负效用.

- 张三的 vNM 效用取决于他的工资 $w \in \mathbb{R}$ 和行动 $a \in \{0, 1\}$:

$$v(w, a) = u(w) - al$$

模型设定

张三的可能行动

- 努力工作: $a = 1$
- 消极怠工: $a = 0$

努力工作会提高项目的成功率, 但会给张三带来负效用.

- 张三的 vNM 效用取决于他的工资 $w \in \mathbb{R}$ 和行动 $a \in \{0, 1\}$:

$$v(w, a) = u(w) - al$$

- u 是严格凹的 (张三厌恶风险)
- $l > 0$ 表示努力干活带给张三的负效用

项目可能成功或失败:

- 如果张三消极怠工 ($a = 0$), 成功的概率为 $p_0 > 0$
- 如果张三努力工作 ($a = 1$), 成功的概率为 $p_1 > p_0$

公司收益:

- 项目失败时, 公司获得利润 $\pi_0 \geq 0$
- 项目成功时, 公司获得利润 $\pi_1 > \pi_0$;

模型设定: 小结

$2 \times 2 \times 2$ 模型:

- 2 种可能行动: a_0, a_1
- 2 种可能成功概率: p_0, p_1
- 2 种可能项目收益: π_0, π_1

基准情形: 无道德风险时的最优合同

假设公司可以根据张三的行动 $a \in \{0, 1\}$ 设定工资.

为了让张三能努力工作 ($a = 1$), 公司应支付的最低工资 w^* 为:

$$u(w^*) - l = \underline{u}$$

- $\underline{u}$ 为张三辞职所能取得的效用水平 (**保留效用**).
- $\underline{u}$ 是外生给定的

公司仅仅在张三努力工作时才支付工资 w^* , 否则不支付任何工资(甚至可能对张三进行“罚款”).

道德风险下的激励设计

当张三行动不可观测时, 公司需设计绩效工资合同 (w_1, w_0) 来激励张三努力干活.

- w_0 : 项目失败时张三的工资
- w_1 : 项目成功时张三的工资

道德风险下的激励设计

当张三行动不可观测时, 公司需设计绩效工资合同 (w_1, w_0) 来激励张三努力干活.

- w_0 : 项目失败时张三的工资
- w_1 : 项目成功时张三的工资

合理的绩效工资合同应满足两个约束:

1. **激励相容** (Incentive Compatibility, **IC**)
2. **个体理性** (Individual Rationality, **IR**)

激励相容约束和个体理性约束是本讲的核心概念.

道德风险下的激励设计

当张三行动不可观测时, 公司需设计绩效工资合同 (w_1, w_0) 来激励张三努力干活.

- w_0 : 项目失败时张三的工资
- w_1 : 项目成功时张三的工资

合理的绩效工资合同应满足两个约束:

1. **激励相容** (Incentive Compatibility, **IC**)
2. **个体理性** (Individual Rationality, **IR**)

激励相容约束和个体理性约束是本讲的核心概念.

- 激励相容约束确保张三会努力干活, 而不是消极怠工
- 个体理性约束确保张三不会辞职

定义 (激励相容). 称工资合同 (w_1, w_0) 是**激励相容**的, 若张三选择接受合同时, 他会努力干活而非消极怠工.

激励相容约束

定义 (激励相容). 称工资合同 (w_1, w_0) 是**激励相容**的, 若张三选择接受合同时, 他会努力干活而非消极怠工.

- 激励相容约束的数学表示:

$$p_1 u(w_1) + (1 - p_1)u(w_0) - l \geq p_0 u(w_1) + (1 - p_0)u(w_0)$$

激励相容约束

定义 (激励相容). 称工资合同 (w_1, w_0) 是**激励相容**的, 若张三选择接受合同时, 他会努力干活而非消极怠工.

- 激励相容约束的数学表示:

$$p_1 u(w_1) + (1 - p_1)u(w_0) - l \geq p_0 u(w_1) + (1 - p_0)u(w_0)$$

$$\Leftrightarrow (p_1 - p_0)[u(w_1) - u(w_0)] \geq l$$

激励相容约束

定义 (激励相容). 称工资合同 (w_1, w_0) 是**激励相容**的, 若张三选择接受合同时, 他会努力干活而非消极怠工.

- 激励相容约束的数学表示:

$$p_1 u(w_1) + (1 - p_1)u(w_0) - l \geq p_0 u(w_1) + (1 - p_0)u(w_0)$$

$$\Leftrightarrow (p_1 - p_0)[u(w_1) - u(w_0)] \geq l$$

- 文字解释: 张三努力干活的边际收益大于边际成本.

$$(p_1 - p_0)[u(w_1) - u(w_0)] \geq l \quad (\text{IC})$$

$$(p_1 - p_0)[u(w_1) - u(w_0)] \geq l \quad (\text{IC})$$

决定张三是否干活的, 是 $\Delta u = u(w_1) - u(w_0) \approx u'(w_0)(w_1 - w_0)$.

- 它取决于合同 (w_1, w_0) 本身以及张三的货币偏好 u .

个体理性约束 (or, 参与约束)

个体理性约束 (也叫参与约束):

- 张三接受合同的效用不会低于他辞职的保留效用 ($\underline{u}$)

个体理性约束 (or, 参与约束)

个体理性约束 (也叫参与约束):

- 张三接受合同的效用不会低于他辞职的保留效用 ($\underline{u}$)

对于满足激励相容约束的合同 (w_1, w_0) , 它对应的个体理性约束为:

$$p_1 u(w_1) + (1 - p_1) u(w_0) - l \geq \underline{u}$$

- 当张三的保留效用 $\underline{u}$ 变大时, 为了让张三能投入工作, 公司必须支付更高的工资.

公司的目标: 最大化期望利润

$$p_1(\pi_1 - w_1) + (1 - p_1)(\pi_0 - w_0)$$

- 当我们说公司 (或委托人) 的目标是最大化期望利润时, 其实假设了公司是**风险中性**的.

公司的最优化问题

公司的最优化问题: 选择满足**激励相容约束**和**个体理性约束**的合同 (w_1, w_0) 来最大化期望利润.

公司的最优化问题

公司的最优化问题: 选择满足**激励相容约束**和**个体理性约束**的合同 (w_1, w_0) 来最大化期望利润.

$$\max_{w_1, w_0 \in \mathbb{R}} p_1(\pi_1 - w_1) + (1 - p_1)(\pi_0 - w_0)$$

subject to:

$$p_1 u(w_1) + (1 - p_1)u(w_0) - l \geq \underline{u} \quad (\text{IR})$$

$$(p_1 - p_0)[u(w_1) - u(w_0)] \geq l \quad (\text{IC})$$

- 记这个带约束的最优化问题为 $\mathcal{M}$.