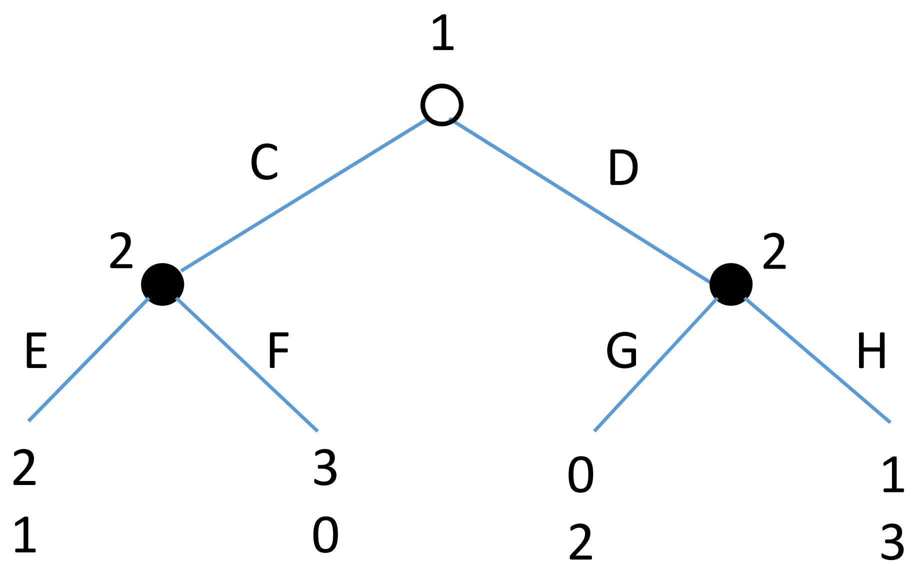
第六讲：动态博弈应用

回顾

- 动态博弈的延展型
- 动态博弈的策略型及纳什均衡
- 子博弈完善均衡(SPE)
  - 该策略组合需要在所有的子博弈里都构成纳什均衡
  - 换言之，在每一个决策点，决策者所作的都必须是最优选择
- 有限博弈中，可用**逆推归纳法**找出子博弈完善均衡



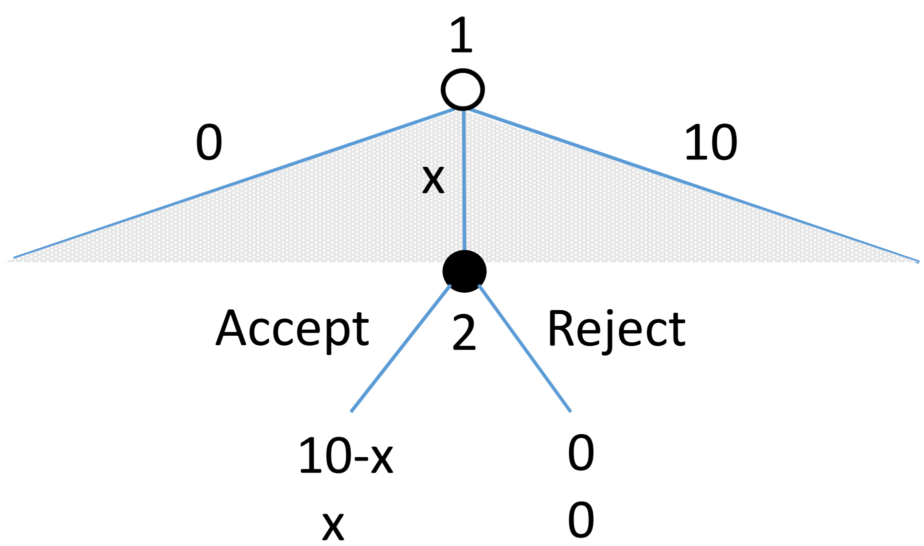
		博弈者2			
		EG	EH	FG	FH
博弈者1	C	2，1	2，1	3，0	3，0
	D	0，2	1，3	0，2	1，3

今天内容

- 讨价还价(bargaining)
- 斯塔克伯格竞争(Cournot 竞争模型的动态版)
- 委托—代理模型(principal-agent model)

“最后通牒”(The Ultimatum Game)

- 两名博弈者和一定数量的资源：比如10元
- 博弈者1（提议方）提出如何分配的方案：分给对方的数目为 x（$0 \le x \le 10$），自己获得10 - x
- 博弈者2（回应方）选择接受或拒绝：接受则按此分配；拒绝则两人都得到0



“最后通牒”(The Ultimatum Game)

- 逆推归纳法
  - 2如果拒绝，收益为0，所以2会接受任何大于等于0的提议
  - 那么1会提议自己全得，2获得0
- 子博弈完善均衡：(propose x=0， accept iff x≥0)
- 1这里占有某种“先行者”优势
  - 他的行动（威胁）具有可信度：因为根据规则，1的提议是最后提案（final offer），无法撤回或修订

*现实中博弈者的行为？

- 博弈者1一般会提议分给博弈者2大约40%左右；大部分的博弈者2会拒绝过分不平等的方案（例如分给2的少于20%）

讨价还价(Bargaining Games)

- 更一般的讨价还价博弈，不妨将待分配的总价值为记为1
- 如果谈判进行N轮，总额恒定
  - 1提出方案，2决定接受还是拒绝；接受则博弈结束，如果拒绝就进入下一轮，由2提出方案，1决定接受还是拒绝；这样交替进行……
  - 直到最后第N轮，一方提出方案，另一方选择接受或拒绝；拒绝的话（不能达成协议），双方获得“保留收益”(s_1，s_2)

  这里的s_1，s_2代表各自获得的份额：$s_1 + s_2 \le 1$，$0 \le s_1, s_2 \le 1$

逆推归纳法

保留收益（份额）：(s_i, s_j)　　$s_i + s_j \le 1$

	提议方	博弈者i	博弈者j
Period N	***i***	$1\text{-}s_j$	s_j
Period N-1	***j***	$1\text{-}s_j$	s_j
… …		…	…
Period 1	***i or j***	$1\text{-}s_j$	s_j

讨价还价：有限轮次、不变的馅饼

- 均衡**结果**：无论在哪一轮达成协议，结果都是$(1 - s_j, s_j)$，即最后的回应方 j 只获得其保留收益s_j，而最后一轮的提议方i获得了$1 - s_j$
- 影响双方所得的两个结构性特征
  - **谁最后回应**：最后的回应方“吃亏”，只能获得自己的保留收益；最后一轮的提议方有利，有可能获得高于保留收益的收益
    - 现实中，每个人都希望强迫对方作“最后回应方”：“take it or leave it”
  - **保留收益**：谈判一旦破裂会导致的结果，s_i 和 s_j

讨价还价：缩小的馅饼(Shrinking Pie)

- 如果在讨价还价的过程中，结果的价值随着时间的推移而降低
  - 折扣系数（discount factor）δ： $0 < \delta < 1$
  - 对于今天的决策者而言，同样的结果（“票面”价值为v），今天立即获得该结果的收益价值为v；明天获得该结果的收益价值只有δv；后天获得该结果的收益价值只有$\delta^2 v$；……
  - δ 可以理解为个人心理层面的“耐心”，也可以理解为物理意义上的“缩水”
- 在对于未来打折的模型里，决策者在比较各种行动方案时，会把它们导致的结果都折合成“现值”，有两种（等价的）折算方式
  - 一种是灵活地把正在考虑的某个决策阶段（t_m）看作“当前”，把其后的阶段的收益折算为这个阶段的“现值”
  - 另一种是把第一阶段（t_1）固定为“当前”，把所有其他阶段的收益都折算为阶段1的“现值”，在任何阶段决策时，都使用这个“现值”

- 考虑一个3轮讨价还价的例子
  - 双方如果不能达成协议的各自保留收益的份额是(s_i, s_j)
  - 为简单起见，假定博弈者可以接受提议时总是接受

缩小的馅饼：N=3，$0 < \delta < 1$

保留收益（份额）：(s_i, s_j) $s_i + s_j \le 1$

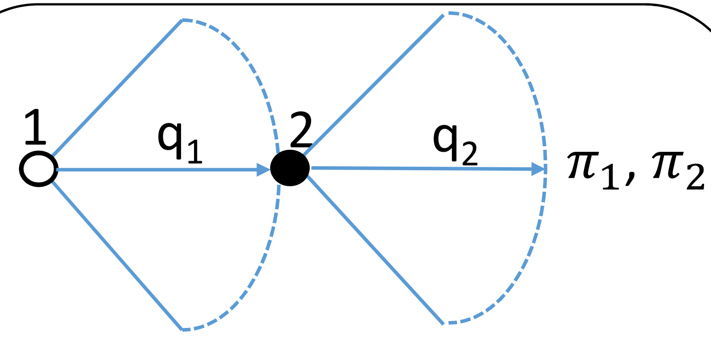
	提议方	博弈者i	博弈者j
Period 3	$\boldsymbol{i}$	$1 - s_j$	s_j
Period 2	$\boldsymbol{j}$	$\delta(1 - s_j)$	$1 - \delta(1 - s_j)$
Period 1	$\boldsymbol{i}$	$1 - \delta(1 - \delta + \delta s_j)$	$\delta(1 - \delta + \delta s_j)$

- 结果：在第一轮达成协议，提议方i获得$1-\delta+\delta^2-\delta^2 s_j$，j获得$\delta-\delta^2+\delta^2 s_j$
- 这个逻辑，可以推广到n轮(n>3)的博弈（请课下自己确认）
  - 第一轮达成协议，提议方获得$S=1+\sum_{k=1}^{n-1}(-1)^k\delta^k+(-1)^n\delta^{n-1}s_j$（这里的$s_j$是最后回应者j的保留收益），另一方获得1-S
- 当$n\rightarrow\infty$，首先提议者的收益 → 1/(1+δ)，而对方的收益 → δ/(1+δ)（课下确认，用等比数列求和公式）
- 可以看到，**当允许谈判的轮次n非常大时**，谁是最后回应者及她的保留收益不再重要，最重要的是馅饼缩小的速度(δ)，其次是第一轮的角色分配
  - 馅饼缩小很慢（或双方都很耐心）的情况下，双方的收益接近平分
  - 馅饼缩小很快（或双方都没有耐心）的情况下，首先提议方获得较多，另一方获得较少

斯塔克伯格模型（Stackelberg Duopoly）

- 将前面（第三讲）静态的库诺特竞争模型 (Cournot Duopoly) 变为动态博弈
- 两家企业生产同一种产品；企业1**先决定**自己的产量q_1并进行生产；观察到企业1的行动以后，企业2**再决定**自己的产量q_2
- 参数设定如前：企业 i 的生产成本为 $c \cdot q_i$；双方总产量为$Q=q_1+q_2$，市场对产品的需求函数为 $P = \alpha - Q$（其中P是价格，$\alpha > c$），企业i的收益为：$\pi_i = Pq_i - cq_i$

逆推归纳法



$$Q = q_1 + q_2$$
$$P = \alpha - Q$$
$$\pi_i = Pq_i - cq_i$$

- 先考虑**企业2**的最优反应，针对企业1的产量q_1
- 最大化 $\pi_2 = Pq_2 - cq_2$

$$= (\alpha - q_1 - q_2) \cdot q_2 - cq_2$$
$$= (\alpha - c - q_1) \cdot q_2 - {q_2}^2$$

- 根据一阶条件（一阶导数为零）：$(\alpha - c - q_1) - 2q_2 = 0$
- 企业2的最优反应是

$$\boxed{q_2 = \frac{1}{2}(\alpha - c - q_1)} \quad when \;\; q_1 \le \alpha - c$$
$$q_2 = 0 \quad when \;\; q_1 > \alpha - c$$

(不难看出，$q_2 = 0$, $q_1 > \alpha - c$ 不可能是均衡；因为此时$Q > \alpha - c$，导致价格 $P < c$，企业1的利润为负）

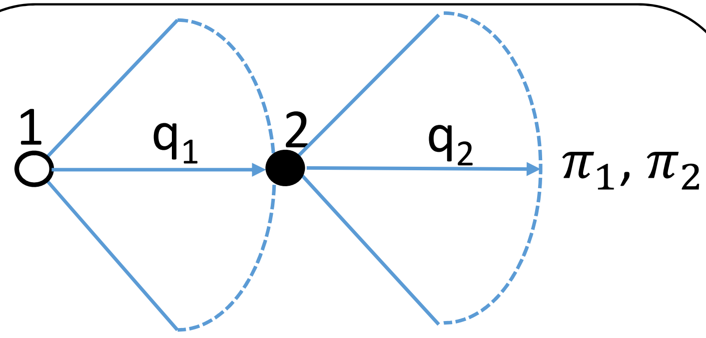
- 已知企业2的选择 $q_2 = \frac{1}{2}(\alpha - c - q_1)$
- 企业1最大化 $\pi_1 = Pq_1 - cq_1$

$$= (\alpha - q_1 - q_2) \cdot q_1 - cq_1$$
$$= (\alpha - c - q_2) \cdot q_1 - {q_1}^2$$

- 代入已知的q_2

$$\pi_1 = \left(\alpha - c - \frac{1}{2}(\alpha - c - q_1)\right) \cdot q_1 - {q_1}^2$$
$$= \frac{1}{2}(\alpha - c) \cdot q_1 - \frac{1}{2}{q_1}^2$$

- 一阶条件（一阶导数为零）：$\frac{1}{2}(\alpha - c) - q_1 = 0$
- 解出企业1的最优产量： $q_1 = \frac{1}{2}(\alpha - c)$
- 再代入企业2的最优反应函数，解出企业2的产量： $q_2 = \frac{1}{4}(\alpha - c)$
- 均衡结果：$(\frac{1}{2}(\alpha - c), \frac{1}{4}(\alpha - c))$（注意：均衡中2的策略是$\frac{1}{2}(\alpha - c - q_1)$）



$$Q = q_1 + q_2$$
$$P = \alpha - Q$$
$$\pi_i = Pq_i - cq_i$$

与之前的静态博弈比较

- 产量：静态博弈 $(\frac{1}{3}(\alpha - c),\ \frac{1}{3}(\alpha - c))$；动态博弈 $(\frac{1}{2}(\alpha - c),\ \frac{1}{4}(\alpha - c))$
  - 这里，企业1的产量增加，企业2的产量减少
- 利润：静态博弈 $(\frac{1}{9}(\alpha - c)^2, \frac{1}{9}(\alpha - c)^2)$; 动态博弈 $(\frac{1}{8}(\alpha - c)^2,\ \frac{1}{16}(\alpha - c)^2)$
  - 这里，企业1的利润增加，企业2的利润减少
- 企业1在这里有“先行者优势”，可以选择较高的产量，迫使2降低产量
  - 在同时型博弈中，1如果宣称他会选择这种高产量，2不会相信。因为如果2选择最优反应(低产量)，1有动力降低产量；而2知道1会降低产量，就不会选择这么低的产量
  - 而在序贯博弈中，由于1是first mover，他的选择是无法更改的，所以是可信的，2不得不选择较低的产量

带有“道德风险”的博弈

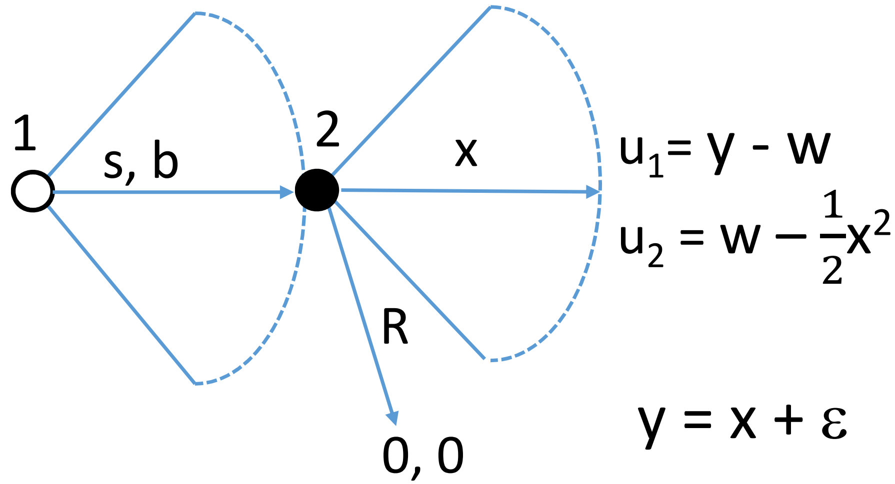
- “道德风险”(Moral Hazard) 广义上指一方的行动使得另一方有选择负面行动的可能性或动力
  - 例如：借钱给对方，对方有动力不还
  - 例如：保险令受保人忽视健康或不顾风险
- 道德风险往往导致动态博弈中的“合作困境”
- 下面分析一个常见的含有道德风险的模型：委托代理模型（principal-agent model)

委托—代理问题（Principal-Agent Problem）

- 委托人聘请代理人进行工作，提供一份合同，合同中的报酬为w
- 代理人可以选择行动，如努力程度，记为x
- 但是代理人的行动x本身无法通过合同约定 (non-contractible)
  - 因为信息缺乏：委托人无法直接观察到x，或者无法向第三方证明x
- 可观察的是工作的成果y：$y = f(x) + \varepsilon$

    其中 ε 是随机变量，例如可以是正态分布的，$\varepsilon \sim N(0, \sigma)$

委托—代理模型



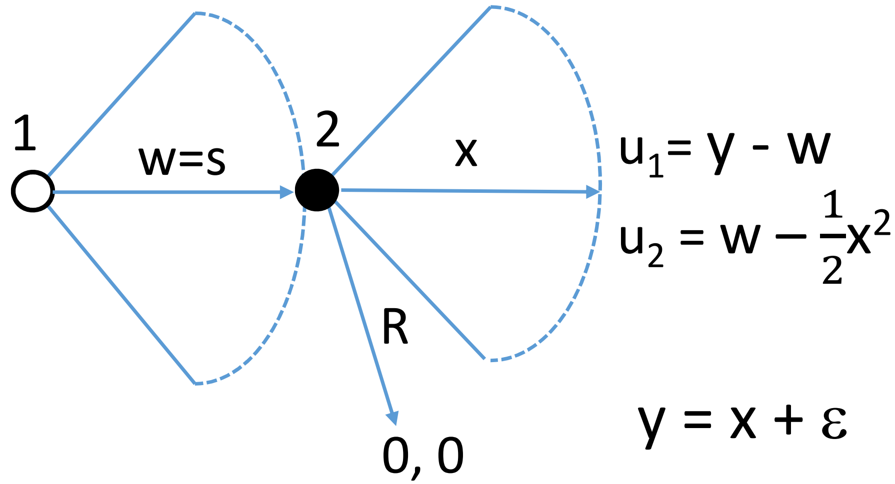
- 委托人首先提出合同，报酬w的形式为

  $w = s + b \cdot y \qquad (s \geq 0，\ b \geq 0)$

- 代理人可以选择接受或者拒绝合同
  - 拒绝合同双方的收益为0
  - 如果接受的话，代理人选择行动 x，生产出 $y = f(x) + \varepsilon$，为简单起见我们设定 $y = x + \varepsilon$， $\varepsilon \sim N(0, \sigma)$

- 博弈结束，委托人获得产品并按照合同支付报酬，双方收益为
  - 委托人获得y，付出w，收益为：$u_1 = y - w$
  - 代理人获得w，付出c(x)；这里设定 $c(x) = \frac{1}{2}x^2$，他的收益为：$u_2 = w - \frac{1}{2}x^2$

固定工资（b=0）



- 假定委托人提出的是固定工资

  $w = s$（s为一常数）

- 逆推归纳法，先考虑**代理人**

  给定w=s，他希望最大化 $u_2 = s - (1/2)\,x^2$

  显然，对于任何固定的s，$x = 0$是最优的

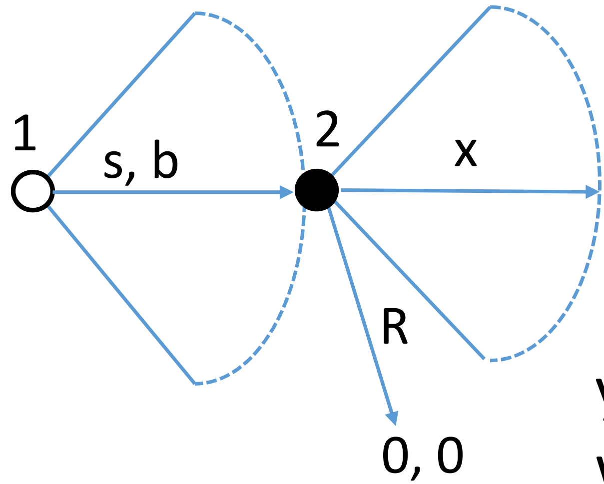
- 给定代理人的策略 x=0，考虑**委托人**

  他的期望收益：$E(u_1) = E(x + \varepsilon - w) = E(0 + \varepsilon - s) = -s$

  委托人的最优选择是 $s = 0$

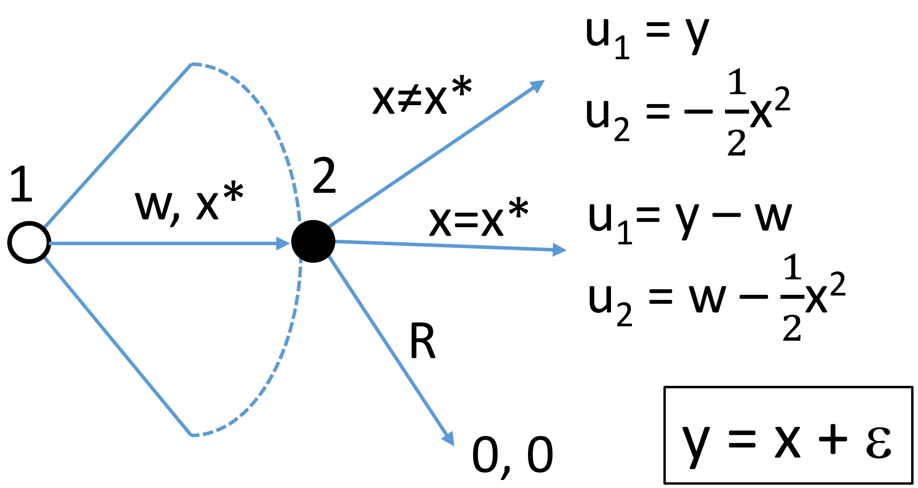
- 子博弈完善均衡：(s=0, x=0)，双方收益为(0, 0)

利润分成（b>0）



- 合同 $w = s + b \cdot y$ $(b > 0, s \geq 0)$
- **代理人的收益**：$E(u_2) = E(s + b \cdot (x + \varepsilon) - (1/2)\, x^2)$
  $= s + b \cdot x - (1/2)\, x^2$
  最大化的一阶条件（导数为零）：$b - x = 0$
  代理人的最优选择：$x = b$
- **委托人的收益**：$E(u_1) = E(y\text{-}w) = E(x + \varepsilon - s - b \cdot (x + \varepsilon)) = x - s - bx$
  已知 x=b，所以 $E(u_1) = b - b^2 - s$
  委托人对于b、s的最优选择（$s \geq 0$），根据一阶条件
  b：$1 - 2b = 0 \rightarrow b = 1/2$
  而对于s可看出：$s = 0$
- 均衡结果: 委托人给出的合同是 $s = 0, b = 1/2$，代理人的努力程度是 $x = b = 1/2$
- 收益：委托人 $E(u_1) = 1/4$，代理人 $E(u_2) = 1/8$，总收益3/8

比较：最优(First Best)



- **如果**代理人的行动x是**可以**通过合同约定的
  - 委托人可指定x*；如果达到，付给固定的w，如果达不到，付给0
- 从**代理人**的角度：如果接受合同，最优选择就是付出$x = x^*$；接受合同的条件是 $u_2 = w - (1/2)\,x^{*2} \geq 0$，即$w \geq (1/2)\,x^{*2}$
- **委托人**的期望收益：$E(u_1) = E(x^*+\varepsilon - w) = x^* - w$
- 求出最优选择x*和w （约束条件下求极值的一般方法不要求掌握）
  - 这里委托人希望w越小越好，x*越大越好；而w“最小”为 $w = (1/2)\,x^{*2}$
  - 代入， $E(u_1) = x^* - (1/2)\,x^{*2}$，用一阶条件可解出：$x^*=1$， $w=1/2$
- 结果：委托人提出合同 $x^*=1, w=1/2$，代理人接受并付出$x=1$
- $E(u_1)=1/2$， $E(u_2)= 0$，总体收益为1/2，全部为委托人所有

比较

- 最优 (First Best)：
  - 合同：指定x*=1，s=1/2，b=0
  - 努力程度：x=1
  - 收益：委托人1/2, 代理人0，总收益1/2

- 次优(Second Best)
  - 合同：不指定x，s=0，b=1/2
  - 努力程度：x=1/2
  - 收益：委托人1/4，代理人1/8，总收益3/8

- 在行为无法观察时，需要采用基于结果的(outcome-based)激励机制
- 这种激励是一种次优选择；信息不对称使得总效用受损
  - 总收益下降；代理人投入的努力不足
- 信息的不对称在这里使得双方的得失并不相同
  - 代理人的收益上升：获得信息租金（information rent)
  - 委托人的收益下降：付出信息租金

- 作业3：请查教学网，下周（4月3日）课堂上交