Price Theory III
Spring 2019

(Lars Stole)

Simon Sangmin Oh

University of Chicago

Note:

These lecture notes are based on Professor Stole's lectures in Theory of Income III, Spring Quarter.

Contents

1 Games of Incomplete Information

"We have whiteboards [as opposed to blackboards] at Booth." – Lars Stole

1.1 Static Games of Incomplete Information

We will review static games of incomplete information that are essential to this quarter's material.

1.1.1 Setup

The game consists of the following elements:

- ▷ players: $i \in N = \{1, ..., n\}$
- ▷ types: $\theta = (\theta_1, ..., \theta_n) \in \Theta$ accompanied by an underlying cdf F or pdf ϕ.
- ▷ action space: $s_i \in S_i$
- ▷ payoff: $u_i (s_i, s_{-i}, \theta)$ depends on everyone's action (strategy) and type

Then the **bayesian game** is defined as the following:

$$\Gamma = \{N; S_1, ..., S_n; u_1, ..., u_n; \Theta, F\}$$

1.1.2 First-Price Auction (FPA) with Two Players

Let's map a first-price auction into the above framework. The tricky part is the payoff part:

$$u_i = \begin{cases} \theta_i - b_i & b_i > b_j \\ \frac{1}{2}(\theta_i - b_j) & b_i = b_j \\ 0 & b_i < b_j \end{cases}$$

Then we can write the FPA as the following game:

$$\Gamma_{FPA} = \{\{1, 2\}; \mathbb{R}_+, \mathbb{R}_+; u_i; \Theta, F\}$$

We can define different notions of strategies in this game:

- ▷ A **pure strategy** is denoted as $b_i : \Theta_i \to \mathbb{R}_+$.
- ▷ A **mixed strategy** is denoted as $\sigma_i (s_i|\theta_i) \in \Delta (s_i)$ where $\Delta (s_i)$ is the set of distributions over s_i.

We say that $\{\sigma_1^*, ..., \sigma_n^*\}$ is a **Bayesian Nash Equilibrium** if for every player and every type, i.e. $\forall i, \forall \theta$:

$$\sigma_i^* (s_i|\theta_i) > 0 \Rightarrow \mathbb{E}_{\theta_{-i}} \left[u_i \left(s_i, \sigma_{-i}^*, \theta\right) |\theta_i\right] \geq \mathbb{E}_{\theta_{-i}} \left[u_i \left(\hat{s}_i, \sigma_{-i}^*, \theta\right) |\theta_i\right], \forall \hat{s}_i \in S_i$$

- ▷ Note that this expectation is essentially a double expectation since

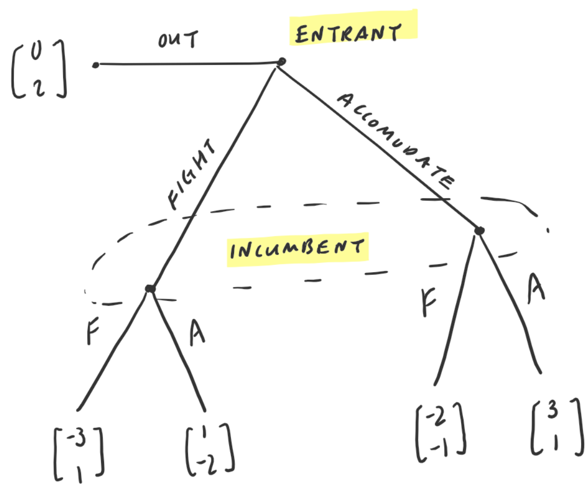
$$u_i \left(s_i, \sigma_{-i}^*, \theta\right) = \sum_{s_{-i} \in S_{-i}} u_i (s_i, s_{-i}, \theta) \, \sigma_{-i}^* (s_{-i}|\theta_{-i})$$

1.2 Dynamic Games of Incomplete Information

Typically: nature moves, the leader then moves, and then the follower moves.

1.2.1 Example

Consider the game of the following form:



Note that we can write this as the following normal form game:

	F	A
Out	$\mathbf{(0,2)}$	$(0,2)$
F	$(-3,1)$	$(1,-2)$
A	$(-2,-1)$	$\mathbf{(3,1)}$

Solving this normal form game has two equilibria: (Out, F) and (A, A).

1.2.2 Weak Sequential Equilibrium (Weak SE) ⇔ Weak PBE

Recall that sequential equilibria are defined in pairs (σ^*, μ^*). Then a **weak sequential equilibrium** (σ^*, μ^*) satisfies:

1. Sequential rationality: σ^* is sequentially rational given μ^*
2. Bayes rule: If $p(h|\sigma^*) > 0$, then
$$\mu^* = \frac{p(x \in h|\sigma^*)}{p(h|\sigma^*)}$$
where h is the information set. Note that we cannot say anything about information sets that are never reached.

1.2.3 Extending WSE #1: Perfect Bayesian Equilibrium (PBE)

The **perfect bayesian equilibrium** (σ^*, μ^*) satisfies the following[1]:

[1] See Ch.8 of FT

1. Sequential rationality (same as before)
2. Bayes rule (same as before)
3. Independent posteriors (which only makes sense if you have independent priors)
   - ▷ This says that if nature chooses beliefs independently at the start of the game, then it's not right if the resulting posteriors turn out to be correlated in some manner.
4. Don't Signal What You Don't Know (DSWYDK)
   - ▷ Suppose you're playing a game twice, where the game history between the two games is player 3 playing left vs. right. Then the posterior on player 2 should be the same in both games.
   - ▷ The idea is that player 3's action should not be giving you any information about player 2.
5. Common μ for out-of-equilibrium actions
   - ▷ This condition is a little contentious. This is done for convenience – if you have firms (agents) to have different out-of-equilibrium beliefs, then you can't really solve the games quite easily.

Note that this turns out to be a little weaker than sequential equilibrium.

- ▷ For our purposes, the set of sequential equilibria and the set of PBE are the same (e.g. nature-leader-follower games).

1.2.4 Extending WSE #2: Sequential Equilibrium (SE)

The **sequential equilibrium** (σ^*, μ^*) satisfies the following

1. Sequential rationality (same as before)
2. Consistency: Take a sequence $\{\sigma_k\}$ and $\{\mu_k\}$ such that $\{\sigma_k\} \to \sigma^*$ and $\{\mu_k\} \to \mu^*$

2 Adverse Selection in Competitive Markets

"This is not a course on game theory. It's a course on information economics, which is far more interesting" – Lars Stole

2.1 Adverse Selection in Labor Markets

We assume that the object being traded is exogenously fixed. In the labor market example, there is a single type of job and the worker is either employed or not; in the insurance example, there is only a full insurance contract available and it is either purchased or not.

2.1.1 Setup

Workers: type $\theta_i \sim F$ on $\left[\underline{\theta}, \bar{\theta}\right]$ which could be discrete or continuous.

▷ The worker's utility is

$$u = \begin{cases} w & \text{if work} \\ r\left(\theta\right) & \text{if no work} \end{cases}$$

where $r\left(\cdot\right)$ represents the reservation wage.

▷ Denote $\Theta\left(w\right) = \left\{\theta | r\left(\theta\right) \leq w\right\}$ is the set of workers who work. This is the **supply of labor**.

Firms: they employ workers and produce.

▷ **Demand for labor** is expressed as:

$$D\left(w\right) = \begin{cases} 0 & \text{if } \mathbb{E}\left[\theta | w\right] < w \\ \infty & \text{if } \mathbb{E}\left[\theta | w\right] > w \\ [0, \infty) & \text{if } \mathbb{E}\left[\theta | w\right] = w \end{cases}$$

2.1.2 Competitive Equilibrium

The competitive equilibrium must satisfy the following two equations:

$$\begin{aligned} (1): &\quad \mathbb{E}\left[\theta | r\left(\theta\right) \leq w^*\right] = w^* \\ (2): &\quad \Theta^* = \left\{\theta | r\left(\theta\right) \leq w^*\right\} \end{aligned}$$

Note that the conditional expectation is a form of rational expectation since it's conditioned on those who show up. We consider two cases:

▷ Suppose r$(\theta) = r$. There are two possibilities in this case:

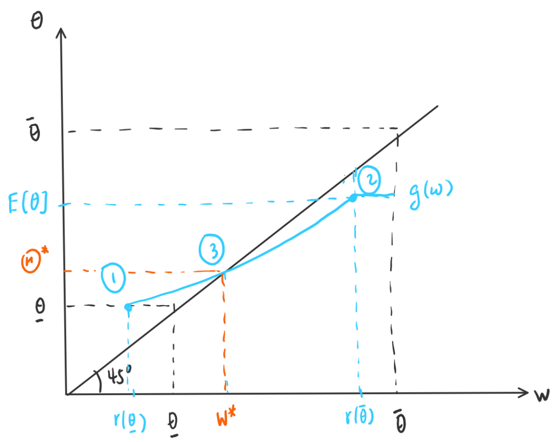
- Case $w^* \geq r$: Everyone works, so $\Theta^* = \Theta$. To get this outcome, we need $w^* = \underline{\mathbb{E}\left[\theta\right] \geq r}$. Note that we can take an unconditional expectation since everyone shows up.
- Case $w^* < r$: Nobody shows up, so $\Theta^* = \emptyset$. To get this outcome, we need $\underline{r > \mathbb{E}\left[\theta\right]} = w^*$

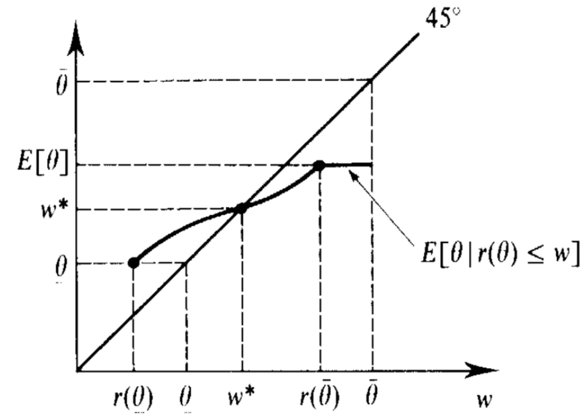
▷ Suppose $r\left(\theta\right)$ is increasing in θ. This immediately introduces **adverse selection.**

- Mathematically, you have $g\left(w\right) \equiv \mathbb{E}\left[\theta | r\left(\theta\right) \leq w\right]$ is increasing in w since as you raise your wage, better people will show up.
- We want to see how this impacts our competitive equilibrium.

2.1.3 Existence of Equilibrium

We will *assume* that full employment is always optimal, i.e. $\theta \geq r(\theta)$. We want to find the point at which $g(w) = w$.

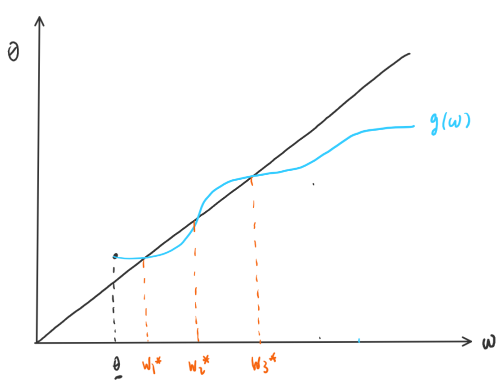


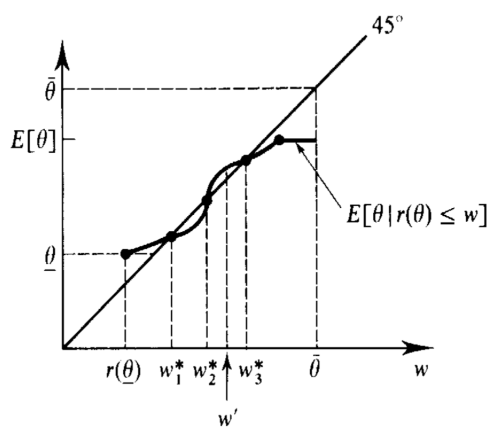


- ▷ We have (1) assuming $\underline{\theta} > r(\underline{\theta})$ and (2) since the conditional expectation becomes an unconditional expectation.
- ▷ A monotonically increasing function connecting (1) and (2) must be crossing the 45-degree line.

Essentially, this is the argument for Tarski's Fixed Point Theorem that applies to a monotonically increasing function. Note that we don't necessarily need continuity assumption. However, we cannot guarantee that the existence of an equilibrium or its uniqueness.

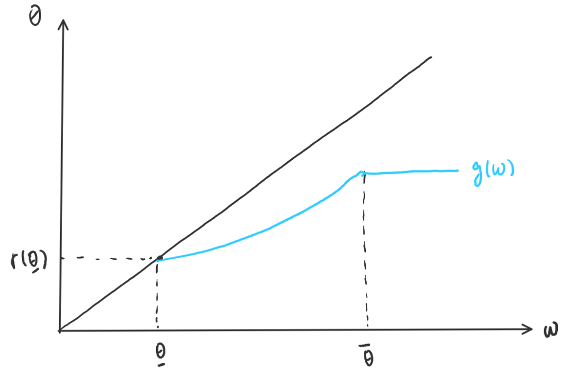
- ▷ Multiple Equilibria: There is no reason why $\mathbb{E}[\theta | r(\theta) \leq w]$ crosses the 45-degree line once. In this case, the Pareto optima choice is w_3^*.

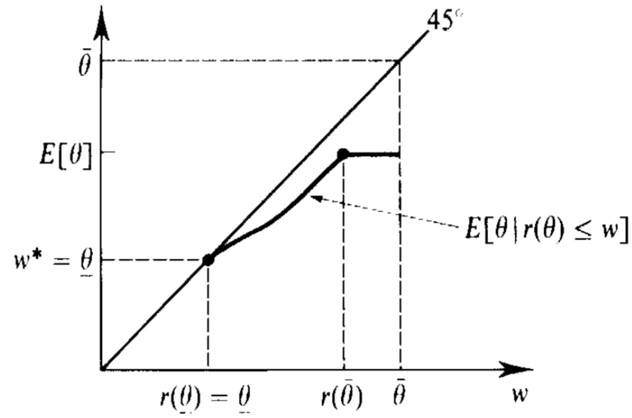




  * Firms are indifferent between the different equilibria since they all make zero profits, but the workers do care.
  * Note: There is a result in game theory that says you will always get an odd number of equilibria in games.

- ▷ Non-existence: You can have no equilibria if $\underline{\theta} = r(\underline{\theta})$:





What if we let the firms deviate? We will explore this in the next section.

Remark 2.1. We have adverse selection because $r(\theta)$ is increasing. If it is decreasing, we would have **advantageous selection.**

2.1.4 Modified Setup: Allowing Firms' Deviation

Now consider the following game where 1) firms choose wages, and 2) workers work at best wage or "UBER." Furthermore, denote

$$\bar{w}^* = \max_w CE_w$$

be the "best" equilibria in the mutliple equilibria case. Then we have the following result:

Proposition 2.1. *(MWG 13.B.1) If $\bar{w}^* > r(\underline{\theta})$ and $\exists \epsilon$ such that*

$$g(w') > w', \forall w' \in (\bar{w}^* - \epsilon, \bar{w}^*)$$

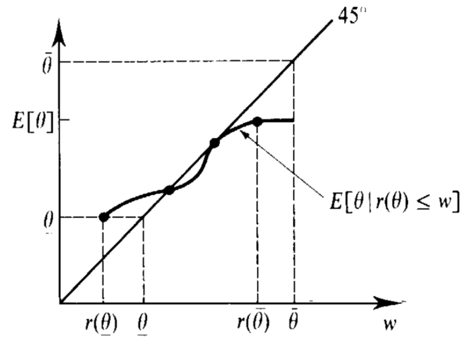
Then every SPNE has max wage equal to $\bar{w}^*$.

Proof. Broadly speaking, there are two steps.

1. From competitive equilibrium setup, it must be the case that all firms make zero profits. So the SPNE must be one of the equilibria.
2. You also cannot have an equilibrium left of the "bubble" since at the lower equilibrium, there is an incentive to deviate upward and make profit (where the $g(w)$ curve is above the 45-degree line)

Note that in the pathological case, there is no profitable upward deviation possible. This is why we needed that condition to rule out this case. ■

In other words, if you let the firms engage in the wage-setting game, they will naturally push you up to the "best" equilibrium. The condition also says that if you look left to $\bar{w}^*$, you have a portion of $g(w)$ cutting the 45-degree line from above. Notice that this rules out the following pathological case:



2.1.5 Can Goverment Improve the PO?

Suppose that the government understood all of the details of the adverse-selection market, except the government (like the firms) does not observe the workers' types. Specifically, can it find an allocation that is Pareto superior to a market allocation (weakly better for all and strictly better for some)? If not, then we say that the market allocation is constrained Pareto Optimal. Details aside, we have the following results:

Proposition 2.2. *(MWG 13.B.2) In words, there is nothing government can do – without knowing θ and maintaining budget balance – that can Pareto improve $\bar{w}^*$.*

Proof. Broadly speaking, this is because for all other people employed, the firms are breaking-even. So subsidizing makes these people worse off by construction. ■

2.2 Adverse Selection in Insurance Markets

Jehle and Reny has a model on insurance markets.

2.2.1 Setup

Consumers: risk $\pi \in [\underline{\pi}, \bar{\pi}] \equiv \Pi \subseteq [0, 1]$ which is the probability of an accident

- ▷ π is associated with cdf F
- ▷ loss L is same for everybody (once accident is realized)
- ▷ Expected loss = πL
- ▷ Risk-averse

Firms: risk-neutral, and only full-insurance policies can be sold.

2.2.2 Competitive Equilibrium

Now suppose consumers observe π while firms are uninformed. The π-type consumer's decision is the following – it is optimal to buy a full-insurance policy at price p if:

$$u(y-p) \geq \pi u(y-L) + (1-\pi) u(y)$$

▷ Rearranging, we have

$$\pi \geq \frac{u(y) - u(y-p)}{u(y) - u(y-L)} \equiv h(p)$$

▷ Note that $h(p)$ is increasing in p:

$$\frac{\partial h(p)}{\partial p} > 0$$

▷ The higher the price, the higher $h(p)$ so the higher the risk of the group that buys insurance. This is the setting of adverse selection because the consumer's who are most interested in buying full insurance are those with higher accident probabilities.

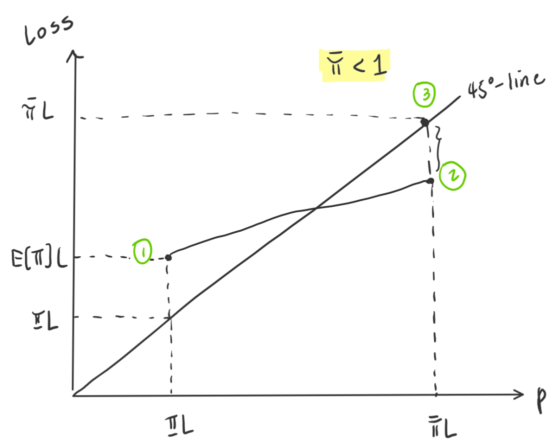
As in the case of the labor market setting, a competitive equilibrium is a price (p^*) and a set of types (Π^*) such that:

$$\begin{aligned} &(1) : p^* = \mathbb{E}[\pi | \pi \in \Pi^*] L \\ &(2) : \Pi^* = \{\pi | \pi \geq h(p^*)\} \\ &\Rightarrow p^* = \mathbb{E}[\pi | \pi \geq h(p^*)] L \end{aligned}$$

where

$$h(p) \equiv \frac{u(y) - u(y-p)}{u(y) - u(y-L)}$$

Furthermore, define $g(p) \equiv \mathbb{E}[\pi | \pi \geq h(p)] \cdot L$. $g(p)$ works similarly as $g(w)$ from the labor market example, and $g(p)$ is also increasing in p. We can draw an equilibrium similarly as before:



▷ In looking at this graph, we have to compare the type (π) against $h(p)$ to see if people buy the insurance or not.

▷ At (1) with $p^* = \underline{\pi}L$, everyone buys the insurance if

$$\underline{\pi} \geq h(\mathbb{E}[\pi] L)$$

▷ At (2) with $p^* = \bar{\pi}L$, the highest type $\bar{\pi}$ will definitely purchase the insurance. Since they would buy if $\pi \geq h(p)$, we have

$$h(\bar{\pi}L) \leq \bar{\pi}$$

which implies that

$$\mathbb{E}[\pi|\pi \geq h(p)]\, L \leq \bar{\pi}L$$

▷ If $\bar{\pi} = 1$, then we are guaranteed an inefficient equilibrium in which only the highest type $(\bar{\pi})$ purchase full insuarnce.

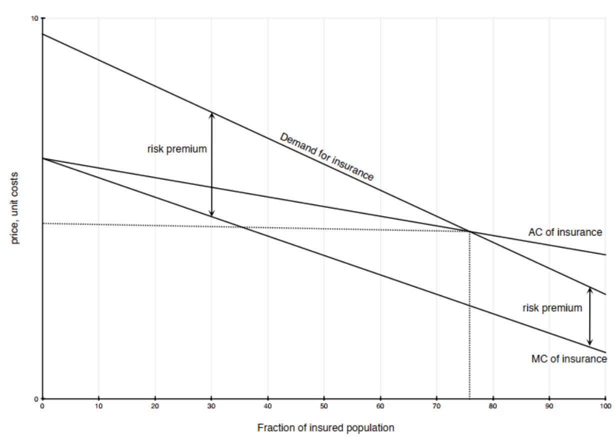
* This is in contrast with the labor market setting in which it was possible that the unique competitive equilibrium has full employment (i.e. when $r(\theta) = r < \mathbb{E}[\theta]$).

2.3 Advantageous vs. Adverse Selection

If selection is adverse, we mean that those who are willing to pay the most for insurance also have the highest risks. Alternately, if we assume high-risk individuals are also less risk-averse, then we have advantageous or beneficial selection.

2.3.1 Adverse Selection and Unit Cost Curve

We can think about putting the fraction of the population who buys insurance on the x-axis and the costs on the y-axis.



▷ Average Cost (AC) of insurance: Since the average cost of insuring the population of consumers with $\pi \in [\hat{\pi}, \bar{\pi}]$ is given as

$$\mathbb{E}[\pi L|\pi \geq \hat{\pi}]$$

he average cost of a policy falls as $\hat{\pi}$ decreases,

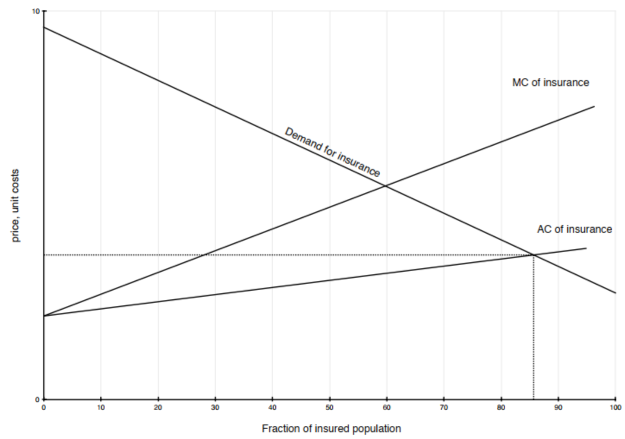
▷ Demand (D) for insurance: We can illustrate the demand as the fraction of the population who have insurance, i.e. $\pi \geq h(p)$ which is also a downward sloping curve.

▷ Marginal Cost (MC) is also falling, but it is everywhere below the AC curve because the AC falls.

▷ The market equilibrium is driven by the intersection of D and AC, but the efficiency requires coverage to occur at D and MC. Hence, the market provides too little insurance.

2.3.2 Advantageous Selection and Unit Cost Curve

Now assume advantageous selection. One example is when rich people buy insurance and also choose lower-risk environments, while poor people are more likely to pass on insurance and choose high-risk environments. In this case, the MC is higher than the AC as low-risk consumers are inframarginal. As a result, the average cost is upward sloping and marginal cost lies above average cost, which leads to the following figure:



3 Signaling

3.1 Signaling by Disclosing Certifiable Information

Given the Pareto inefficiencies caused by private information in the previous models of competitive markets, it seems natural that the existence of a costless signal or certification technology – either for worker productivities or consumer risks – might improve welfare. In this section, we assume that such a costless technology exists, but leave the control of the technology in the hands of the party with private information.

3.1.1 Setup

Consider the following setup:

1. Nature chooses θ_i for worker i
2. Worker can choose to reveal the type and it is credible (i.e. verifiable from the other side)
3. Firms choose wages conditional on certified information.

Theorem 3.1. *In every SPNE, full information is revealed.*

Proof. Suppose there is a pool of people (≥ 2) who do not reveal. Suppose there are two.

- ▷ Denote the higest value person θ_k and the other person θ_m. Then $\theta_k > \mathbb{E}\left[\theta_k \middle| \text{ no evidence}\right]$ so θ_k will reveal.
- ▷ Generally, the highest guy will want to deviate.

If there is only one guy, it's trivial. ■

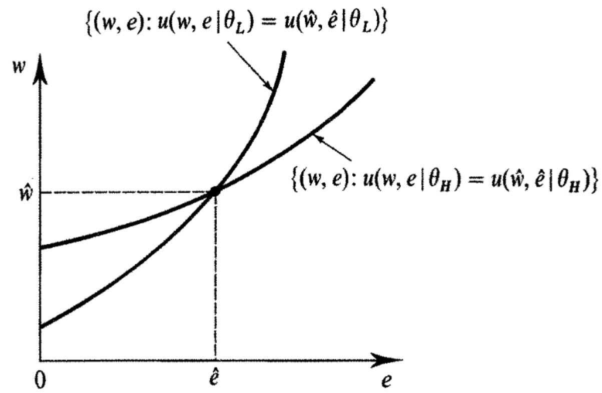
Every game has the following elements: each type can produce evidence of their type (credibility), and everyone wants to be thought "good" (monotonicity).

3.1.2 Toy Model of Labor Market (MWG)

Setup Workers: $u = w - c\left(e, \theta\right)$ where $c\left(e, \theta\right)$ is the cost of effort

- ▷ $c\left(e, \theta\right)$ is increasing and convex in e
- ▷ $c\left(0, \theta\right) = 0$
- ▷ $c\left(e, \theta\right)$ is decreasing in θ
- ▷ $c_{e\theta} < 0$ which gives us the single crossing property i.e. the higher productivity workers have lower marginal costs of education.

We also assume that workers take on two productivity values: $\theta_H > \theta_L > 0$ and that $r\left(\theta\right) = 0$ i.e. it is efficient for both types of workers to be employed. Graphically, the single-crossing property can be represented as:



▷ The marginal rate of substitution can be written as

$$MRS^L_{e,w} = c_e\left(e, \theta_L\right), \quad MRS^H_{e,w} = c_e\left(e, \theta_H\right)$$

▷ For the high productivity person, you don't have to compensate as much i.e. they have lower marginal costs of education.

Timing of the game:

▷ Workers choose $e \Rightarrow$ Firms offer wages conditional on $e \Rightarrow$ Workers accept / reject

The workers are looking for some strategy $e\left(\theta\right)$ for the high producivity and the low productivity guy. In the equilibrium, however, we have to write down what wage the firms are offering for any given level of education, $w\left(e\right)$. We will be finding the PBE / SE, so we also have to find the beliefs, denoted as $\mu\left(\theta_H|e\right) = \mu\left(e\right)$, and $w\left(e\right)$ has to be optimal given $\mu\left(e\right)$.

Separating Equilibria In solving for the separating equilibria, the following result is helpful:

Lemma 3.1. *In every separating equilibria,* $e^*_l = 0, w^*_L = e_L$ *and* $u^*_L = \theta_L$

Proof. You can't get the low guy to do any education, because the worst you can think of the low guy is that he's the low guy, so in a separating equilibria, he won't go to college. ■

Therefore, the conditions for separating equilibria are:

1. $w\left(e^*_H\right) = \theta_H$ and $w\left(e^*_L\right) = \theta_L$
2. $u^*_L \geq \theta_H - c\left(e^*_H, \theta_L\right)$ i.e. pretending to be the high guy is costly for the low guy.
3. $u^*_H \geq \theta_L$ i.e. pretending to be the low guy is costly for the high guy.

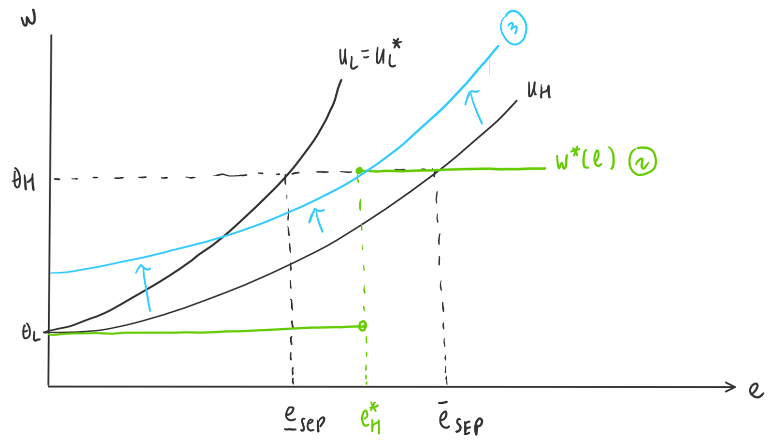
Construct the belief as

$$\mu^*\left(e\right) = \begin{cases} 0 & e < e^*_H \\ 1 & e \geq e^*_H \end{cases}$$

which immediately gives us:

$$w\left(e\right) = \begin{cases} \theta_H & e \geq e^*_H \\ \theta_L & e < e^*_H \end{cases}$$

Graphically, this is represented as



- ▷ The "best" one is $\underline{e}_{sep}$, but any $e_h \in \left[\underline{e}_{sep}, \bar{e}_{sep}\right]$ will do.
- ▷ The highest level of separating equilibrium $(\bar{e}_{sep})$ is determined by solving the θ_H's non-deviation constraint:

$$u_L^* = \theta_H - c\left(e_H^*, \theta_H\right)$$

Assuming $c\left(e, \theta\right) = e\left(k - \theta\right)$, then we have

$$\theta_L = \theta_H - e\left(k - \theta_H\right) \Rightarrow \bar{e}_{sep} = \frac{\theta_H - \theta_L}{k - \theta_H}$$

- ▷ The lowest level of separating equilibrium $\left(\underline{e}_{sep}\right)$ is determined by solving the θ_L's non-deviation constraint:

$$\theta_L = \theta_H - e\left(k - \theta_L\right) \Rightarrow \underline{e}_{sep} = \frac{\theta_H - \theta_L}{k - \theta_L}$$

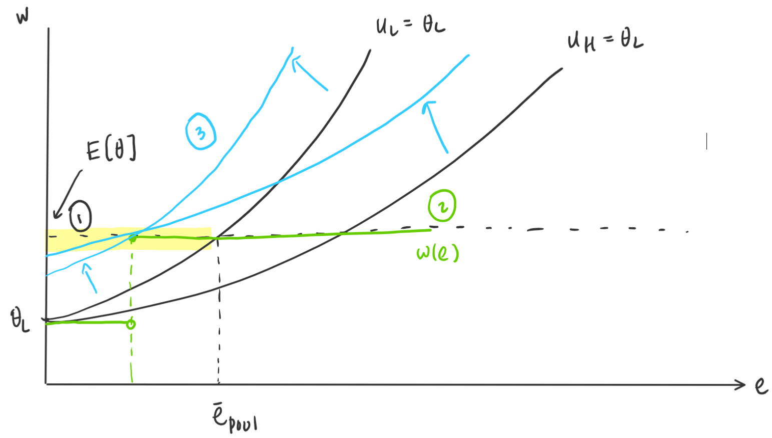
Pooling Equilibria We need the following conditions for pooling equilibria:

1. $w^*\left(e_p^*\right) = \mathbb{E}\left[\theta\right]$
2. $u_L^* \geq \theta_L$
3. $u_H^* \geq \theta_L$

Construct the beliefs as

$$\mu^*\left(e\right) = \begin{cases} \phi & \text{if e} \geq e_p^* \Rightarrow w^*\left(e\right) = \mathbb{E}\left[e\right] \\ 0 & \text{if } e < e_p^* \Rightarrow w^*\left(e\right) = \theta_L \end{cases}$$

Graphically, this is represented as



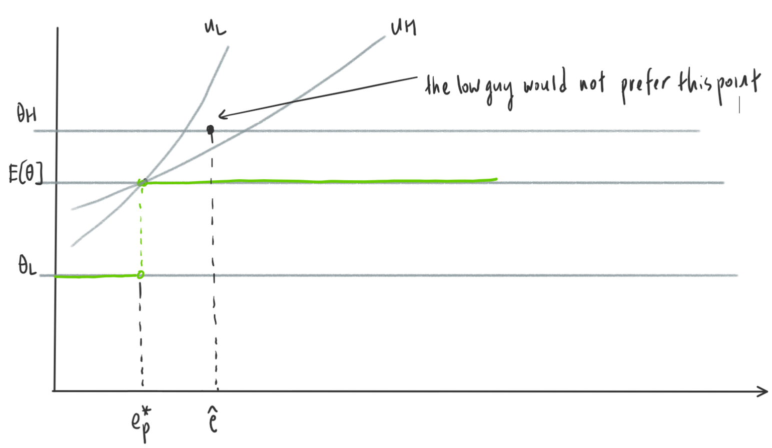
▷ If the wage is not in the yellow region, the low type will set $e = 0$.

▷ Note that e_p^* is the threshold that the firm uses and the level of effort for both high and low type.

$$\mathbb{E}[\theta] - e_p^* (k - \theta_L) \geq \theta_L \Rightarrow \bar{e}_p = \frac{\mathbb{E}[\theta] - \theta_L}{y - \theta_L}$$

▷ The Pareto-optimal pooling equilibrium is $e_p^* = 0$. This also corresponds to the no-signal outcome.

The pooling equilibrium is "unstable" since the high person can deviate to $\hat{e} > e_p^*$, a move that would only make sense for a high person to do. This is credible. The graphical depiction is shown below:



but since beliefs are pinned down, this will not happen in our setup.

Separating vs. Pooling We don't have one Pareto-dominating the other. But we can make some arguments. For example, if $\phi \approx 1$ i.e. most of the people are of the high type, they would prefer the pooling equilibrium since they will get something close to $\mathbb{E}[\theta] \approx \theta_H$ whereas in the separating equilibrium, they still have to pay a fixed cost.

3.2 Multiple Equilibria and Refinements

There are many, many equilibria in the labor-market signaling game. Are all reasonable?

3.2.1 Sender-Receiver Game

Consider the following setup:

- ▷ Nature chooses θ and shows to sender $\theta \in \Theta$
- ▷ Sender chooses a $m \in M$: $\sigma_S(m|\theta)$
- ▷ Receiver chooses action $a \in A : \sigma_R(a|m)$ with associated belief $\mu(\theta|m)$

The payoffs are $u^S(\theta, m, a)$ and $u^R(\theta, m, a)$ respectively.

3.2.2 Intuitive Criterion

We want the following criteria:

1. Define the set of people that will deviate since the payoff will be greater than their wildest dreams: $\hat{\Theta}(m) \in \Theta$ such that
$$\theta \in \hat{\Theta}(m) \Leftrightarrow u^*(\theta) > \max_{a \in BR(\Theta, m)} u^S(\theta, m, a)$$
This expression captures the idea in the graph shown previously. It now remains to define $BR(\cdot)$:
$$BR(\mu, m) \equiv \arg\max_{a \in A} \sum_\theta u^R(\theta, m, a)\,\mu(\theta|m)$$
and
$$BR(S, m) = \bigcup_{\mu \in \Delta(S)} BR(\mu, m)$$
for $S \in \Theta$ i.e. the all the possible best responses restricting our focus to $S \in \Theta$.

2. $\nexists \tilde{\theta}$ such that
$$u^*\left(\tilde{\theta}\right) < \min_{a \in BR\left(\Theta \setminus \hat{\Theta}(m), m\right)} u\left(\tilde{\theta}, m, a\right)$$
i.e. there exists no type such that worst possible choice by deviating gives more payoff than his current utility.

3.2.3 Refinements

Define the following:

$$D(\theta, m) = \left\{\sigma_R \in MBR(\Theta, m)\,|u^*(\theta) < \sum_a u^S(\theta, m, a)\,\sigma_R(a|m)\right\}$$

where MBR stands for mixed best respones i.e. all possible responses by the receiver that would make θ willing to deviate.

- ▷ D is the "deviation set" that is strictly attractive to type θ

Similarly,

$$D^\circ(\theta, m) = \left\{\sigma_R \in MBR(\Theta, m)\,|u^*(\theta) = \sum_a u^S(\theta, m, a)\,\sigma_R(a|m)\right\}$$

The $\boldsymbol{D1}$ **criterion** is the following:

$$\text{If } D(\theta, m) \cup D^\circ(\theta, m) \subseteq D\left(\tilde{\theta}, m\right), \text{ then } \mu(\theta|m) = 0$$

Essentially we are forcing beliefs to have zero probability on those less likely to deviate. The **intuitive criterion** version is:

$$\text{If } D(\theta, m) = \emptyset, D\left(\tilde{\theta}, m\right) \neq \emptyset, \text{ then} \mu(\theta|m) = 0$$

3.3 Revisiting the Insurance Model

We revisit the insurance model to find the pooling and the separating equilibria.

3.3.1 Setup

Recall that we had

$$1 > \pi_H > \pi_L > 0$$

where probability of high is ϕ. The utilities are defined as:

$$u_i\left(B, p\right) = \pi_i u\left(y - L + B - p\right) + \left(1 - \pi_i\right) u\left(y - p\right), \forall i \in \{H, L\}$$

▷ Nature chooses π ⇒ Consumer offers (B, p) ⇒ Firms accept (A) or reject (R)

The big difference is that the consumer is choosing both coverage B and price p whereas in the MWG game, the firm chooses the wage level. In this model, you thus have some equilibrium profits since there is an equilibrium where the consumer has to provide contracts that earn firms money which stems from bad beliefs.

We introduce a useful lemma:

Lemma 3.2. *Let (u_L^*, u_H^*) be the equilibrium payoffs to the two consumer types. Then:*

$$u_L^* \geq \underline{u}_L = \max_{(\pi_H B, B)} u_L\left(B, p = \pi_H B\right)$$
$$u_H^* \geq \underline{u}_H = \max_{(\pi_H B, B)} u_H\left(B, \pi_H B\right) = u_H\left(L, \pi_H L\right)$$

In other words, in any pure-strategy equilibrium, the players always have the option to offer an insurance contract that earns zero profit for π_H beliefs.

In this case, the high-risk consumer will purchase full insurance (because it is actuarially fair), and thus we have

$$u_H^* = u\left(L, \pi_H L\right)$$

but the low-risk consummer will typically buy less than full insurance, as the terms are not fair for π_L. The analog to the labor market setting is that each workers earns at least

$$\underline{u}_L = \underline{u}_H = \theta_L - c\left(0, \theta_i\right) = \theta_L$$

in any pure-strategy equilibrium.

Separating Equilibria We will write down the conditions and check that the conditions are indeed sufficient.

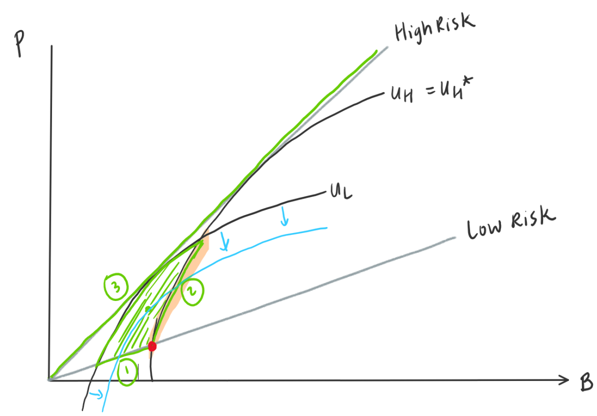
Proposition 3.1. *The policies $\psi_L = (B_L, p_L)$ and $\psi_H = (B_H, p_H)$ are accepted in a separating equilibrium if and only if*

1. *$B_H^* = L$ and $p_H^* = \pi_H \cdot L$ i.e. high-risk consumer purchases full-insurance*
2. *$p_L^* \geq \pi_L \cdot B_L^*$ i.e. low-risk policy earns non-negative profits*
3. *$u_H\left(L, \pi_H L\right) \geq u_H\left(B_L^*, p_L^*\right)$ i.e. high-risk type does not prefer low-risk contract*

4. *$u_L(L, \pi_H L) \geq \underline{u}_L$ i.e. from previous lemma*

▷ Note that the firm will accept $p \geq \pi_H \cdot B$. $p = \Pi_H B$ will always be accepted, so the high-risk guy will ask for $B_H^* = L$ and $p_H^* = \pi_H \cdot L$.

The separating equilibria can be shown as below:



▷ The (p, B) for the low guy:

* (1) comes from condition 2 since $p_L^* \geq \pi_L \cdot B_L^*$
* (2) so that H does not act like the L guy
* (3) so that L does not act like the H guy

▷ The firm will prefer any line on the orange, and the low guy will choose the red dot.

We have to finish specifying the beliefs:

$$\mu(\psi) = \begin{cases} 0 & \psi_H^* \\ 1 & \psi_L^* \\ 0 & \psi_H^* \neq \psi_L^* \end{cases}$$

i.e. if you don't choose the contract, I will assume you are the high risk guy. Thus, the acceptance decision will be:

$$\sigma(A|\psi) = \text{accept } \Leftrightarrow \psi_L, \psi_H, p \geq \pi_H \cdot L$$

Pooling Equilibria Now both consumers offer $\psi_p = \psi_L = \psi_H$ and the firm believes $\mu(\psi_p) = \phi$, the prior probability that the consumer is of low risk. Consequently, the firm will only accept such a policy if it is not profitable:

$$p \geq \mathbb{E}[\pi] B$$

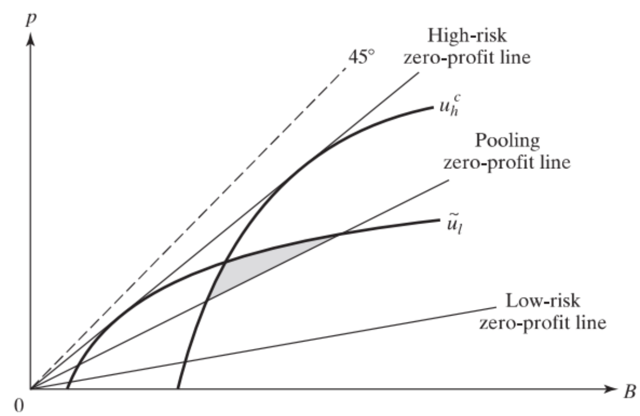
It's useful to do the analysis by introducing a zero-profit line for pooling contracts. It must satisfy the following conditions:

1. $\bar{p} \geq \mathbb{E}[\pi] \bar{B}$
2. $u_H^* \geq \underline{u}_H = u_H(L, \pi_H L)$

3. $u_L^* \geq \underline{u}_L = \max_{B \geq 0} u_L\left(B, \pi_H B\right)$

A large set of ψ_p may arise in any pooling equilibrium:

Figure 8.9. Pooling equilibria. The shaded region depicts the set of policies that can arise as pooling equilibria.



- ▷ As ϕ decreases (i.e. proportion of low risk consumers increases), the pooling zero-profit line gets closer to the high-risk zero-profit line. This shrinks the shaded area and eventually it disappears altogether.
- ▷ As ϕ increases, the shaded region expands as the slope of the pooling zero-profit line decreases. When ϕ is sufficiently large enough, there are pooling equilibria that make both consumers better off than the separating equilibrium, even the low-risk guy:

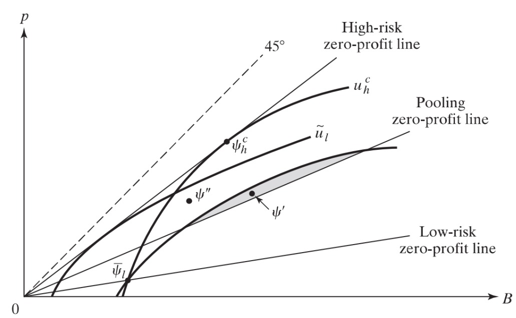


Figure 8.10. Pooling may dominate separation. The best separating equilibrium for consumers yields policies $\psi_l = \tilde{\psi}_l$ and $\psi_h = \psi_h^c$. The pooling equilibrium outcome $\psi_l = \psi_h = \psi'$ in the shaded region is strictly preferred by both risk types. Other pooling equilibrium outcomes, such as $\psi_l = \psi_h = \psi''$, are not.

Unlike the labor market signaling game, it is possible that firms make positive profit in the JR insurance-market game.

Application of the Intuitive Criterion Now we apply the intuitive criterion to refine our equilibria.

3.4 Cheap Talk

Our previous analysis of signaling games costly signals (e.g., education in labor markets and incomplete coverage insurance markets). Here, we consider the case where signals are costless (i.e., they have no direct payoff effects for either Sender or Receiver). One can therefore think of the signals as “Cheap Talk”. That said, is it possible that the signals/messages may nonetheless signal information?

3.4.1 Setup

Consider the following setup:

1. Nature chooses $\theta \in \Theta$ for S
2. S chooses $m \in M$ according to $\sigma_S(m|\theta)$
3. R forms a belief $\mu(\theta|m)$ and chooses an appropriate action $\sigma_R(a|m), a \in A$

We are looking for ***meaningful communication*** i.e. the mapping from Θ to A is not constant. For this, we need:

1. Sender's preferences over $a \in A$ must depend on θ.
2. Receiver's optimal action depends on θ
3. Commonality of preferences: if they are diametrically opposed, you will never have communication.

3.4.2 Crawford and Sobel (1982)

Their primary assumptions are:

- ▷ the set of messages is $\mathcal{M} = [0, 1]$ and there exists a n-type partition
- ▷ Sender of type θ randomly chooses a message m from the interval in the partition containing his type
- ▷ Receiver's beliefs $\mu(\theta|m)$ are prior beliefs conditioned on the partition of the message $m_i \in [x_i, x_{i+1}]$:

$$\mu(\theta|m_i) = \begin{cases} \frac{\mu(\theta)}{\mu([x_i, x_{i+1}])} & \text{if } \theta \in [x_i, x_{i+1}] \\ 0 & \text{otherwise} \end{cases}$$

  and given these beliefs, for each interval, we have a sequentially rational action:

$$\bar{a}^R([x_i, x_{i+1}]) \equiv \max_{a \in \mathbb{R}} \int_{[a_i, a_{i+1}]} u^R(\theta, a)\, dF(\theta)$$

  chosen.

As a final step, since we assume that a^S and a^R is increasing in θ, it suffices to check that at the boundaries, people (those with $\theta = x_i$) are indifferent between the two choices of the message:

$$u^S(x_i, \bar{a}_i) = u^S(x_i, \bar{a}_{i-1})$$

This is ensuring that the sender does not have any incentive to send a message outside the interval.

Example with Quadratic Preferences Suppose $u^S(\theta, a) = -(1/2)(a - (\theta + b))^2$ and $u^R(\theta, a) = -(1/2)(a - \theta)^2$.

- ▷ The babbling equilibrium always exists, i.e. the receiver chooses $1/2$ and the sender babbles.
- ▷ What about a two-partition equilibrium?
  - * The partitions are $[0, x_1)$ and $(x_1, 1]$.
  - * Sender's Actions:

- The actions are $\bar{a}_1 = x_1/2$ and $\bar{a}_2 = (1 + x_1)/2$. Still, x_1 is unknown.
- We also need $u^S(x_1, \bar{a}_1) = u^S(x, \bar{a}_2)$. Solving this yields:

$$-\frac{1}{2}\left(\frac{x_1}{2} - x_1 - b\right)^2 = -\frac{1}{2}\left(\frac{1 + x_1}{2} - x_1 - b\right)^2$$

which yields:

$$x_1 = \frac{1}{2} - 2b$$

Note that if $b \geq 1/4$, then the two-partition equilibrium cannot exist.

In short, for any b, we have a one-step equilibrium; for $b < 1/4$, we have a two-step equilibrium.

▷ What about n-partition equilibrium?

* Then we require:

$$x_1 + (x_1 + 4b) + (x_1 + 8b) + \cdots + (x_1 + (n-1)4b) = 1$$

which yields

$$n(n-1)2b < 1$$

This implies that the "closer" the two preferences are – the "smaller" the bias is – more equilibria we can generate.

3.5 Signal-jamming (Holmstrom 1982/99)

The name "signal-jamming" was coined by Fudenberg and Tirole (1986) in a paper about predation. The general idea is that the market generates observations that are noisy signals about the state of the world and a player may attempt to "jam these signals" so that another player would misinterpret the observations. We will

3.5.1 2-Period Setup

In this setup, the party strategically does something to mess up the other party's ability to learn.

▷ Manager of type $\theta \in \Theta$ where $\theta = \mu_\theta + \eta, \quad \eta \sim N(0, \sigma_\eta^2)$

▷ This is a two-period game:

* Period 1: Manager is employed
  - Manager chooses a_1 at cost $\psi(a_1)$
  - Market observes

$$X_1 = \theta + a_1 + \epsilon_1, \qquad \epsilon_1 \sim N(0, \sigma_\epsilon^2)$$

so the manager will work hard to generate a higher signal of his ability.

* Period 2: Market offers the wage:

$$w_2 = \mathbb{E}[X_2|X_1]$$

Note that $a_2^* = 0$ since the manager will die at period 2, so we have

$$w_2 = \mathbb{E}[\theta|X_1]$$

We know that

$$X_1 = \theta + a_1^* + \epsilon_1 \Rightarrow \theta = (X_1 - a_1^*) - \epsilon_1$$

We also know that

$$\theta = \mu_\theta + \eta$$

▷ So we will use these two signals to form a posterior of minimum variance:

$$\mathbb{E}\left[\theta|X_1\right]=\lambda\mu_\theta+(1-\lambda)\left(X_1-a_1^*\right), \qquad \lambda=\frac{\sigma_\epsilon^2}{\sigma_\epsilon^2+\sigma_\eta^2}$$

3.5.2 Solving the Model

At $t=2$, the worker has no incentive to choose any effort, so $a_2^*=0$. How should we determine a_1^*?

▷ Worker's payoff is

$$\begin{aligned} & -\psi\left(a_1\right)+\delta w_2 \\ = & -\psi\left(a_1\right)+\delta\left\{\lambda\mu_\theta+(1-\lambda)\left(X_1-a_1^*\right)\right\} \end{aligned}$$

Taking the first-order condition:

$$\psi'\left(a_1^*\right)=\delta(1-\lambda)=\frac{\delta\sigma_\eta^2}{\sigma_\epsilon^2+\sigma_\eta^2}>0$$

In other words, the first-period effort is positive (due to career concerns), but it's still lower than the first-best effort given by

$$\psi'\left(a^{FB}\right)=1$$

3.5.3 Extensions to T-period Model

In each subsequent period, the firm learns the additional signal

$$z_t=q_t-a_t^*\left(q^{t-1}\right)=\theta+\epsilon_t$$

which is distributed normally with mean θ and precision $1/\sigma_\epsilon^2$. The best estimate of θ after observing q^{t-1} for $t>1$ is easily determined to be:

$$\mathbb{E}\left[\theta|z_0,z_1,...,z_t\right]=\frac{h_\theta\mu_\theta+h_\epsilon\sum_{s=1}^{t-1}z_s}{h_\theta+th_\epsilon}$$

and the precision of this posterior is $h_t=h_{t-1}+h_\epsilon$.

▷ With large T , the firms learn the worker's type in the longrun.

4 Screening

4.1 Screening in Labor Markets

Now we let the uninformed parties move first and make offers to the informed side of the market. Such endogenous offers will (sometimes) separate or "screen" out the different types, and hence we refer to them as screening contracts.

4.1.1 Setup

There are two types $\theta_H > \theta_L > 0$ with probability ϕ of being ϕ_H. The timing is as follows:

1. The nature selects θ
2. Two firms $(n = 2)$ will offer contracts: $\left\{ \left(w_H^j, t_H^j\right), \left(w_L^j, t_L^j\right) \right\}$ for $j = 1, 2$
3. Workers will accept the best offer.

We also require the single crossing property: $c_{t\theta}(t, \theta) < 0$ i.e. higher types can do work more efficiently. Tasks have no value besides screening people. Finally, the type-θ worker's payoff from accepting a wage w and task t is

$$u(w|t, \theta) = w - c(t, \theta)$$

4.1.2 Complete Information Game

Suppose θ is public information. The result is that in equilibrium, it has to be the case that

$$(w_L^* = \theta_L, t_L^* = 0), \qquad (w_H^* = \theta_H, t_H^* = 0)$$

The proof is as follows:

▷ Wages cannot be below the productivity value, since the other firm can provide something ϵ less and grab all the workers. This shows $w_L^* \geq \theta_L$ and $w_H^* \geq \theta_H$.

▷ Wages cannot be above the productivity value since firms will incur a loss. This shows $w_L^* \leq \theta_L$ and $w_H^* \leq \theta_H$.

▷ Tasks cannot be positive since the other firm can offer a task level ϵ less and wage δ less such that the worker still prefers the new contract. This shows $t_L^* \leq 0$ and $t_H^* \leq 0$.

Note that $\{(\theta_H, 0), (\theta_L, 0)\}$ cannot be part of an equilibirum in an incomplete-information game since both types of workers would choose $(\theta_H, 0)$ and the firms would lose money on the θ_L-type workers.

4.1.3 Incomplete Information Game

If we offered the same contracts as before, it will not be a separating equilibrium since θ_L can pretend to be θ_H. (If tasks are productive, then it may have some separating properties by themselves). We derive the equilibrium through the following series of results:

1. In any pure-strategy equilibrium (pooling or separating), both firms earn zero profits in expectation.

   Proof. Suppose that in fact that the expected profit is positive for firm i.

   ▷ If firm j deviates as the following: $w'_H = w_H^* + \epsilon, w'_L = w_L^* + \epsilon$ then this firm captures all the workers. Since firm i was earning profits previously, this firm will capture all the profits.

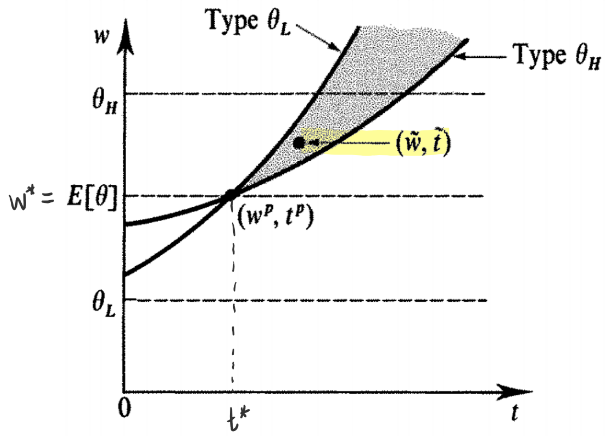
- ▷ Thus firm j will deviate and this cannot be sustained as an equilibrium.

■

2. No pooling equilibria exist. (See Figure)

Proof. Suppose that a pooling equilibrium exists.

- ▷ We know what the wage is $w^* = \mathbb{E}[\theta]$ and we know that the indifference curves have to pass through (w^*, t^*) since this is a pooling equilibrium.
- ▷ Consider $(\tilde{w}, \tilde{t})$ in the wedge. The high type θ_H has the incentive to deviate to $(\tilde{w}, \tilde{t})$ and thus this cannot be sustained as an equilibrium.

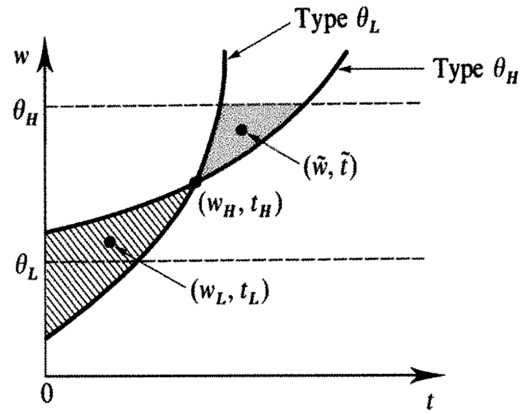


■

3. In any separating equilibrium, $w_H^* = \theta_H$ and $w_L^* = \theta_L$.

Proof. We consider two cases.

- ▷ Suppose $w_L^* < \theta_L$.
  - * We are making zero profits in equilibrium. Firm j thus has the incentive to deviate and offer $(w_L^* + \epsilon, t_L)$ thereby making profit off of the low type. Maybe the high type will come in and this will even further reinforce firm j's profit.
  - * Thus, it has to be the case that $w_L^* \geq \theta_L$.
- ▷ Combined with result (1), this implies $w_H^* \leq \theta_H$.
  - * In other words, I can't make money on the low guy, and I know I have to be net zero, so I need to make money on the high guy.
- ▷ Now we know that $w_H^* \leq \theta_H$. We show that $w_H^* < \theta_H$ is not possible.
  - * Suppose it is the case. Then the firm at (w_H, t_H) has the incentive to deviate to $(\tilde{w}, \tilde{t})$.

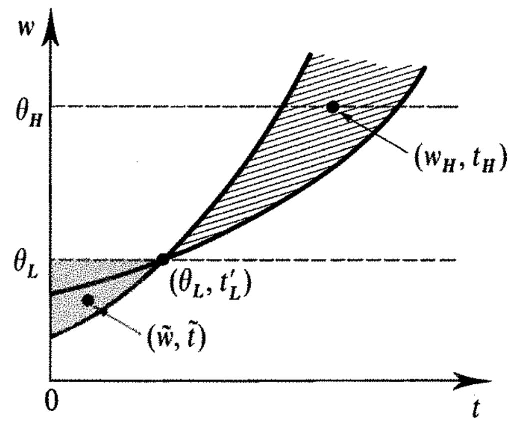


■

4. In any separating equilibrium, $(w_L^*, t_L^*) = (\theta_L, 0)$

   Proof. We show that t_L^* cannot be positive.

   ▷ If it is positive, then the firm at (θ_L, t_L^*) has the incentive to deviate to $\left(\tilde{w}, \tilde{t}\right)$ since the θ_L person prefers it,and the firm makes profits on the θ_L guy, and nothing changes for the θ_H guy.

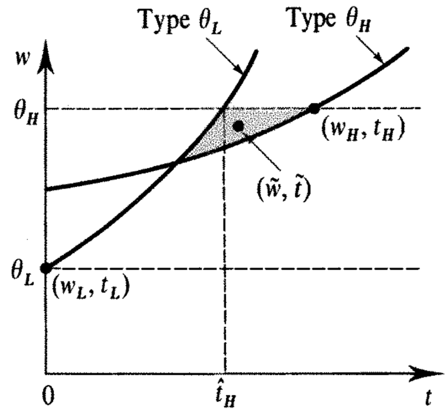


■

5. Define $\hat{t} : \theta_L - c\,(0, \theta_L) = \theta_H - c\left(\hat{t}, \theta_L\right)$ i.e. the task level that makes the low guy indifferent to pretending to be the high guy. Then the high-ability workers are offered and accept $\left(\theta_H, \hat{t}\right)$ i.e. $\hat{t} = t_H^*$.

   Proof. Note that $\hat{t}$ is the intersection between the θ_L indifference curve starting at $(0, \theta_L^*)$ and $w = \theta_H$ horizontal line.

   ▷ Suppose $t_H^* > \hat{t}$. Then the high type has an incentive to deviate to $\left(\tilde{w}, \tilde{t}\right)$. This drives $t_H^* = \hat{t}_H$.



■

Combining the preceding results, we have that in any SPE, firms offer the menu

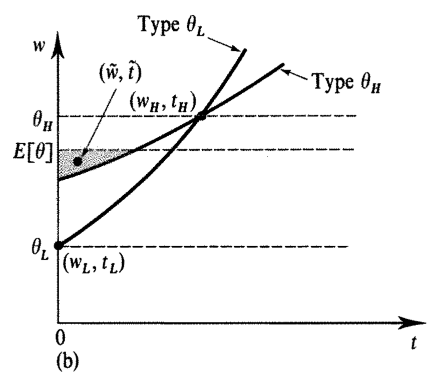
$$\{(\theta_H, t_H), (\theta_L, 0)\}$$

where t_H satisfies:

$$\theta_H - c(t_H, \theta_L) = \theta_L - c(0, \theta_L)$$

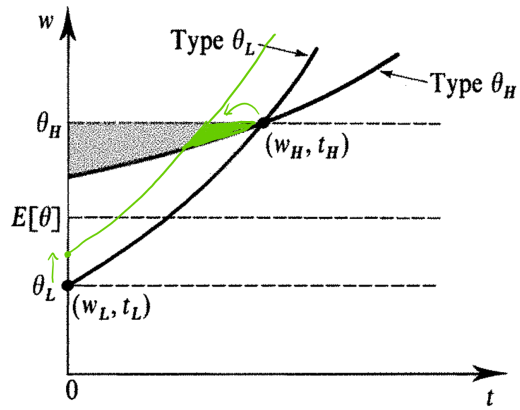
and low-ability workers choose $(\theta_L, 0)$ while high-ability workers choose (θ_H, t_H).

▷ Note that separating equilibria can fail to exist when the probability of the high-ability type is sufficiently high that a pooling contract can be constructed which is favorable to both types.



* As ϕ becomes sufficiently close to 1, $E[\theta]$ becomes close to θ_H and things get fucked up (in the second panel of the pictture below).

▷ In the following picture, the deviation is not obvious immediately because $E[\theta]$ is below the indifference curve for type θ_H. Specifically, we see that the firm can't deviate to the shaded region since the low guy will also take advantage of this new contract.



(a)

▷ In this case, you can bump up the wage of the w_L guy (and incur loss on the low-type), but you can also reduce the task for the high guy and lower the wage . This deviation can also happen.

In other words, there are multiple ways in which the separating equilibria may not exist.

Extensions We can refine the concept of deviation.

▷ Inderst-Wambach (EER 2001) show that you can get an adverse selection upon deviation. In other words, only the people who have a lot to gain will show up when a deviated contract is presented.

4.2 Screening in Insurance Markets

We now consider the associated screening game in which the (two) insurers move first and offer contracts, followed by the informed consumers' choice from the available offers.

4.2.1 Setup

We have the following setup:

1. Two insurance companies simultaneously offer a menu of offers:
$$\left\{\Psi_L^j = \left(B_L^j, p_L^j\right), \Psi_H^j = \left(B_H^j, p_H^j\right)\right\}$$
2. Nature chooses the consumer's type π with probability ϕ that the consumer is low-risk and $(1-\phi)$ probability that the consumer is high-risk.
3. Consumer chooses a single policy from the presented menus or chooses no insurance.

4.2.2 Complete Information Game

Under full information, consumer's type is known at stage 1 when offers are chosen by the insurers. Thus, the low-risk consumer will be offered
$$(L, p_L = \pi_L L)$$
by both firms and high-risk consumer will be offered
$$(L, p_H = \pi_H L)$$
by both firms.

4.2.3 Results

We present the results with minimal proofs. We focus on pure-strategy equilibria.

1. In every pure-strategy equilibrium, expected profits are zero.

   Proof. In pooling, this is easy to prove.

   ▷ In separating equilibrium: WLOG, the following must be true:

   $$u_L(B_L^*, p_L^*) > U_L(B_H^*, p_H^*)$$
   $$u_H(B_H^* + \delta, p_H^*) \geq U_H(B_L^* + \epsilon, p_L^*)$$

   ▷ Consider $B_L' = B_L^* + \epsilon$ and $B_H' = B_H^* + \delta$ where ϵ, δ are arbitrarily chosen.

   ▷ Then we can have

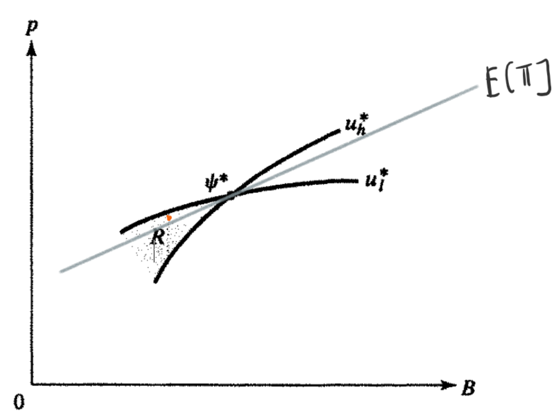
   $$u_L(B_L^*, p_L^*) > U_L(B_H^*, p_H^*)$$
   $$u_H(B_H^* + \delta, p_H^*) > U_H(B_L^* + \epsilon, p_L^*)$$

   which implies that the deviation allows the firm to grab all profits. ■

2. There is no pooling equilibrium.

   Proof. Let (B^*, p^*) be the pooling policy. Since expected profits are zero, it must be that

   $$p^* = (\phi\pi_L + (1-\phi)\pi_H) B^*$$



   In this case, a deviating firm can offer a single contract in the shaded region. Since (B^*, p^*) generates zero expected profit when both types purchase it, it yields strictly positive profits when only the low-risk consumer purchases it. ■

3. In any separating equilibrium, the high-risk consumer must obtain utility at least $u_H(L, \pi_H L)$.

   Proof. If not, the deviating firm could offer the high-risk guy $(L, \pi_H L + \epsilon)$ and earn ϵ as profit. ■

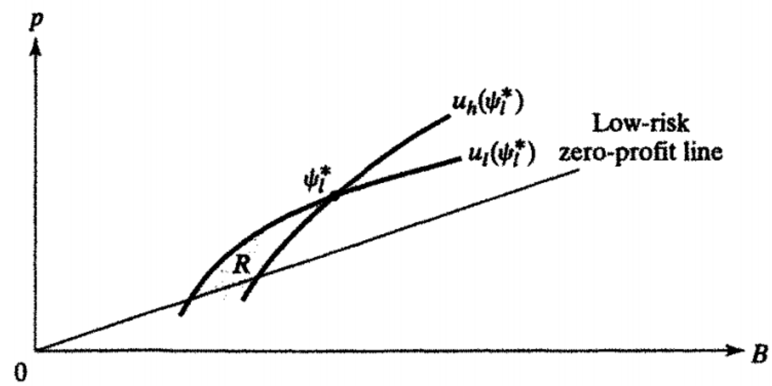
4. In any separating equilibrium, it must be the case that the low-risk consumer's contract must lie on the low-risk zero-profit line: $p_L^* = \pi_L B_L^*$.

Proof. First, note that (B_L^*, p_L^*) must lie on or above the low-risk zero-profit line.

▷ This is because the high-risk type must earn $u_H(L, \pi_H L)$ in payoff, which implies that (B_H^*, p_H^*) must lie on or below the high-risk zero-profit line.

▷ This implies that the firm earns non-positive profits on the high-risk guy, so it must earn non-negative profits on the low-risk guy.

Suppose it lies strictly above. Then there is a shaded region R where one firm could offer a single policy in the shaded region and attract only the low-risk type, making positive profit. ■



5. In any separating equilibrium, the high-risk consumer's contract is $(B_H^*, p_H^*) = (L, \pi_H L)$.

Proof. Since we showed that firms earn zero profits for the low-risk guy, it must also earn zero profits on the high-risk guy. ■

6. In any separating equilibrium, the low-risk consumer's contract is characterized as

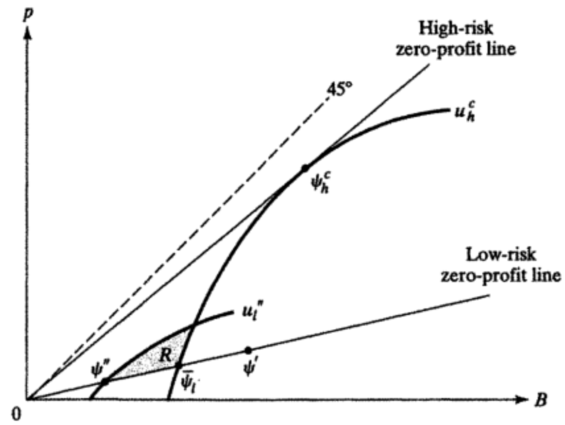
$$u_H(B_L, \pi_L B_L) = u_H(L, \pi_H L)$$

Proof. We want to show that only one B_L^* can arise in any equilibrium.

▷ Suppose B_L^* does not satisfy the hypothesis:

$$u_H(B_L, \pi_L B_L) < u_H(L, \pi_H L)$$

▷ Suppose ψ'' is chosen by the low-risk type. This gives a rise to a shaded region R that is west of the u_H indifference curve where a deviating firm can offer a contract and attract the low-risk type, earning positive profit.



■

Pooling equilibria are also susceptible to deviations:

Combining the preceding results, we have that in any SPE, the chosen contracts are

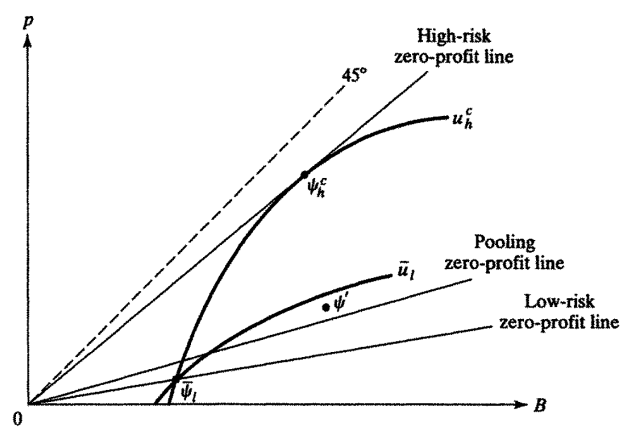
$$(B_H, p_H) = (L, \pi_H L)$$

and $(B_L, p_L) = (B^*, \pi_L B^*)$ where B^* satisfies

$$u_H(B^*, \pi_L B^*) = u_H(L, \pi_H L)$$

Of course, this does not imply that a separating equilibrium exists. It only says that if one does exist, then it necessarily is the least-cost separating equilibrium.

Existence As in the case of labor-market screening games, if the probability that the informed type is the good type (i.e. low risk) is sufficiently high, then a pooling contract deviation exists which prevents separation from arising in equilibrium.



▷ Both firms and the low-risk type have an incentive to deviate from $\bar{\psi}_l$ to ψ'.

Extensions

▷ Non-exclusive contracts: $p = \pi_H B$ akin to Akerlof's lemon model.

5 Moral Hazard

We analyze the contracting problem in which the principal hires the agent to performa task, and the agent chooses the effort intensity which affects performance. The principal only cares about the performance, and since the effort is costly to the agent, the principal has to compensate the agent for incurring these costs. When effort is unobservable, the best principal can do is to relate compensation to performance, and this compensation scheme will typically entail a loss since performance is only a noisy signal of effort.

5.1 Preliminaries

The basic moral hazard problem has a fairly simple structure, yet general conclusions have been difficult to obtain. The characterization of optimal contracts in the context of moral hazard is still somewhat limited, and typically applications have put more structure on the moral hazard problem under consideration, thus enabling a sharper characterization of the optimal incentive contract.

5.1.1 Setup

The timing is as follows:

1. P offers A: $w(x) : X \to [\underline{w}, \bar{w}] \subseteq \mathbb{R}$
2. A accepts or rejects. If he accepts, he can choose $e \in \mathcal{E}$ which improves the distribution of x.
3. x is realized according to $F(X|e)$

The payoffs are:

$$\begin{aligned} [P] &: V = v(x - w(x)) = x - w(x) \text{ : risk netural} \\ [A] &: U = u(w(x)) - \psi(e) : \text{ risk averse} \end{aligned}$$

Principal's problem:

$$\begin{aligned} \max_{e \in \mathcal{E}, w \in [\underline{w}, \bar{w}]} & \mathbb{E}[x - w(x)|e] \equiv \int_{\mathcal{X}} (x - w(x)) f(x|e)\, dx \\ \text{s.t. } & \mathbb{E}[u(w(x))|e] - \psi(e) \geq \underline{U} \qquad [IR] \\ & e \in \arg\max_{\tilde{e} \in \mathcal{E}} \mathbb{E}[u(w(x))|\tilde{e}] - \psi(\tilde{e}) \qquad [IC] \end{aligned}$$

▷ The agent's expected utility from choosing e given contract w is

$$\mathbb{E}[u(w(x))|e] - \psi(e) = \int_{\mathcal{X}} u(w(x)) f(x|e)\, dx - \psi(e)$$

▷ If the value of this program/problem is negative, then the agent will not choose this (= IR constraint)

We also assume that effort (e) improves the distribution of output in the FOSD sense:

$$\tilde{e} > e \Leftrightarrow F(x|\tilde{e}) \succsim_{FOSD} F(x|e), \forall e \in \mathcal{E}$$

5.1.2 Examples with No Moral Hazard

Here are some examples in which moral hazard does not prevent the first-best outcome from happening.

1. Contractable e

   ▷ IC disappears, and the principal can pay $w = u^{-1}\left(\bar{U} + \psi(e)\right)$ i.e. set wage as low as possible to satisfy the agent's participation.

   ▷ The principal can now choose to solve

   $$\max_e \mathbb{E}[x|e] - u^{-1}\left(\underline{U} + \psi(e)\right)$$

   Taking the FOC with respect to e:

   $$\int_{\mathcal{X}} x f_e(e|x)\, dx = \frac{1}{u'(w)} \psi'(e)$$

   ▷ We will refer the resulting $e^* = e^{fb}$. Notice that if agents were risk-neutral as well, then $u'(w) = 1$ and the first-best effort corresponds to what the principal would choose if she directly chose effort at her own personal cost of $\psi(e)$.

2. No uncertainty

   ▷ For every x you want to implement, there is a corresponding e i.e. $e = G^{-1}(x)$

   ▷ Then the principal's maximization problem is

   $$x - u^{-1}\left(\bar{U} + \psi\left(G^{-1}(x)\right)\right)$$

   which is essentially the same as the problem from before.

3. Agent is risk-neutral i.e. $u''(\cdot) = 0$.

   ▷ The first-best level of output can then be implemented by essentially selling the enterprise to the agent in the contract w. Specifically, the principal offers the agent the following sales contract:

   $$w(x) = x - \pi + \underline{U}$$

   where π represents the value of the enterprise when run efficiently:

   $$\pi = \max_e \mathbb{E}[x|e] - \psi(e)$$

   ▷ Why does this wage schedule work?

   * If the agent accepts and chooses the first-best effort defined by

     $$\int_{\mathcal{X}} x f_e(e|x)\, dx = \psi'(e)$$

     then she receives in expectation

     $$\mathbb{E}[w(x)|e] = \max_e \mathbb{E}[x|e] - \psi(e) - \pi + \underline{U} = \underline{U}$$

     so the agent accepts and all surplus is captured by the principal.

4. Shifting support

- ▷ The support is given as $\mathcal{X}$ but suppose that if you work really hard, the lower values of x will never be materialized. In other words, the support of x shifts as a function of e. Then e may be indirectly contractible, and the first-best may be implementable.
- ▷ For example, suppose x is uniformly distributed on $[e, e+K]$, $K > 0$, in which the principal can implement any $\hat{e}$ by forcing a contract

$$w(x) = \begin{cases} \underline{w} & \text{if } x \geq \hat{e} \\ -\infty & \text{otherwise} \end{cases}$$

where $u(\bar{w}) = \underline{U} + \psi(\hat{e})$.

5.1.3 Solving the Model

Let $\mathcal{E} = \{e_L, e_H\}$ with $\psi(e_L) = 0$ and $\psi(e_H) = \Delta$ and $\underline{U} = 0$. There are two possibilities: $e^* = e_L$ or $e^* = e_H$. Recall the principal's problem:

$$\begin{aligned} \max_{e \in \mathcal{E}, w \in [\underline{w}, \bar{w}]} & \mathbb{E}[x - w(x) | e] \equiv \int_{\mathcal{X}} (x - w(x)) f(x|e) \, dx \\ \text{s.t. } & \mathbb{E}[u(w(x)) | e] - \psi(e) \geq \underline{U} \qquad [IR] \\ & e \in \arg\max_{\tilde{e} \in \mathcal{E}} \mathbb{E}[u(w(x)) | \tilde{e}] - \psi(\tilde{e}) \qquad [IC] \end{aligned}$$

We can consider the following two cases:

1. Principal desires to implement e_L.
   - ▷ Immediately, the $[IC]$ constraint can be ignored since the agent cannot choose an effort lower than e_L. And since incentives are unneeded, it is optimal to insure the agent against wage risk.
   - ▷ When the agent implements e_l, the payoff to the principal is

   $$\mathbb{E}[x|e_L] - \bar{w} = \mathbb{E}[x|e_L] - u^{-1}(\underline{U})$$

   where $\bar{w}$ is implicitly defined by $u(\bar{w}) = \underline{U} + \psi(e_L) = \underline{U}$, which implies a wage offer of

   $$\bar{w} = u^{-1}(\underline{U} + \psi(e_L)) = u^{-1}(0) = 0$$

2. Principal wants to implement e_H.
   - ▷ What is the cheapest way to do this?

   $$\begin{aligned} \max_{w(\cdot), e} & \mathbb{E}[x - w(x) | e_H] \\ \text{s.t. } & \mathbb{E}[u(w(x)) | e] - \psi(e) \geq \underline{U} \qquad [IR] \\ & \mathbb{E}[u(w(x)) | e_H] - \Delta \geq \mathbb{E}[u(w(x)) | e_L] \qquad [IC] \end{aligned}$$

     * It turns out that you can covert this to a convex-linear program by defining

     $$z(x) \equiv u(w(x))$$

     which allows us to write:

     $$w(x) = u^{-1}(z(x)), \quad z(x) \in [u(\underline{w}), u(\bar{w})]$$

* This allows us to rewrite the program as

$$\begin{aligned}
&\max_{z(\cdot)} \mathbb{E}\left[x - u^{-1}\left(z\left(x\right)\right)|e_H\right] \\
&\text{s.t. } \mathbb{E}\left[z\left(x\right)|e_H\right] - \Delta \geq \mathbb{E}\left[z\left(x\right)|e_H\right] \qquad [IC] \\
&\quad \mathbb{E}\left[z\left(x\right)|e_H\right] - \Delta \geq \underline{U} \qquad [IR]
\end{aligned}$$

▷ Let's suppose that a solution exists for now. Writing out the Lagrangian:

$$\begin{aligned}
\mathcal{L} = &\int_{\underline{x}}^{\bar{x}} \left[x - w\left(x\right)\right] f\left(x|e_H\right) dx \\
&+ \mu\left(\int_{\underline{x}}^{\bar{x}} u\left(w\left(x\right)\right)\left(f\left(x|e_H\right) - f\left(x|e_L\right)\right) dx - \Delta\right) \\
&+ \lambda\left(\int_{\underline{x}}^{\bar{x}} u\left(w\left(x\right)\right) f\left(x|e_H\right) - \Delta - \underline{U}\right)
\end{aligned}$$

▷ We will fix x and maximize over $w\left(x\right)$ pointwise, leading to:

$$-f\left(x|e_H\right) + \mu\left(u'\left(w\left(x\right)\right)\right)\left(f_H\left(x\right) - f_L\left(x\right)\right) + \lambda u'\left(w\left(x\right)\right) f_H\left(x\right) = 0, \quad \forall x$$

Leading to

$$\underline{\frac{1}{u'\left(w\left(x\right)\right)} = \lambda + \mu\left(\frac{f_H\left(x\right) - f_L\left(x\right)}{f_H\left(x\right)}\right)}$$

where $f_H\left(x\right) = f\left(x|e_H\right)$ and $f_L\left(x\right) = f\left(x|e_L\right)$.

▷ Using this result, we can say a lot.

* $\lambda > 0$: If $\lambda = 0$, then the agent's IR constraint is slack. In this case, the principal could reduce the agent's wage schedule from $w\left(x\right)$ to $w_\epsilon\left(x\right)$ where $u\left(w\left(x\right)\right) - \epsilon = u\left(w_\epsilon\left(x\right)\right)$. Then this reduces expected wage payments without impacting the agent's incentives to choose e_H, so we have a contradiction.
* $\mu > 0$: If $\mu = 0$, then wage would be constant (since the RHS is independent of x). Then the agent will not exert effort and $e = e_L$, a contradiction.
* To implement high effort, wages should be higher when the output is more indicative of f_H than f_L. This is NOT the same as saying $w\left(x\right)$ is increasing in output.
* If the principal were also risk averse, the condition is modified to:

$$\frac{v'\left(x - w\left(x\right)\right)}{u'\left(w\left(x\right)\right)} = \lambda + \mu\left(\frac{f_H\left(x\right) - f_L\left(x\right)}{f_H\left(x\right)}\right)$$

5.1.4 Monotone-Likelihood Ratio Property (MLRP)

If $f\left(x|e\right)$ is differentiable in e, then we say that the distribution satisfies MLRP if and only if for any e,

$$\frac{f_e\left(x|e\right)}{f\left(x|e\right)} \text{ is increasing in } x$$

If e is discrete, then this is equivalent to

$$\frac{f_H\left(x\right) - f_L\left(x\right)}{f_H\left(x\right)} \text{ is increasing in } x$$

Note that Milgrom (1981) also has an alternative definition of MLRP

$$x \succ \tilde{x} \rightarrow G\left(e|x\right) \succsim_{FOSD} G\left(e|\tilde{x}\right)$$

where $G\left(e|x\right)$ is a posterior distribution formed upon observing x. This condition can be interpreted as: higher x provides good news about the agent's efforts.

- ▷ If x is distributed normally with mean e, then the distribution satisfies MLRP.
- ▷ We can show that MLRP ⇒ FOSD, but not the reverse. Note that the expectation of f_e/f is zero:

$$\begin{aligned}\int_{\underline{x}}^{\bar{x}} \frac{f_e\left(x|e\right)}{f\left(x|e\right)} f\left(x|e\right) dx &= \int_{\underline{x}}^{\bar{x}} f_e\left(x|e\right) dx \\ &= \frac{\partial}{\partial e} \int_{\underline{x}}^{\bar{x}} f\left(x|e\right) dx = 0\end{aligned}$$

so for any $x < \bar{x}$, we have

$$\mathbb{E}\left[\frac{f_e\left(x|e\right)}{f\left(x|e\right)}\right] = \int_{\underline{x}}^{x} \frac{f_e\left(x|e\right)}{f\left(x|e\right)} f\left(x|e\right) dx \equiv F_e\left(x|e\right) < 0$$

which is the definition of FOSD.

5.1.5 Non-existence

Assume that $x = e + \epsilon, \quad \epsilon \sim N\left(0, \sigma^2\right)$.

- ▷ The normal function has the following property:

$$\lim_{x\rightarrow -\infty} \frac{f_L\left(x\right)}{f_H\left(x\right)} = +\infty, \quad \lim_{x\rightarrow -\infty} \frac{F_L\left(x\right)}{F_H\left(x\right)} = +\infty$$

where f_L has a lower mean than f_H. Economically, this means that a lower value of x gives you more information.

- ▷ Take a risk-averse person. If I give some really low wage with some probability, but if the probability goes to zero more quickly than your risk aversion, then you can actually approximate the first-best solution (the probability decreases so quickly that it does not bother the risk-averse individual). This is the idea in Mirrlees (1975 / 1999).

To see this more clearly, assume the following wage schedule

$$w\left(x\right) = \begin{cases} w_1 & \text{if } x \geq k \\ w_0 & \text{if } w < k \end{cases}$$

where w_0, w_1, k are to be determined by the principal.

- ▷ The idea is that one can make simultaneously make k and w_0 very low while satisfying the IC and IR constraints as low output is very, very indicative of low effort. This makes the IC constraint arbitrarily cheap to satisfy, even with a risk-averse agent.
- ▷ Suppose the principal wants to induce e_H using the contract above. The $[IC]$ and $[IR]$, simplified (and suppoinsg that tbey both bind), are:

$$\begin{aligned}[IC] &: \left(1 - F\left(k|e_H\right)\right) u\left(w_1\right) + F\left(k|e_H\right) u\left(w_0\right) - \Delta = 0 \\ [IR] &: \left[F\left(k|e_L\right) - F\left(k|e_H\right)\right]\left(u\left(w_1\right) - u\left(w_0\right)\right) = \Delta\end{aligned}$$

Substituting out w_0 so that the constraint only depend on (w_1, k):

$$[IC\&IR] : \left(\frac{F\left(k|e_L\right)}{F\left(k|e_H\right)} - 1\right)\left(u\left(w_1\right) - \Delta\right) = \Delta$$

- ▷ Now consider a sequence of $k \to \infty$. For each k, if we choose w_1 according to the above equation, we guarantee e_H will be chosen and the contract accepted by the agent.

- ▷ In such case, the expected wage paid to the agent is

$$(1 - F(k|e_H))\, w_1 + F(k|e_H)\, w_0 \leq w_1$$

but as k approaches $-\infty$, the ratio $F(k|e_L)/F(k|e_H)$ approaches infinity, and thus $\underline{u(w_1) \to \Delta}$ i.e. the wages get infinitely close to the first-best wage, but not actually get there. In other words, the solution does not exist.

5.1.6 Informativeness Principle

So far, we assumed that there is a single signal x that the principal can use in her contract with the agent. Suppose instead that two outcomes were observed: x and y. When will the principal want to make the wage schedule depend upon both variables?

Start by using $f(x, y|e)$ as the joint distribution, and suppose that the principal wishes to implement e_H rather than e_L. Then she will solve

$$\begin{aligned}
&\min_{w(\cdot)} \int_{\mathcal{X}} \int_{\mathcal{Y}} w(x,y)\, f(x,y|e_H)\, dydx \\
&\text{such that} \int_{\mathcal{X}} \int_{\mathcal{Y}} u(w(x,y))\, f(x,y|e_H)\, dydx - \psi(e_H) \geq \underline{U} \\
&\int_{\mathcal{X}} \int_{\mathcal{Y}} u(w(x,y))\, f(x,y|e_H)\, dydx - \psi(e_H) \geq \int_{\mathcal{X}} \int_{\mathcal{Y}} u(w(x,y))\, f(x,y|e_L)\, dydx - \psi(e_L)
\end{aligned}$$

Assuming a solution exists, we obtain the following condition:

$$\frac{1}{u'(w(x,y))} = \lambda + \mu \left(1 - \frac{f(x,y|e_L)}{f(x,y|e_H)}\right)$$

which implies that if f_L/f_H is independent of y, then $w(x,y)$ is independent of y .

In other words, if "x is sufficient for y with respect to e" i.e.

$$f(x,y|e) = g(x|e)\, h(y|x)$$

then we have that the wage w is independent of y since $h(y|x)$ will be cancelled from both the numerator and the denominator.

- ▷ Informativeness principle: An additional signal y is valuable for incentives if and only if x is not sufficient for y with respect to e i.e. y carries additional information about the agent's effort that is not contained in x .

- ▷ This result also implies that randomization is not optimal. Randomization will never come into these contracts due to the sufficiency principle – randomization (y) does not convery any additional information about the effort. (Note that this result does rely on our assumption that the agent's preferences are additively separable in effort and wage).

- ▷ Relative performance evaluation is optimal when the second agent's output is informative about the first agent's effort.

5.2 Solving the General Model

In this section, we will show that if the first-order approach is valid, then everything goes through smoothly. Then we will characterize the conditions for the first-order approach to be valid.

5.2.1 The First-order Approach

Principal's problem:

$$\begin{aligned} &\max_{w(\cdot),e} \mathbb{E}\left[x - w\left(x\right)|e\right] \\ &\text{s.t. } \mathbb{E}\left[u\left(w\left(x\right)\right)|e\right] - \psi\left(e\right) \geq \underline{U} \qquad [IR] \\ &\quad e \in \arg\max_{\tilde{e}\in\mathcal{E}} \mathbb{E}\left[u\left(w\left(x\right)\right)|\tilde{e}\right] - \psi\left(\tilde{e}\right) \qquad [IC] \end{aligned}$$

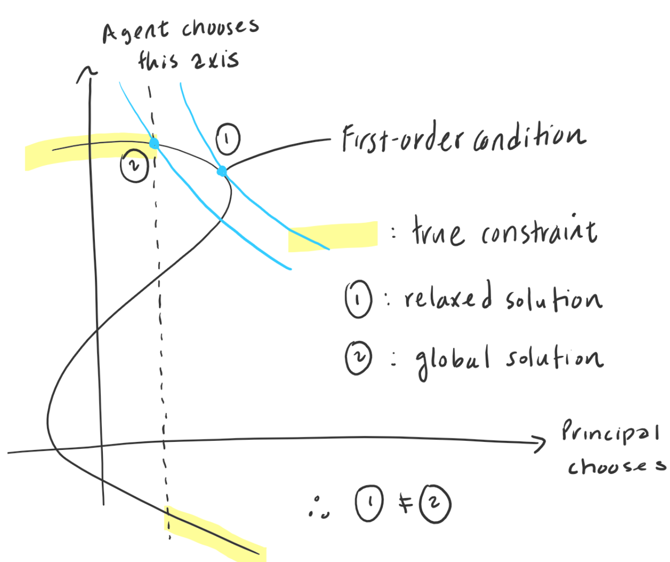
Or writing out the expectation explicitly:

$$\begin{aligned} &\max_{w(\cdot),e} \int_X \left[x - w\left(x\right)\right] f\left(x|e\right) dx \\ &\text{s.t. } \int_X u\left(w\left(x\right)\right) f\left(x|e\right) dx - \psi\left(e\right) \geq \underline{U} \qquad [IR] \\ &\quad e \in \arg\max_{\tilde{e}\in\mathcal{E}} \int_X u\left(w\left(x\right)\right) f\left(x|\tilde{e}\right) - \psi\left(\tilde{e}\right) \qquad [IC] \end{aligned}$$

We can relax the problem by replacing $[IC]$:

$$\begin{aligned} &\max_{w(\cdot),e} \int_X \left[x - w\left(x\right)\right] f\left(x|e\right) dx \\ &\text{s.t. } \int_X u\left(w\left(x\right)\right) f\left(x|e\right) dx - \psi\left(e\right) \geq \underline{U} \qquad [IR] \\ &\quad \int_X u\left(w\left(x\right)\right) f_e\left(x|e\right) dx - \psi'\left(e\right) = 0 \qquad [IC - FOC] \end{aligned}$$

It is important to note that $[IC - FOC]$ is a weaker constraint. Thus, when you solve the relaxed problem, you need to check if it solves the global program as well. This point can be made through a figure:



5.2.2 Solving the Relaxed Problem

We can set up the Lagrangian

$$\mathcal{L} = \int_X (x - w(x)) f(x|e)\, dx + \lambda \left(\int_X u(w(x)) f(x|e)\, dx - \psi(e) \right) + \mu \left(\int_X u(w(e)) f_e(x|e)\, dx - \psi'(e) \right)$$

and differentiate with respect to w:

$$\frac{1}{u'(w(x))} = \lambda + \mu \left(\frac{f_e(x|e)}{f(x|e)} \right)$$

Previously, we had

$$\frac{1}{u'(w(x))} = \lambda + \mu \left(\frac{f_H(x) - f_L(x)}{f_H(x)} \right)$$

Similar observations as before can be made:

- ▷ If first-order approach is okay, and f satisfies MLRP, then optimal wage is increasing in output.
- ▷ Wage will not depend on a signal y if and only if x is sufficient for y with respect to e.

We can also show that $\mu > 0$ and $\lambda > 0$. This is in Jeweitt (1988)

- ▷ First, note that $\underline{\mu \neq 0}$ (since the wage cannot be constant given $e > 0$)
- ▷ Second, write:

  $$f_e(x|e) = \left(\frac{1}{u'(w(x))} - \lambda \right) \frac{f(x|e)}{\mu}$$

  Since f_e is a density,

  $$0 = \int f_e(x|e)\, dx = \int \left(\frac{1}{u'(w(x))} - \lambda \right) \frac{f(x|e)}{\mu} dx$$

  Since $1/u'(w)$ is positive, it must be the case that $\underline{\lambda > 0}$.
- ▷ Furthermore, examine the agent's first-order condition:

  $$\int_X u(w(e)) f_e(x|e)\, dx = \psi'(e)$$

  Plugging in the expression for $f_e(x|e)$:

  $$\begin{aligned} \int_X u(w(e)) \left(\frac{1}{u'(w(x))} - \lambda \right) \{f(x|e)\}\, dx &= \mu\psi'(e) \\ \Leftrightarrow \text{Cov}\left[u(w(x)), \frac{1}{u'(w(x))} \right] &= \mu\psi'(e) \end{aligned}$$

  since the expectation of $[1/u'(w(x))] = \lambda$ (since expectation of f_e/f is zero). The covariance has to be positive, and $\psi'(e)$ is also positive, so $\underline{\mu > 0}$.
- ▷ Note that this proof doesn't work if the principal's risk-averse. We also didn't use MLRP here.

In general, the IR constraint binds. The one place where it is no longer correct is when you have constraint on the wage i.e. suppose you can't pay a worker below a certain wage. To give an incentive to work for you, you can't lower the wage to extract rent, so you have slack and $\lambda = 0$.

5.2.3 Sufficient Conditions for Validity of First-order Approach

The problem is that the relaxed program may pick out non-optimal solutions. Therefore, we are interested in finding conditions that make the program concave, in which case the solution will be unique.

Approach 1. Convex Distribution Function Condition (CDFC) (Assuming F differentiable): Distribution F satisfies the convex distribution function if $F_{ee}(x|e) \geq 0$.

- ▷ A lot of people add this to their mix of assumptions.
- ▷ Example #1: Suppose $F(x|e) = e \cdot F_1(x) + (1-e) F_0(x)$ with $f_1(x)/f_0(x)$ increasing in x. (This will satisfy MLRP).
- ▷ Example #2: $G(x)^{h(e)}$ where $h''(e)$ is concave.

We can see this using integration by parts:

$$\begin{aligned}
&\int_{\underline{x}}^{\bar{x}} u(w(x)) f(x|e)\, dx - \psi(e) \\
=& u(w(x)) F(x|e) \,|_{\underline{x}}^{\bar{x}} - \int_{\underline{x}}^{\bar{x}} u(w(x)) w'(x) F(x|e)\, dx - \psi(e) \\
=& u(w(\bar{x})) - \int_{\underline{x}}^{\bar{x}} u'(w(x)) w'(x) F(x|e)\, dx - \psi(e)
\end{aligned}$$

is strictly concave (note that we are also using $w'(x) > 0$ which we obtained from MLRP plus the continuity of f_e/f).

Approach 2. Restrictions on the Utility Functions Jewitt (ECTA 1988) proposes a weakening of the CDFC condition by placing restrictions on the agent's utility functions, but these conditions are restrictive in different ways.

5.3 Limited Liability and Optimality of Debt Contracts

In some circumstances, it makes sense to model the agency problem as one where we have limited liability. One critical difference is that we can no longer use the previous argument that IR constraint must also bind. We will illustrate this using Innes (1990), which applies the ideas of optimal incentive design to demonstrate circumstances in which a debt contract is the optimal security for an entrepreneur to use to raise funds and to simultaenously induce high-powered incentives for effort.

5.3.1 Setup

There's an entrepreneur who wants to solve

$$\max_{e, r(\cdot)} \int (x - r(x)) f(x|e)\, dx - \psi(e)$$

where $r(x)$ represents the capital-market repayment schedule (how much you pay as a function of how much money you make) and $\psi(e)$ is the loss function, subject to

$$\int_0^{\bar{x}} r(x) f(x|e)\, dx \geq I \qquad [IR]$$

where interest rate is zero and I is the initial outlay, and the enterpreneur chooses e in away that satisfies:

$$e \in \arg\max_{\tilde{e}} \int (x - r(x)) f(x|\tilde{e})\, dx - \psi(\tilde{e}) \qquad [IC]$$

and limited liability:

$$0 \leq r(x) \leq x \qquad [LL]$$

▷ The market will try to infer how hard she will work based on the effort level she chooses.

▷ Limited liability can be interpreted as the person being infinitely risk-averse at a certain wage. If we are told, for example, that $w \geq 0$ in an optimal contract, another way to think about this is to think of an agent with the following utility:

$$u(w) = \begin{cases} w & w \geq 0 \\ -\infty & w < 0 \end{cases}$$

which implies that the person's infinitely risk averse.

5.3.2 Assumptions

1. There's a solution that's interesting to the enterpreneuer i.e. the value function is non-negative (an additional IR).
2. The effort level will be less than first-best, where first-best corresponds to the case when the entrepreneur maximizing with her own money $r(x) = 0$.
3. The following problem is strictly concave:

$$\int_{z}^{\bar{x}} x f(x|e)\, dx - \psi(e)$$

This assumption allows us to proceed with the first-order approach.

5.3.3 Relaxed Problem

Construct the relaxed problem:

$$\max_{e, r(\cdot)} \int (x - r(x)) f(x|e)\, dx - \psi(e)$$

$$\begin{aligned} [IR] &: \int_0^{\bar{x}} r(x) f(x|e)\, dx = I \\ [IC - FOC] &: \int (x - r(x)) f_e(x|e)\, dx - \psi'(e) \geq 0 \\ [LL] &: 0 \leq r(x) \leq x \end{aligned}$$

Writing out the Lagrangian ignoring the $[LL]$ constraint:

$$\begin{aligned} \mathcal{L} &= \int (x - r(x)) f(x|e)\, dx - \psi(e) + \lambda \left\{ \int_0^{\bar{x}} r(x) f(x|e)\, dx - 1 \right\} + \mu \left\{ \int (x - r(x)) f_e(x|e)\, dx - \psi'(e) \right\} \\ &= \int_0^{\bar{x}} r(x) \left\{ \lambda - \mu \frac{f_e(x|e)}{f(x|e)} - 1 \right\} f(x|e)\, dx + (\text{ stuff that doesn't depend on } r(x)) \end{aligned}$$

▷ Note that $r(x)$ appears linearly. This implies that depending on the sign of

$$\lambda - \mu \frac{f_e(x|e)}{f(x|e)} - 1$$

then $r(x) = 0$ or $r(x) = \bar{x}$.

▷ If $\mu = 0$, we get a first-best solution, so assume $\mu > 0$.

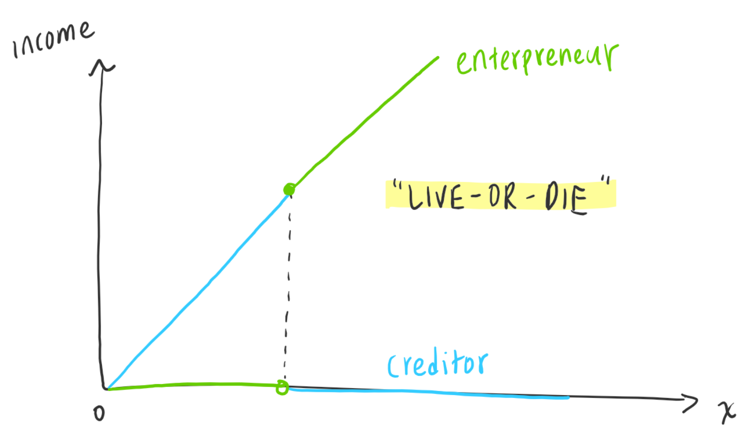
▷ We know that f_e/f is increasing in x (from MLRP), which implies that the above value is decreasing in x. This implies the existence of a threshold value $\hat{x}$ that yields:

$$r(x) = \begin{cases} x & \text{if } x \leq \hat{x} \\ 0 & \text{if } x > \hat{x} \end{cases}$$

where

$$\hat{x} = \lambda - 1 - \mu \frac{f_e(\hat{x}|e)}{f(\hat{x}|e)} = 0$$

Graphically:



▷ This is NOT a debt contract – It repays everything for $x \leq \hat{x}$ but for $x > \hat{x}$, it pays absolutely nothing to the investors.

▷ This is optimal given MLRP because it concentrates the residual income in the right tail where MLRP gives stronger incentives.

5.3.4 Optimal Monotonic Repayment Contracts

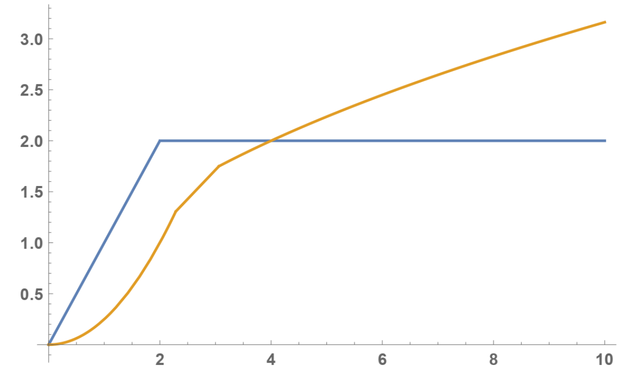
The previous contract we derived is a non-linear one. Now we impose the requirement that $r(x)$ is non-decreasing. Innes (1990) argues that the optimal monotonic contract must be debt.

▷ To see this, suppose $(r^*(\cdot), e^*)$ is an optimal repayment solution that implements e^* in equilibrium, and conisder a debt contract that repays investors I under the assumption that effort is unchanged. Specifically, choose the face value of debt D such that

$$\int_0^{\bar{x}} r^*(x) f(x|e^*)\, dx = \int_0^{\bar{x}} r^D(x) f(x|e^*)\, dx$$

given an effort. Then we can show that this is superior to r^*, thereby yielding a contradiction to its optimality:

1. Given the choice of D as above, $e^D > e^*$ i.e. effort on debt contract is always greater.



   * Because r^* is increasing, there is a point where to its right $r^*(x) > r^D(x)$ and to its left, $r^D(x) < r^*(x)$.
   * This says that the average value of the left gap is equal to the average value of the right gap.
   * As you're switching it from r^* to r^D, you're shifting the *repayments* to the left, which means *residual income* is shifted to the right and MLRP guarantees that enterpreneur's efforts will increase.
2. Under a debt contract, integration by parts reveals that investors are better off with the debt contract.

▷ Thus, we conclude that the optimal repayment schedule for entrepreneurial incentives, given the monotonicity constraint, is debt.

5.4 Holmstrom-Milgrom (ECTA 1987)

The optimal solution in the basic principal-agent model is arguably complex because there is an imbalance between the agent's one-dimensional effort space and the infinitedimensional control space of the principal. The goal of Holmstrom and Milgrom (1987) is to construct a dynamic model of agency that allows the agent to respond to the history of production over an interval of time in such a way that the principal's optimal contract becomes linear in some sense.

5.4.1 Setup

We are going to have $t = 1, ..., T$ periods, and in each period the output can be $x_t \in \{x_0, ..., x_n\}$. The agent chooses $\phi \in \Phi \subseteq \Delta^n$ i.e. ϕ is a probability distribution over $n+1$ things, and the associated cost of choosing ϕ_t is $c(\phi_t)$. The principal is risk-netural, and the agent is risk-averse:

$$u(y) = -e^{-ry}$$

where r is the risk aversion problem. Utility is increasing and concave, starting from a negative number. Furthermore, the cost of effort is monetary (as opposed to having cost of effort separate from the utility of wage).

▷ The principal will pay wage based on the observed series of outputs

$$W(x_1, ..., x_T)$$

but the agents' utility is determined by the final wage:

$$-e^{-r\left(w\left(X^T\right)-\sum_{t=1}^T c(\phi_t)\right)} \text{ where } X^t = (x_1, ..., x_t)$$

- ▷ CARA utility has a feature that the risk preferences stay the same no matter the amount of wealth.
  - * We need this because suppose the agent had a series of good results early on. We don't want the agent to exert less effort going forward.
  - * Authors argue that this is more applicable to monthly contracts and inappropriate for contracts 5- or 10- years going forward.
- ▷ Also note that marginal utility of income relative to cost of effort is constant.
- ▷ The above two conditions ensures that the problem is stationary, i.e. independent of wealth. This is a **crucial element** in the model.

5.4.2 The One-Period Problem

The principal solves:

$$\max_{\{w_0,\ldots,w_n\},\phi} \sum_{i=0}^{n} \phi_i \left(x_i - w_i\right)$$

$$\text{subject to } [IC] : \phi \in \arg\max_{\phi'} \sum_{i=1}^{n} \phi_i u\left(w_i - c\left(\phi'\right)\right)$$

$$[IR] : \sum_{i=0}^{n} \phi_i u\left(w_i - c\left(\phi\right)\right) \geq u\left(\underline{w}\right)$$

where $\underline{w}$ is the alternative reservation wage. We obtain the following results:

1. (Theorem 3) Under some reasonable assumptions on the cost function, you can ignore the objective function i.e. if you choose a ϕ, it pins down $\{w_0, \ldots, w_n\}$ entirely through $[IC]$.
   - ▷ In other words, for any $\phi \in \Phi$, other ϕ is not implementable or $\exists! w^* = (w_0^*, \ldots, w_n^*)$ that satisfies IC and IR with equality.
2. Suppose (w^*, ϕ^*) solves one-period program with certainty equivalent $\underline{w}$. Then if $CE = \hat{w}$, then $(w^* - \underline{w} + \hat{w}, \phi^*)$ is optimal.
   - ▷ If you choose $\hat{w}$, ϕ does not change and only the wage schedule gets bumped up.

5.4.3 The T-Period Problem

We have the same problem repeat every day for year. We ask two questions: what is the cheapest way to implement a path, and what is the optimal path?

- ▷ Define path to be $\left\{\phi_t\left(X^{t-1}\right)\right\}$ – I will look at the history and choose ϕ_t.
- ▷ Define $U_t\left(X^t\right)$, the continuation value of the agent to be

$$U_t\left(X^t\right) \equiv \mathbb{E}\left[u\left(w\left(\tau\right) - \sum_{\tau=t+1}^{T} c\left(\phi_\tau\right)\right) | X^t\right]$$

  and define the certainty equivalent to be

$$w_t\left(X^t\right) = CE\left(U_t\left(X_t\right)\right)$$

▷ Thus the agent must choose $\phi_t\left(X^{t-1}\right)$ that satisfies:

$$\phi_t\left(X^{t-1}\right) \in \arg\max_{\phi'} \sum_i \phi_i u\left(\underline{w}_t\left(X^{t-1}, x_{it}\right) - c\left(\phi_t\right)\right)$$

Using the result 1 from the one-period problem, we know that choosing ϕ_t pins down the wages. Denote the difference between the new set of wages and the old set to be $\tilde{w}_t$:

$$\tilde{w}_t\left(x_{it}|\phi_t\left(X^{t-1}\right)\right) \equiv \underline{w}_t\left(X^{t-1}, x_{it}\right) - \underline{w}_{t-1}\left(X^{t-1}\right)$$

which yields the optimal wage that the principal pays:

$$w\left(X^T\right) = \sum_{t=1}^{T} \tilde{w}_t\left(x_{it}|\phi_t\left(X^{t-1}\right)\right) + \underline{w}_0$$

▷ Basically, we are taking as given $\left\{\phi_t\left(X^{t-1}\right)\right\}$. This uniquely pins down the wages sequence, so just trivially addup the differences $(\tilde{w})$ to get $w\left(X^T\right)$.

▷ Remember that we are simply computing the cost of implementing $\left\{\phi_t\left(X^{t-1}\right)\right\}$.

What about optimality? We have the following theorem:

▷ Optimal $\left\{\phi_t\left(X^{t-1}\right)\right\}_{t=1..T}^{X^{t-1}\in H^{t-1}} = \left(\phi_1^*, ..., \phi_T^*\right)$ i.e. you repeat the one-period contract every period.

Another way to write the optimal wage is:

$$w\left(X^T\right) = w^* \cdot A^T$$

where

$$\begin{aligned} w^* &= \left(w_0^*, ..., w_n^*\right) \\ A^T &= A_0^T, A_1^T, ..., A_n^T \end{aligned}$$

where A_j^T is the number of times x_j happened during the year. w^* is the price you're paying for each of the outcome. Further denote w_0^* as the worst outcome and define

$$\begin{aligned} \alpha_i &= w_i^* - w_0^* \\ \beta_i &= T \cdot w_0^* \end{aligned}$$

and

$$\underline{w\left(X^T\right) = \sum_{i=1}^{n} \alpha_i A_i^T + \beta}$$

▷ Suppose you had only two outcomes: 0 and 1. Then you just have a binomial distribution and if you pack in a lot, you get a normal distribution in some well-defined limit:

$$w\left(x\right) = \alpha x_T + \beta, \quad x \sim N\left(0, \sigma^2\right)$$

i.e. you get a linear result.

▷ Note that the limit we're taking is we're increasing the number of intervals but keeping the actual length of the interval.

5.4.4 Application: Holstrom & Milgrom (JLEO, 1991)

Now the setup is

$$x = \mu + \epsilon, \epsilon \sim N\left(0, \sigma^2\right)$$

where μ is the agent's action choice and an associated cost function $c(\mu) = \frac{k}{2}\mu^2$ and $w(x) = \alpha x + \beta$. The agent has exponential utility (CARA utility) with parameter r and the principal is risk-neutral.

- ▷ **Full Information**: where e is contractable, the principal would offer the agent with a fixed wage equal to

$$w = \underline{w} + c(\mu)$$

and maximize $\mathbb{E}[x|\mu] - \underline{w} - c(\mu)$ which is equivalent to maximizing $\mu - c(\mu)$. Assuming $\underline{w} = 0$ for simplicity, this means that the first-best contract is

$$\mu^{FB} = \frac{1}{k}$$

which implies

$$w^{FB} = \frac{1}{2k}, \qquad \pi = \frac{1}{2k}$$

- ▷ **Imperfect Information:** Holmstrom and Milgrom's result tells us that we can restrict attention to linear contracts of the form

$$w(x) = ax + \beta$$

  * The agent's certainty equivalent upon choosing action μ is computed by

$$\begin{aligned}\mathbb{E}\left[-e^{-r(w(x)-C(\mu))}\right] &= \mathbb{E}\left[-e^{-r\left(a(\mu+\epsilon)+\beta-\frac{k}{2}\mu^2\right)}\right] \\ &= -e^{-r\left(a\mu+\beta-\frac{k}{2}\mu^2\right)}\mathbb{E}\left[-e^{-ar\epsilon}\right] \\ &= -e^{-r\left(a\mu+\beta-\frac{k}{2}\mu^2\right)+\frac{1}{2}\alpha^2r^2\sigma^2}\end{aligned}$$

  and thus the certainty equivalent is

$$\alpha\mu + \beta - \frac{k}{2}\mu^2 - \frac{r}{2}\alpha^2\sigma^2$$

  * Since the agent's CE function is globally concave in μ, the FOC is both necessary and sufficient, thereby yielding:

$$\mu^* = \frac{\alpha}{k}$$

  This implies that the utilities-possibility frontier is independent of β. Independence of β is an artifact of CARA utility, whereas linearity is due to the combination of CARA utility and normally distributed errors. Thus, β can be thought of as a free parameter that allows us to satisfy the agent's $[IR]$ constraints.

  * Thus, we can now reduce the principal's original problem

$$\begin{aligned}\max \quad & \mathbb{E}[x - \alpha x - \beta|\mu] \\ \text{s,t, } & \mu \in \arg\max \mathbb{E}[u(\alpha x + \beta - c(\mu))] \\ & \mathbb{E}[u(\alpha x + \beta - c(\mu))] \geq u(\underline{w})\end{aligned}$$

  by assuming that the $[IR]$ binds to rewrite $\alpha\mu + \beta$ as the following:

$$\begin{aligned}\mathbb{E}[u(w - c(\mu))] &= u(\underline{w}) \\ \Rightarrow \alpha\mu + \beta - \frac{k}{2}\mu^2 - \frac{r}{2}\alpha^2\sigma^2 &= \underline{w} \\ \Rightarrow \alpha\mu + \beta &= \underline{w} + \frac{k}{2}\mu^2 + \frac{r}{2}\alpha^2\sigma^2\end{aligned}$$

Plugging this into the original problem and replacing $[IC]$ with $\alpha = \mu k$ yields:

$$\max_{\alpha,\mu} \quad -\bar{w} + \mu - \frac{k}{2}\mu^2 - \frac{r}{2}\alpha^2\sigma^2$$
$$\text{s.t. } \alpha = \mu k$$

Solving this model yields

$$\alpha^* = \frac{1}{1 + rk\sigma^2}$$

and

$$\mu^* = \frac{1}{k\left(1 + rk\sigma^2\right)} = \alpha^*\mu^{FB} < \mu^{FB}$$
$$\pi^* = \frac{1}{(2k)\left(1 + rk\sigma^2\right)} = \alpha^*\pi^{FB} < \pi^{FB}$$

* As either r, k, or σ^2 decreases, the power of the optimal incentive scheme increases (i.e. α^* increases). In other words, we move towards the first best when risk aversion, the uncertainty in measuring effort, or the curvature of the agent's effort function decreases.

5.4.5 Role of Additional Information

Suppose the principal observes an additional signal (y) which is correlated with ϵ that satisfies

$$\mathbb{E}\left[y\right] = 0, \text{Var}\left[y\right] = \sigma_y^2, \text{Cov}\left[\epsilon, y\right] = \rho\sigma_y\sigma_\epsilon$$

Letting the optimal wage contract be linear in both aggregates, we have

$$w\left(x, y\right) = \alpha_1 x + \alpha_2 y + \beta$$

where

$$\alpha_1^* = \frac{1}{1 + rk\sigma_\epsilon^2\left(1 - \rho^2\right)}, \quad \alpha_2^* = -\alpha_1\frac{\sigma_\epsilon}{\sigma_y}\rho$$

Compared to the previous result, it is as if the outside signal reduces the variance on ϵ from σ_ϵ^2 to $\sigma_\epsilon^2\left(1 - \rho^2\right)$.

5.4.6 Multi-Task Incentive Contracts

Now suppose that the principal can contract on following k vector of aggregates: $x = \mu + \epsilon, \quad \epsilon \sim \mathcal{N}\left(0, \Sigma\right)$. The agent chooses a vector of efforts, μ, at a cost of $C\left(\mu\right)$, and the agent's utility is exponential with CARA parameter of r. As before, CARA utility and normal errors implies that the optimal contract solves:

$$\max_{a,\mu} B\left(\mu\right) - C\left(\mu\right) - \frac{r}{2}\alpha'\Sigma\alpha$$
$$\text{s.t. } \mu \in \arg\max_{\tilde{\mu}} \alpha'\tilde{\mu} - C\left(\tilde{\mu}\right)$$

where once agin β is determined so as to meet the agent's IR constratin, given the optimal (α, μ). We can thus show that the optimal contract satisfies:

$$\alpha^* = \left(I + r\left[C_{ij}\left(\mu^*\right)\right]\Sigma\right)^{-1} B'\left(\mu^*\right)$$

Now suppose that there are two tasks but the effort of only the first task can be measured: $\sigma_2 = \infty$ and $\sigma_{12} = 0$. Furthermore, suppose that under the optimal contract $\mu^* > 0$ i.e. both tasks will be provided at the optimum.

▷ It turns out that the optimal contract satisfies $\alpha_2^* = 0$ and

$$\alpha_1^* = \left(B_1 \left(\mu^* \right) - B_2 \left(\mu^* \right) \frac{C_{12} \left(\mu^* \right)}{C_{22} \left(\mu^* \right)} \right) \left(1 + r\sigma_1^2 \left(C_{11} \left(\mu^* \right) - \frac{C_{12} \left(\mu^* \right)^2}{C_{22} \left(\mu^* \right)} \right) \right)^{-1}$$

▷ If effort levels across tasks are complements (i.e. $C_{12} < 0$), then the more complementary, the higher the α_1^*.

▷ This has the flavor of the public finance results that when the government can only tax a subset of goods, it should tax them more or less depending upon whether the taxable goods are substitutes or complements with the un-taxable goods. See, for example, Atkinson and Stiglitz [1980 Ch. 12] for a discussion concerning taxation on consumption goods when leisure is not directly taxable.

It is also possible that the optimal contract may have $\alpha_1^* = 0$.

▷ If the technologies are symmetric, i.e. $C(\mu) = c(\mu_1 + \mu_2)$ and $B(\mu_1, \mu_2) \equiv B(\mu_2, \mu_1)$, then $\alpha_1^* = \alpha_2^* = 0$ since no incentives are provided.

6 Monopolistic Screening

We explore the model of optimal (monopolistic) screening contracts. Because we assume a monopoly principal rather than a market of competitive firms offering screening contracts, we avoid many of the issues in competitive screening games (e.g., a unique equilibrium contract will be generally exist in the monopolistic screening environment).

We also continue with our assumption of a single agent. We will later apply and extend the techniques from monopolistic screening to multi-agent settings (e.g., public goods games, bilateral trading mechanisms and auctions). When we do so, it is common to refer to the program as optimal mechanism design (rather than optimal screening contracts).

Preview of Steps for Solving a Screening Problem Here are the following steps:

1. Rewrite the implementability constraint.
2. Simplify $\mathbb{E}\left[U\left(\theta\right)\right]$ using integration by parts to obtain the following equation::
$$\mathbb{E}\left[U\left(\theta\right)\right]=U\left(\underline{\theta}\right)+\mathbb{E}\left[U_{\theta}\left(q\left(\theta\right),\theta\right)\frac{1-F\left(\theta\right)}{f\left(\theta\right)}\right]$$
3. Incorporate $\mathbb{E}\left[U\left(\theta\right)\right]$ into principal's objective to maximize pointwise for $q\left(\cdot\right)$ assuming regularity.
4. Recover $U\left(\cdot\right),t\left(\cdot\right)$ from $q\left(\cdot\right)$ using $t\left(\theta\right)=u\left(q\left(\theta\right),\theta\right)-U\left(\theta\right)$
5. Find $\underline{P}\left(q\right)$ that implements $\left\{q\left(\cdot\right),U\left(\cdot\right)\right\}$.
   - ▷ This is often called as the "taxation principle" – reverse of the revelation principle – which says:
   $$\underline{P}\left(q\right)=\begin{cases}t\left(\theta\right) & \text{if } q=q\left(\theta\right)\text{ for some }\theta\\ \infty & \text{otherwise}\end{cases}$$
   i.e. when you're done, you can go from the direct mechanism to something that is more tangible.
   - ▷ Thus,
   $$\bar{\theta}\left(q\right)=\sup_{\theta}\left\{\theta|q\left(\theta\right)=q\right\}$$
   $$\bar{P}\left(q\right)=t\left(\bar{\theta}\left(q\right)\right)\text{ if }\bar{\theta}\left(q\right)\neq\emptyset$$

6.1 Monopoly Non-Linear Pricing

A single firm wishes to sell its product to a single consumer who has unknown preferences for the seller's good. The buyer knows his preference for the good, but the seller only knows the probabilitiy distribution for the buyer's type.

6.1.1 Setup

Both buyer and seller are risk-neutral.

- ▷ Buyer consuming q units at total price t (transfer) yields utility $u\left(q,\theta\right)-t$ where
  - * u is increasing and concave in $q\in\left[\underline{q},\bar{q}\right]$
  - * u is increasing in θ and u_{θ} is bounded on Θ
  - * single-crossing property holds: $u_{q\theta}\left(q,\theta\right)>0,\forall\theta\in\Theta,q>0$

▷ Seller: Makes profit of $V = t - C(q)$ where C is increasing and convex

The conditions on the buyer demand imply that the demand curve is given as $p = u_q(q,\theta)$. Furthermore, if $\theta \to \theta'$, the demand curve shifts upwards since $u_{q\theta} > 0$, which implies that we have an ordering of different demand curves in our economy.

▷ For example, we can consider unit demand $u(q,\theta) = \theta q, \qquad q \in \{0,1\}$.

Furthermore, we define the first-best q for each θ as:

$$q^{FB}(\theta) = \max_{q \in Q} u(q,\theta) - C(q)$$

The timing is as follows:

1. Nature chooses $\theta \in \Theta$ according to $F(\theta)$. Buyer sees θ.
2. Seller offers Buyer a "mechanism" (Γ).
3. Buyer rejects or accepts

Remark 6.1. What if we put (2) and (3) before (1)?

▷ Having 2 before 1 implies the first-best outcome.

▷ To see this, compute the consumer's wilingness to pay who does not know θ:

$$\pi^{FB} = \mathbb{E}_\theta\left[\max_q u(q,\theta) - C(q)\right]$$

Then the seller can a contract telling the buyer the option to consume q at a cost $C(q)$.

▷ Intuitively, this makes sense since both the Seller and Buyer are unware of θ.

6.1.2 Mechanisms

What mechanism can the seller choose? We will define mechanism as an extensive-form game Γ in which the seller's strategy is pre-committed. In other words, all the seller's moves are locked down at his decision nodes (does not need to be sequentially rational). Here are some examples:

1. Seller posts a price p and buyer then can choose to purchase the good or not, i.e. $q \in \{0,1\}$.
2. Seller posts $P(q) = q, q \in Q$, and buyer then can choose the quantity.
3. Seller offers a menu of lotteries $P(\phi)$ where $\phi(q,t) \in Q \times \mathbb{R}$ and buyer chooses.
4. Buyer does cheap talk; Depending on the cheap talk, seller offers a mechanism; Buyer makes a decision. (Extended versions of this form would also make sense.)

A **direct mechanism** $\tilde{\Gamma}$ consists of the following form:

▷ Buyer reports his type $\hat{\theta} \in \Theta$,

▷ Distribution of (q,t) is generated by $\hat{\theta}$ – $\{\phi(q,t|\theta)\}_{\theta \in \Theta}$ – which the agent can plug into his expected utility.

A **deterministic direct mechanism** is a deterministic version of ϕ_θ: $\left\{q\left(\hat{\theta}\right), t\left(\hat{\theta}\right)\right\}_{\hat{\theta}\in\Theta}$.

Example 6.1. (Unit Demand) Suppose $u(q,\theta) = \theta q, \quad q \in \{0,1\}$ and we have

$$q\left(\hat{\theta}\right) = \begin{cases} 1 & \text{if } \hat{\theta} \geq p \\ 0 & \text{if } \hat{\theta} < p \end{cases}, \qquad t\left(\hat{\theta}\right) = \begin{cases} p & \text{if } \hat{\theta} \geq p \\ 0 & \text{if } \hat{\theta} < p \end{cases}$$

Note that this is equivalent to a seller posting a price p. Note that the Buyer has a weak incentive to report the truth.

6.1.3 Revelation Principle (Deterministic)

Proposition 6.1. *The statement is as follows: For any Γ and deterministic equilibrium* outcome, $\{q^*(\theta), t^*(\theta)\}_{\theta\in\Theta}$, *then there exists a direct mechanism $\tilde{\Gamma}$ such that*

1. *Buyer reports his type truthfully: $\tilde{\sigma}_B^* = \theta$*
2. *You achieve the same outcome: $\tilde{q}^*(\theta) = q^*(\theta), \tilde{t}^*(\theta) = t^*(\theta)$*

Proof. We will construct the direct mechanism by taking the deterministic outcome from the game and give you whatever outcome you would have gotten if you are truly what you report:

- ▷ If $\hat{\theta} = \theta$, then we're done.
- ▷ If $\hat{\theta} \neq \theta$, then he earns $u\left(q^*\left(\hat{\theta}\right), \theta\right) - t^*\left(\hat{\theta}\right)$ which is less than $u(q^*(\theta), \theta) - t^*(\theta)$ since this is not the optimal strategy in the original game.

■

Note that an equivalent result holds for $\{\phi^*(y|\theta), t^*(\theta)\}_{\theta\in\Theta}$ i.e. a non-deterministic equilibrium allocations.

6.1.4 Optimal Deterministic Mechanism

Consider the following seller's program. Note that using the revelation principle, the seller can restrict her attention to the space of deterministic direct mechanisms in which truth-telling is induced in equilibrium:

$$\max_{\{q(\cdot),t(\cdot)\}} \mathbb{E}_\theta\left[t(\theta) - C(q(\theta))\right]$$
$$\text{such that } \theta \in \arg\max_{\hat{\theta}\in\Theta} u\left(q\left(\hat{\theta}\right), \theta\right) - t\left(\hat{\theta}\right) \quad \text{[truth-telling is optimal]}$$
$$u(q(\theta), \theta) - t(\theta) \geq 0, \forall t \quad \text{[IR]}$$

- ▷ If we have an IR constraint with a constant outside option, we can just embed it into the utility function. If the outside option depends on θ, the assumption $u_\theta > 0$ may prevent us from simply rolling it into the utility function.

Definitions:

▷ The **indirect utility** from truth telling is defined as:

$$U(\theta) \equiv u(q(\theta), \theta) - t(\theta)$$

▷ Direct mechanism $\{q(\cdot), t(\cdot)\}$ is **incentive-compatible** $[IC]$ if and only if for all $\theta \in \Theta$,

$$U(\theta) \geq U\left(\hat{\theta}|\theta\right) = u\left(q\left(\hat{\theta}\right), \theta\right) - t\left(\hat{\theta}\right), \forall \theta, \hat{\theta}$$

▷ Direct mechanism $\{q(\cdot), t(\cdot)\}$ is **implementable** if for $t(\theta) = u(q(\theta), \theta) - U(\theta)$, $\{q(\cdot), t(\cdot)\}$ is incentive-compatible.

Note that incentive compatibility and implementability constraints are challenging to work with in their current form, so we seek an alternative representation. Conveniently enough, if we have the following results:

Proposition 6.2. *A DRM $\{q(\cdot), t(\cdot)\}$ is incentive-compatible and $\{q(\cdot), U(\cdot)\}$ is implementable if and only if*

1. *$q(\cdot)$ is non-decreasing;*
2. *We can write:*

$$U(\theta) = U(\underline{\theta}) + \int_{\underline{\theta}}^{\bar{\theta}} u_\theta(q(s), s)\, ds$$

Proof. First, we show that $[IC]$ implies (1) and (2).

▷ $[IC]$ implies that $U(\theta) \geq U\left(q\left(\hat{\theta}\right), \theta\right) - t\left(\hat{\theta}\right), \forall \hat{\theta}, \theta$. We can rewrite this as:

$$U(\theta) \geq U\left(\hat{\theta}\right) + u\left(q\left(\hat{\theta}\right), \theta\right) - u\left(q\left(\hat{\theta}\right), \hat{\theta}\right)$$

Rearranging:

$$U(\theta) - U\left(\hat{\theta}\right) \geq u\left(q\left(\hat{\theta}\right), \theta\right) - u\left(q\left(\hat{\theta}\right), \hat{\theta}\right)$$

▷ Repeating the same process symmetrically yields:

$$u(q(\theta), \theta) - u\left(q(\theta), \hat{\theta}\right) \geq U(\theta) - U\left(\hat{\theta}\right)$$

▷ Combining yields:

$$u(q(\theta), \theta) - u\left(q(\theta), \hat{\theta}\right) \geq u\left(q\left(\hat{\theta}\right), \theta\right) - u\left(q\left(\hat{\theta}\right), \hat{\theta}\right)$$

Since $u_{q\theta} > 0$, it must be that $q(\theta)$ and θ are moving together (and thus (1) is true). Furthermore, dividing each side by $\theta - \hat{\theta}$ and taking the limit to as $\hat{\theta} \to \theta$:

$$\lim_{\theta \to \hat{\theta}} \frac{u(q(\theta), \theta) - u\left(q(\theta), \hat{\theta}\right)}{\theta - \hat{\theta}} \geq \lim_{\theta \to \hat{\theta}} \frac{u\left(q\left(\hat{\theta}\right), \theta\right) - u\left(q\left(\hat{\theta}\right), \hat{\theta}\right)}{\theta - \hat{\theta}}$$

we see that left and right derivatives exist, and that they are equal at all points where $q(\cdot)$ is continuous.

- ▷ Therefore, assuming continuity of q, u is twice differentiable and the left and right derivatives coincide, we have:

$$U'(\theta) = u_\theta(q(\theta), \theta)$$

- ▷ Given that u_θ is Lipschitz-continuous, $U(\theta)$ will also be Lipschitz-continuous, which also implies absolute continuity. Thus, the fundamental theorem of calculus is valid and we have

$$U(\theta) = U(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} u_\theta(q(s), s)\, ds$$

- ▷ *Interpretation*: If you find me a q that is monotonic, and U satisfies the formula, I can find you a transfer function. If it violates any of these properties, you cannot find such transfer function:

$$t(\theta) = u(q(\theta), \theta) - U(\underline{\theta}) - \int_{\underline{\theta}}^{\theta} u_\theta(q(s), s)\, ds$$

We now know **sufficiency**.

- ▷ To get to $[IC]$, note that we need:

$$\begin{aligned} U(\theta) &\geq u\left(q\left(\hat{\theta}\right), \theta\right) - t\left(\hat{\theta}\right) \quad \forall \theta, \theta' \\ &= U\left(\hat{\theta}\right) + u\left(q\left(\hat{\theta}\right), \theta\right) - u\left(q\left(\hat{\theta}\right), \hat{\theta}\right) \\ \Rightarrow U(\theta) - U\left(\hat{\theta}\right) &= u\left(q\left(\hat{\theta}\right), \theta\right) - u\left(q\left(\hat{\theta}\right), \hat{\theta}\right) \end{aligned}$$

- ▷ Using the (2):

$$\int_{\hat{\theta}}^{\theta} u_\theta(q(s), s)\, ds \geq \int_{\hat{\theta}}^{\theta} u_\theta\left(q\left(\hat{\theta}\right), s\right) ds \quad \forall \theta, \hat{\theta}$$

- ▷ Suppose $\theta > \hat{\theta}$. (1) implies that $q(s) \geq q\left(\hat{\theta}\right), \forall s \in \left[\hat{\theta}, \theta\right]$. Then $u_{q\theta} > 0$ implies that $u_\theta(q(s), s) \geq u_\theta\left(q\left(\hat{\theta}\right), s\right)$, thereby satisfying the inequality.
- ▷ A symmetric argument holds for $\theta < \hat{\theta}$.

Note that in the proof, we used $u_{q\theta} > 0$ both in the necessity and the sufficiency, but $u_\theta \geq 0$ never came up. ■

In settings such as auctions, we will assume $u(q, \theta) = \theta q + \beta$. Then (1) $U(\theta)$ is a convex function, and (2) $q(\theta) \in \partial U(\theta)$ i.e. q is in the subdifferential of $U(\theta)$ (i.e. it is one of the supporting gradients). For linear utility functions specifically, we can use these two conditions to substitute the conditions in the lemma (i.e. these conditions imply the conditions in the lemma.)

6.1.5 Revenue Equivalence

Note that the seller's revenue, $t(\theta)$, is defined by the allocation $q(\cdot)$ and the utility of the lowest-type agent $U(\underline{\theta})$:

$$t(\theta) = u(q(\theta), \theta) - U(\theta)$$

Using the previous result, we have that for any direct mechanism, $\{q(\theta), t(\theta)\}_{\theta \in \Theta}$ that is incentive compatible yields

$$t(\theta) = u(q(\theta), \theta) - U(\underline{\theta}) + \int_{\underline{\theta}}^{\bar{\theta}} u_\theta(q(s), s)\, ds$$

6.1.6 Simplified Seller's Program

Recall the original seller's program:

$$\begin{aligned}
&\max_{\{q(\cdot),t(\cdot)\}} \mathbb{E}_\theta\left[t(\theta) - C(q(\theta))\right] \\
&\text{such that } \theta \in \arg\max_{\hat{\theta}\in\Theta} u\left(q\left(\hat{\theta}\right),\theta\right) - t\left(\hat{\theta}\right) \quad \text{[truth-telling is optimal]} \\
&\qquad u(q(\theta),\theta) - t(\theta) \geq 0, \forall t \quad \text{[IR]}
\end{aligned}$$

Now we can simplify the program using the proposition:

$$\begin{aligned}
&\max_{\{q(\cdot),t(\cdot)\}} \mathbb{E}_\theta\left[u(q(\theta),\theta) - C(q(\theta)) - U(\theta)\right] \\
&\text{such that } [IC1] \quad q(\cdot) \text{ is non-decreasing} \\
&\qquad [IC2] \quad U(\theta) = U(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} u_\theta(q(s),s)\,ds, \forall\theta \\
&\qquad [IR] \qquad U(\theta) \geq 0, \forall\theta\in\Theta
\end{aligned}$$

If we assume $u_\theta \geq 0$, then you can replace $[IR]$ with $U(\underline{\theta}) \geq 0$. We will do that here. Note that we can now write $\mathbb{E}[U(\theta)]$ as

$$\begin{aligned}
\mathbb{E}[U(\theta)] &= U(\underline{\theta}) + \mathbb{E}\left[U(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} u_\theta(q(s),s)\,ds\right] \\
&= U(\underline{\theta}) + \int_{\underline{\theta}}^{\bar{\theta}}\int_{\underline{\theta}}^{\bar{\theta}} u_\theta(q(s),s)\,f(\theta)\,d\theta
\end{aligned}$$

Using integration by parts:

$$= U(\underline{\theta}) + \underbrace{\left[\left(\int_{\underline{\theta}}^{\bar{\theta}} u_\theta(q(s),s)\right)(F(\theta)+K)\right]_{\underline{\theta}}^{\bar{\theta}}}_{[A]} - \int_{\underline{\theta}}^{\bar{\theta}} (F(\theta)+K)\,u_\theta(q(\theta),\theta)\,d\theta$$

By choosing $K=-1$,we can make $[A]=0$ (since $F\left(\bar{\theta}\right)$ is equal to 1).[2] Doing so yields:

$$\mathbb{E}[U(\theta)] = U(\underline{\theta}) + \mathbb{E}\left[u_\theta(q(\theta),\theta)\,\frac{1-F(\theta)}{f(\theta)}\right]$$

Thus, we arrive at the following seller's program:

$$\begin{aligned}
&\max_{q(\cdot),t(\cdot)} \mathbb{E}_\theta\left[u(q(\theta),\theta) - C(q(\theta)) - U(\underline{\theta}) - u_\theta(q(\theta),\theta)\,\frac{1-F(\theta)}{f(\theta)}\right] \\
&\text{such that } [IC1] \quad q(\cdot) \text{ is non-decreasing} \\
&\qquad [IC2] \quad U(\underline{\theta}) \geq 0
\end{aligned}$$

2

- ▷ Once again, note that if we assume $u(\theta) \leq 0$, we will have to use $U\left(\bar{\theta}\right)$ (since this is the value we know) and set $K=0$.

The obvious choice is to set $U(\underline{\theta}) = 0$ which pins down our choice of $t(\cdot)$. Therefore, our **final simplified problem** becomes:

$$\max_{q(\cdot)} \quad \mathbb{E}_\theta \left[u(q(\theta), \theta) - u_\theta(q(s), s) \frac{1 - F(\theta)}{f(\theta)} - C(q(\theta)) \right]$$
$$\text{such that } [IC1] \quad q(\cdot) \text{ is non-decreasing}$$

- ▷ $u(q(\theta), \theta)$: utility of the buyer
- ▷ $u_\theta(q(s), s) \frac{1-F(\theta)}{f(\theta)}$: agent's (expected) information rent
- ▷ The two terms combined, denoted as $\tilde{u}(q, \theta)$, is denoted as "virtual utility of buyer."

6.1.7 Conditions for Pointwise Maximization

We can solve the following problem pointwise:

$$\max_{q(\cdot)} \quad \mathbb{E}_\theta \left[u(q(\theta), \theta) - u_\theta(q(\theta), \theta) \frac{1 - F(\theta)}{f(\theta)} - C(q(\theta)) \right]$$
$$\text{such that } [IC1] \quad q(\cdot) \text{ is non-decreasing}$$

Define the virtual surplus $\Lambda(q, \theta)$ as

$$\Lambda(q, \theta) \equiv u(q(\theta), \theta) - u_\theta(q(\theta), \theta) \frac{1 - F(\theta)}{f(\theta)} - C(q(\theta))$$

Since we are solving the problem pointwise, we want to ensure that the pointwise optimum over q defined by

$$\Lambda_q(q(\theta), \theta) = 0, \forall \theta \in \Theta$$

is (1) indeed a pointwise global maximum of Λ and the (2) corresponding $q(\cdot)$ is non-decreasing:

$$q'(\theta) = -\frac{\Lambda_{q\theta}}{\Lambda_{qq}} \geq 0$$

- ▷ Note that we got this expression for $q'(\theta)$ through taking the derivative of $\Lambda_q(q, \theta) = 0$ which yields:

$$\Lambda_{q\theta} + \Lambda_{qq} q'(\theta) = 0 \Rightarrow q'(\theta) = -\frac{\Lambda_{q\theta}}{\Lambda_{qq}} \geq 0 \Leftrightarrow \Lambda_{q\theta} \geq 0$$

We say that a screening problem defined by $\{u, C, F\}$ is **regular** if the function

$$\Lambda(q, \theta) \equiv u(q(\theta), \theta) - u_\theta(q(\theta), \theta) \frac{1 - F(\theta)}{f(\theta)} - C(q(\theta))$$

is twice continuously differentiable, strictly quasi-concave over $q \in Q$ and $\Lambda_{q\theta}(q, \theta) \geq 0$. One way to achieve this (regularity) is supposing u is quadratic and the inverse hazard rate

$$H(\theta) \equiv \frac{1 - F(\theta)}{f(\theta)}$$

is non-increasing in θ. This is referred to as the monotone-hazard rate condition (MHRC).

6.1.8 Solving the Simplified Problem

Now we set out to solve the following problem pointwise:

$$\max_{q(\cdot)} \quad \mathbb{E}_\theta \left[u\left(q\left(\theta\right),\theta\right) - u_\theta\left(q\left(\theta\right),\theta\right)\frac{1-F\left(\theta\right)}{f\left(\theta\right)} - C\left(q\left(\theta\right)\right) \right]$$
$$\text{such that } [IC1] \quad q\left(\cdot\right) \text{ is non-decreasing}$$

As before, define the virtual surplus $\Lambda\left(q,\theta\right)$ as

$$\Lambda\left(q,\theta\right) \equiv u\left(q\left(\theta\right),\theta\right) - u_\theta\left(q\left(\theta\right),\theta\right)\frac{1-F\left(\theta\right)}{f\left(\theta\right)} - C\left(q\left(\theta\right)\right)$$

while maintaining the solution be monotone.

▷ Taking the derivative with respect to q yields:

$$\left[u_q\left(q\left(\theta\right),\theta\right) - C'\left(q\left(\theta\right)\right)\right] f\left(\theta\right) = \left[1-F\left(\theta\right)\right] u_{\theta q}\left(q\left(\theta\right),\theta\right)$$

The equation can be interpreted as the effect of a marginal increase in q:

* LHS is the social benefit from a marginal increase in q for type θ, weighted by $f\left(\theta\right)$.
* RHS is the cost of such an increase in efficiency of type-θ's consumption. $u_{\theta q}$ is the marginal surplus increase to types above θ, which have mass $1-F\left(\theta\right)$.
* In short, the inefficiencies in consumption are evaluated against increased rents to consumers.

▷ It is illuminating to compare the above equation to the monopolist's first–order condition, which is

$$\left(p-c\right) f\left(p\right) = 1 - F\left(p\right)$$

The first-best $q^{FB}\left(\theta\right)$ is defined as the solution to

$$u_q\left(q^{FB}\left(\theta\right),\theta\right) = C_q\left(q^{FB}\left(\theta\right)\right)$$

Then for $\theta = \bar{\theta}$, there is no distortion in consumption: $q\left(\theta\right) = q^{FB}\left(\theta\right)$. This is referred to as the "no distortion at the top" screening result.

6.1.9 Example with Linear Utility

Consider the following setup: $C\left(q\right) = cq, q \in \{0,1\}$ and $u\left(q,\theta\right) = \theta q$. Then seller chooses ϕ (probability of getting the good that is a function of θ) to maximize:

$$\begin{aligned}\Lambda &\equiv u\left(q\left(\theta\right),\theta\right) - u_\theta\left(q\left(\theta\right),\theta\right)\frac{1-F\left(\theta\right)}{f\left(\theta\right)} - C\left(q\left(\theta\right)\right) \\ &= \phi\left(\theta\right)\left(\theta - c - \frac{1-F\left(\theta\right)}{f\left(\theta\right)}\right)\end{aligned}$$

Assuming F satisfies the MHRC, the program becomes regular, which allows to solve this problem point-wise. Thus, the optimal $\phi\left(\cdot\right)$ is:

$$\phi^*\left(\theta\right) = \begin{cases} 1 & \text{if } \theta - c \geq \frac{1-F(\theta)}{f(\theta)} \\ 0 & \text{if } \theta - c < \frac{1-F(\theta)}{f(\theta)} \end{cases}$$

In other words, even if the seller could randomize, he will have a deterministic mechanism as his optimal choice!

- ▷ Note that we denote
$$J(\theta) = \theta - \frac{1 - F(\theta)}{f(\theta)}$$
as the virtual type.

- ▷ Define θ^* as the type that satisfies $J(\theta^*) = c$ and write the corresponding indirect utility function as
$$U(\theta) = \int_{\underline{\theta}}^{\bar{\theta}} q(s)\, ds = \begin{cases} \theta - \theta^* & \text{if } \theta \geq \theta^* \\ 0 & \text{otherwise} \end{cases}$$
since $u_\theta = q$. Then the corresponding transfer function, $t(\theta) = \theta q(\theta) - U(\theta)$:
$$t(\theta) = \begin{cases} \theta^* & \text{if } \theta \geq \theta^* \\ 0 & \text{otherwise} \end{cases}$$

Thus, we have found the optimal direct mechanism that solves the seller's program. What is the indirect mechanism that we see in the real world which accomplishes the same thing?

- ▷ The monopolist posts the price $p = \theta^*$ and the buyer purchases a unit if and only if $\theta > p$. So at the end of a lot of work, we are back to our simple monopoly program of picking a price.

- ▷ We also showed earlier that the seller can do nothing better (i.e. randomization cannot improve profits).

6.1.10 Non-linear Pricing for General Preferences

We want to derive a non-linear price schedule P that the monopolist offers to consumers, and the consumer decides which q to choose. We want to create the price schedule so that it implements the same output-utility allocation $\{q, U\}$ as the direct mechanism. The fact that such menu can always be constructed is known as the **taxation principle**.

Proposition 6.3. *(Taxation Principle) For any incentive compatible (deterministic) direct mechanism $\{q(\cdot), t(\cdot)\}$ that implements $U(\cdot)$, there exists a non-linear price function $P : Q \to \mathbb{R}$ such that*
$$U(\theta) = \max_{q \in Q} u(q, \theta) - P(q)$$
$$q(\theta) = \arg\max_{q \in Q} u(q, \theta) - P(q)$$

Proof. If $\{q(\cdot), t(\cdot)\}$ is an IC-DSM, then define
$$P(q) = \begin{cases} t(\theta) & \text{if } q = q(\theta) \\ \infty & \text{otherwise} \end{cases}$$
The buyer has the same set of price-output pairs available under $P(q)$ as with the direct mechanism $\{q(\cdot), t(\cdot)\}$. Given the direct mechanism is incentive-compatible, then type θ buys $q(\theta)$. ■

Thus, given functions for $q(\cdot)$ and $t(\cdot)$, it is straightforward to construct $P(q)$ by inverting $q(\cdot)$. Then we can define
$$\theta(q) = \max_{\theta} \{\theta \in \Theta | q(\theta) = q\}$$
and construct $P(q) = t(\theta(q))$. We illustrate with an example.

Example 6.2. Suppose $u(q,\theta) = \theta q - 0.5q^2$ and $C(q) = q$ with $\theta \sim$ Uniform $[1,2]$.

- ▷ If we were to get the first-best solution, we would solve: $\max_q u(q,\theta) - C(q)$ which yields: $q^{FB}(\theta) = \theta - 1$.
- ▷ Now consider the monopolist's problem. Applying steps $(1)-(3)$, we have:

$$\max_{q(\theta)} \mathbb{E}[\Lambda]$$

assuming $\underline{U}$ where

$$\begin{aligned}
\Lambda &\equiv u(q(\theta),\theta) - u_\theta(q(\theta),\theta)\frac{1-F(\theta)}{f(\theta)} - C(q(\theta)) \\
&= \left[\theta q - \frac{1}{2}q^2\right] - \left[q\frac{1-(\theta-1)}{2-1}\right] - [q] \\
&= \left[\theta q - \frac{1}{2}q^2\right] - [q(2-\theta)] - [q] \\
&= -\frac{1}{2}q^2 - 3q + 2\theta q
\end{aligned}$$

- ▷ Solving this pointwise, we attain $q(\theta) = \max\{0, 2\theta - 3\}$ where the max operator exists to guarantee non-negativity.
- ▷ If you graph this out, $q(\theta)$ lies below $q^{FB}(\theta)$. It turns out that

$$U(\theta) = \int_1^\theta q^*(s)\,ds = \begin{cases} \frac{9}{4} - (3-\theta)\theta & \theta \geq \frac{3}{2} \\ 0 & \theta \leq \frac{3}{2} \end{cases} = \max\left\{0, \frac{9}{4} - (3-\theta)\theta\right\}$$

- ▷ Using this expression, we can back out $t(\theta)$ since

$$t(\theta) = u(q(\theta),\theta) - U(\theta) = \cdots = \begin{cases} (6-\theta)\theta - \frac{27}{4} & \text{if } \theta \geq \frac{3}{2} \\ 0 & \text{if } \theta < \frac{3}{2} \end{cases}$$

- ▷ To find $P(q)$, we invert $q(\cdot)$ and obtain, for any $q \geq 0$,

$$\theta(q) = \frac{q+3}{2} \text{ if } q \geq 0$$

which implies

$$\underline{P}(q) = t(\bar{\theta}(q)) = \left(6 - \left(\frac{q+3}{2}\right)\right)\left(\frac{q+3}{2}\right) - \frac{27}{4} = \frac{1}{4}(6-q)q = \frac{3}{2}q - \frac{1}{4}q^2$$

How can we check that this works?

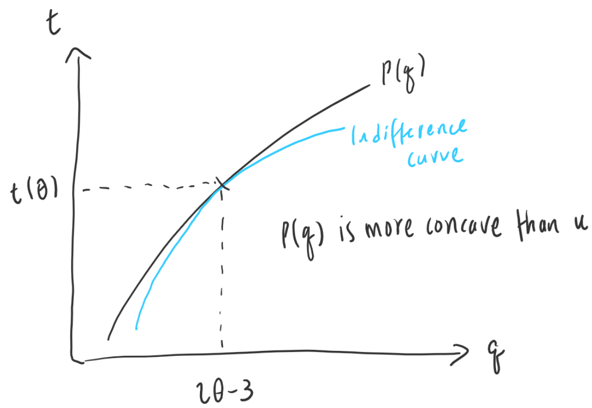
- ▷ Consider the program of the consumer:

$$\max_q \underbrace{\theta q - \frac{1}{2}q^2}_{u(q,\theta)} \underbrace{-\frac{3}{2}q + \frac{1}{4}q^2}_{-P(q)}$$

Solving this program, we have

$$q = 2\theta - 3$$

Graphically, we have:



6.1.11 Wilson (1993)'s Non-linear Pricing

Consider the setup with $u(q,\theta) = \theta v(q)$. Rearrange the optimality condition:

$$u_q - C'(q) = \frac{1-F(\theta)}{f(\theta)} u_{\theta q}(q(\theta), \theta)$$

and using the fact that $P'(q) = u_q(q,\theta) = \theta v'(q)$ yields:

$$\frac{P'(q) - C'(q)}{P'(q)} = \frac{1-F(\theta)}{f(\theta)}$$

We can think of the RHS as the marginal inverse inverse elasticity of demand since the typical monopoly condition is the following:

$$\frac{p - MC(q)}{p} = \frac{1-F(\theta)}{\theta f(\theta)}$$

so we have:

$$\epsilon(q) = \frac{\bar{\theta}(q) f\left(\bar{\theta}(q)\right)}{1 - F\left(\bar{\theta}(q)\right)}$$

Wilson (1993) further defines the *cumulative demand function* as

$$N(p,q) = \mu(\{\theta | D(p,\theta) \geq 1\})$$

6.1.12 Mussa-Rosen (1978)'s Quantity Discounts $(P''(q) < 0)$

Consider the setup where $u(q,\theta) = \theta q$ and $C(q) = 0.5q^2$. This is an example where $P(q)$ is convex. Skipping a bunch of the algebra, the problem is characterized as the following series of equations:

$$P'(q) = u_q\left(q, \bar{\theta}(q)\right)$$
$$P''(q) = u_{qq} + u_{q\theta}\left(\frac{-\Lambda_{qq}}{\Lambda_{q\theta}}\right)$$

Note that $u(q,\theta)$ is strictly concave and quadratic and $c(q) = cq$, then we have $P''(q) < 0$. Note that this is interpreted as a quantity discount i.e. the marginal price is falling in quantity since P is the total price and P' gives you the marginal price.

6.2 Two-Type Model

Sometimes, it is convenient to focus on simple models with only two types. Here we present our non-linear pricing model with two types, $\theta \in \left\{\underline{\theta}, \bar{\theta}\right\}$ to illustrate how the marginal conditions in the continuously-distributed type case appear as simply inequalities. We will later generalize the setup to n types to provide a clearer analog to the continuous case.

6.2.1 Setup

There are two types with $\theta_2 > \theta_1$. Utility is given as $u\left(q, \theta\right) = \theta q - 0.5q^2$ and $C\left(q\right) = cq$. We will *not* follow the five steps since it's difficult to fit this problem into step 1. We will restrict our contracts of the form

$$\left\{\left(q_1, t_1\right), \left(q_2, t_2\right)\right\}$$

such that type θ_1 consumers find it optimal to choose $\left(q_1, t_1\right)$ and θ_2 consumers find it optimal to choose $\left(q_2, t_2\right)$. Thus the firm's optimization program can be wirtten as:

$$\begin{aligned}
\max_{\left\{\left(q_1, t_1\right), \left(q_2, t_2\right)\right\}} & \phi\left(t_2 - cq_2\right) + \left(1-\phi\right)\left(t_1 - cq_1\right) \\
\text{s.t. } & [IR2] : \theta_2 q_2 - 0.5q_2^2 - t_2 \geq 0 \\
& [IR1] : \theta_1 q_1 - 0.5q_1^2 - t_1 \geq 0 \\
& [IC2] : \theta_2 q_2 - 0.5q_2^2 - t_2 \geq \theta_2 q_1 - 0.5q_1^2 - t_1 \\
& [IC1] : \theta_1 q_1 - 0.5q_1^2 - t_1 \geq \theta_1 q_2 - 0.5q_2^2 - t_2
\end{aligned}$$

We want to simplify the constraints.

▷ To proceed, we will use the results that $[IR1] + [IC2] \Rightarrow [IR2]$. That is, I can throw away the high guy's $[IR2]$ constraint.

  * This is because $[IC2]$ says
  $$\theta_2 q_2 - 0.5q_2^2 - t_2 \geq \theta_2 q_1 - 0.5q_1^2 - t_1$$
  and since $\theta_2 > \theta_1$,
  $$> \theta_1 q_1 - 0.5q_1^2 - t_1 \geq 0$$
  where the last inequality is from the $[IR1]$. In general, all the IR constraints except for the worst guy will disappear.

▷ Now let's think about $[IC2]$. We will prove that $[IC2]$ will bind at the optimum.

  * Suppose it were slack. This means that I can raise t_2 and increase profits while not affecting $[IR1]$ and making $[IC1]$ actually easier to satisfy.

▷ If $[IC2]$ binds, then $[IC1]$ will be slack if $q_2 - q_1 \geq 0$ so we can get rid of it.

Thus, we solve a relaxed problem where we maximize subject to $[IR1]$ and $[IC2]$.

▷ Now the program is characterized as

$$\begin{aligned}
\max_{\left\{\left(q_1, t_1\right), \left(q_2, t_2\right)\right\}} & \phi\left(t_2 - cq_2\right) + \left(1-\phi\right)\left(t_1 - cq_1\right) \\
\text{s.t. } & [IR1] : \theta_1 q_1 - 0.5q_1^2 - t_1 = 0 \\
& [IC2] : \theta_2 q_2 - 0.5q_2^2 - t_2 = \theta_2 q_1 - 0.5q_1^2 - t_1
\end{aligned}$$

▷ Using the equalities in the constraints, we have

$$\begin{aligned} t_1 &= \theta_1 q_1 + \frac{1}{2} q_1^2 \\ t_2 &= \theta_2 q_2 - \frac{1}{2} q_2^2 - \left(\theta_2 q_1 - \frac{1}{2} q_1^2\right) + \left(\theta_1 q_1 - \frac{1}{2} q_1^2\right) \\ &= \theta_2 q_2 - \frac{1}{2} q_2^2 - \Delta\theta q_1 \end{aligned}$$

▷ Substituting this into the maximization program:

$$\max_{\{(q_1,t_1),(q_2,t_2)\}} \phi \left[\theta_2 q_2 - c(q_2) - \Delta\theta q_1 - \frac{1}{2} q_2^2\right] + (1-\phi)\left[\theta_1 q_1 - c(q_1) - \frac{1}{2} q_1^2\right]$$

subject to $q_2 \geq q_1$. Ignoring the monotonicity constraints, the first-order constraints imply:

$$\begin{aligned} [q_2] &: \phi(\theta_2 - q_2 - c) = 0 \\ [q_1] &: (1-\phi)(\theta_1 - c - q_1) + \phi\Delta\theta \end{aligned}$$

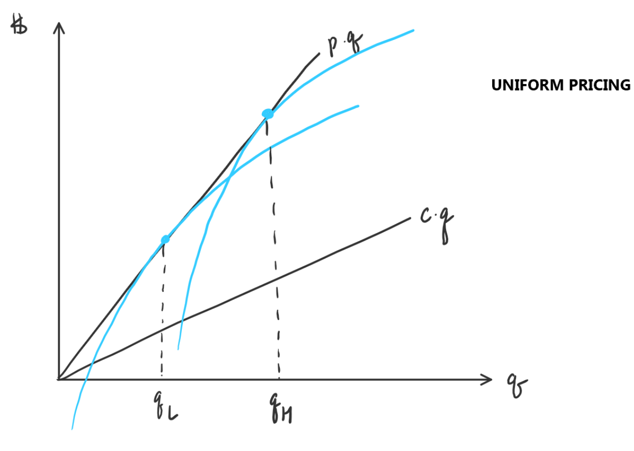
▷ Simplifying, we have

$$\begin{aligned} q_2 &= \theta_2 - c = q_2^{FB} \\ q_1 &= \theta_1 - c - \frac{\phi}{1-\phi}\Delta\theta \end{aligned}$$

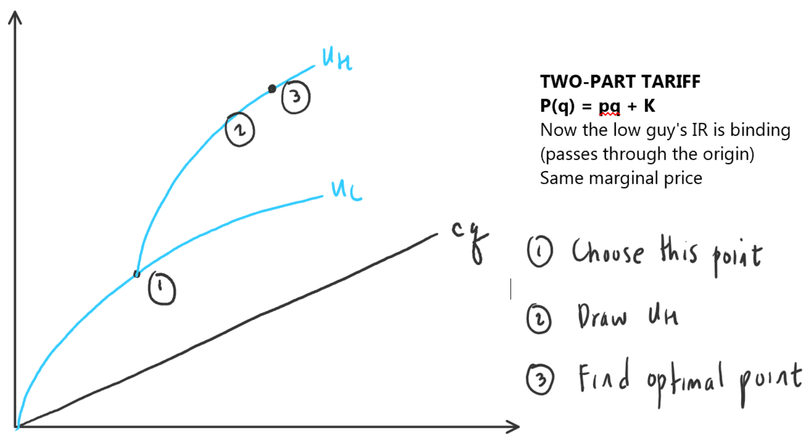
6.2.2 Intuition

We describe the intuition below.

1. Uniform Pricing: The traditional monopoly problem yields the following figure

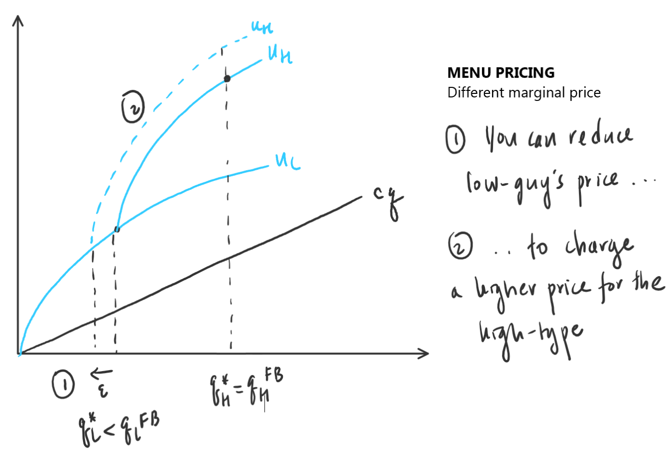


2. Two-Part Tariff: Now we shift the low guy's indifference curve so that it passes through the origin



- ▷ Let u_L pass through the origin.

3. Menu Pricing:



6.2.3 n-type distribution

Now we generalize our results for the 2-type setting with $\theta_1 < \theta_2, ..., < \theta_{n-1} < \theta_n$ with probability $p_i \in (0,1)$ for each type and distribution function $P_i \equiv \sum_{j=1}^{i} p_j$. Therefore, the principal's program is choosing an output-utility allocation $\{q_i, U_i\}_{i=1}^n$ where $U_i = u(q_i, \theta_i) - t_i$:

$$\max_{\{q_i, U_i\}_{i=1}^n} \sum_{i=1}^{n} p_i \{u(q_i, \theta_i) - C(q_i) - U_i\}$$
$$\text{s.t. } U_i \geq u(q_j, \theta_i) - t_j$$
$$U_i \geq \underline{U}$$

Once again, we can eliminate many of the constraints and focus on local incentive compatibility. We skip the details but here is the punchline:

- ▷ In the case of finite distribution of types, the optimal mechanism q_i satisfies:

$$p_i \left(u_q(q_i, \theta_i) - C'(q_i) \right) = [1 - P_i] \left(u_q(q_i, \theta_{i+1}) - u_q(q_i, \theta_i) \right), \quad i = 1, ..., n,$$

and t_i is chosen as the unique solution to the first-order difference equation,

$$U_i - U_{i-1} = u\left(q_{i-1}, \theta_i\right) + u\left(q_{i-1}, \theta_{i-1}\right)$$

with the initial condition $U_1 = \underline{U}$.

▷ Just as before, we have no distortion at the top and a suboptimal level of activity for all lower types.

6.3 Regulation (Baron & Myerson, 1982)

There is a large body of work on regulating firms with private information that was initiated by Baron and Myerson (1982) and brought to fruition by a number of interesting papers written in the 1980s and 1990s by Laffont, Tirole, Lewis, Sappington, McAfee and McMillan. Here, we present a simple version of Baron and Myerson (1982) to illustrate how the methods used in nonlinear pricing can easily be adapted to settings with private information over costs.

6.3.1 Setup

Consider the case where a regulator must set the price for a natural monopolist with private information about marginal cost. Specifically, the regulator sets a price at which consumers purchase output (p), a transfer to the firm t (in addition to the revenues it earns from the consumers) requiring that the monopolist satisfies all demand at the set price.

▷ Firm's cost of production: $C\left(q\right) = \theta q + K$ where K is a fixed capital cost and $\theta \sim F\left(\theta\right)$ on $\left[\underline{\theta}, \bar{\theta}\right]$.

▷ Agent: $\Pi = R\left(q\right) + t - \theta q - K$ where $R\left(q\right)$ is the consumer revenue

▷ Principal (regulator) cares about consumer surplus minus the payments made to the monopolist: $CS(q) - t$

6.3.2 Solving the Model

We proceed through the following steps:

1. **Show implementability**
   Define type-θ monopolist's payoff, when reporting $\hat{\theta}$ as

$$\Pi\left(\hat{\theta}|\theta\right) = R\left(q\left(\hat{\theta}\right)\right) - \theta q\left(\hat{\theta}\right) - K + t\left(\hat{\theta}\right)$$

   We want to characterize the set of implementable $\{q\left(\cdot\right), \Pi\left(\cdot\right)\}$. Following the same arguments as in the case of nonlinear pricing, we have:

   ▷ First, note that

$$\Pi'\left(\theta\right) = \Pi_2\left(\theta|\theta\right) = -q\left(\theta\right) \Rightarrow \underline{\Pi\left(\theta\right) = \Pi\left(\underline{\theta}\right) - \int_{\underline{\theta}}^{\theta} q\left(s\right) ds}$$

   or equivalently:

$$\Pi\left(\theta\right) = \Pi\left(\underline{\theta}\right) + \int_{\theta}^{\bar{\theta}} q\left(s\right) ds$$

   ▷ Second, note that

$$\Pi_{12}\left(\theta|\theta\right) \geq 0 \Rightarrow \underline{q'\left(\theta\right) \leq 0}$$

   Note that this is an environment where the participation constraint binds for the highest-cost firm.

2. **Expression for $\mathbb{E}[\Pi(\theta)]$ that is a function of $(\Pi(\bar{\theta}))$**
   Start from noting that

$$\Pi(\theta) = \Pi(\bar{\theta}) + \int_{\theta}^{\bar{\theta}} q(s)\, ds$$

   So taking the expectation and using integration by parts:

$$\begin{aligned}
\mathbb{E}[\Pi(\theta)] &= \Pi(\bar{\theta}) + \int_{\underline{\theta}}^{\bar{\theta}} \left(\int_{\theta}^{\bar{\theta}} q(s)\, ds \right) f(\theta)\, d\theta \\
&= \Pi(\bar{\theta}) + \left[\left(\int_{\theta}^{\bar{\theta}} q(s)\, ds \right) F(\theta) \right]_{\underline{\theta}}^{\bar{\theta}} + \int_{\underline{\theta}}^{\bar{\theta}} q(\theta) F(\theta)\, d\theta \\
&= \Pi(\bar{\theta}) + \int_{\underline{\theta}}^{\bar{\theta}} q(\theta) F(\theta)\, d\theta \\
&= \Pi(\bar{\theta}) + \mathbb{E}\left[q(\theta) \frac{F(\theta)}{f(\theta)} \right]
\end{aligned}$$

3. **Rewrite the Regulator's Program**
   The regulator solves

$$\max_{q(\cdot)} \quad \mathbb{E}[CS(q) - t]$$

   Since $\Pi = R(q) - \theta q - k + t$ which implies

$$t(\theta) = \Pi(\theta) - R(q(\theta)) + \theta q(\theta) - K$$

   and thus the regulator's problem can be rewritten as

$$\max_{\theta, q(\cdot)} \quad \mathbb{E}[CS(q) + R(q(\theta)) - \theta q(\theta) - K - \Pi(\theta)]$$

   Substituting in the expression for $\Pi(\theta)$, we have:

$$\max_{q(\cdot)} \quad \mathbb{E}\left[CS(q) + R(q(\theta)) - \left(\theta + \frac{F(\theta)}{f(\theta)} \right) q(\theta) - K \right]$$

   Defining $\Lambda(q, \theta)$ as the term inside the bracket, $\Lambda_{q\theta} \leq 0$ if we have MHRC. Therefore, combined with the MHRC assumption, the pointwise-optimum satisfies:

$$\underbrace{CS'(q(\theta)) + R'(q(\theta))}_{p(\theta) \text{ on demand curve}} - \theta = \frac{F(\theta)}{f(\theta)} \geq 0$$

We see that output is distorted below the full information solution for all $\theta > \underline{\theta}$ in an attempt to reduce monopoly rents due to private information.

6.4 Hybrid Models of Hidden Information / Hidden Action

Myerson (1982) introduces moral hazard into hidden information models and extends the revelation principle to include a notion of *obedience*.

6.4.1 The Idea

Suppose that agent has private information θ but also has a private action e (effort) which stochastically impacts the principal's returns x that is captured in a conditional density function $f(x|e)$ as studied in the moral hazard framework. Then Myerson's general revelation principle states that the principal can restrict attention to contracts of the following form:

▷ After θ is realized, the agent reports $\hat{\theta} \in \Theta$.

▷ The mechanism announces an effort recommendation $e\left(\hat{\theta}\right)$ and a compensation scheme $w\left(x|\hat{\theta}\right)$.

▷ For every θ, the agent is truthful (i.e announces $\hat{\theta} = \theta$) and obedient (the agent finds it optimal to choose effort $e(\theta)$).

6.4.2 Laffont and Tirole (JPE, 1986)

We illustrate the above setup using an example. Consider a regulated firm with private information about its cost θ.

▷ Firm: total cost of production is $C(q) = (\theta - e)q$ where the effort is costly at $\psi(e)$. The firm is exerting the effort e and the associated utility is given as

$$U = t - \psi(e)$$

where ψ is increasing, strictly convex, and $\psi'''(e) \geq 0$.

▷ The authors assume a novel contracting structure in which the regulator can observe costs but must determine how much of the costs are attributable to effort and how much are attributed to inherent luck (i.e. type).

* The regulator can ask the firm to report its type and assign the firm a marginal cost target of $c\left(\hat{\theta}\right)$ and an output level of $q\left(\hat{\theta}\right)$ in exchange for compensation equal to $t\left(\hat{\theta}\right)$.

* This means that a firm with type θ that wishes to make the marginal cost target of $c\left(\hat{\theta}\right)$ must expend effort equal to $e = \theta - c\left(\hat{\theta}\right)$.With such a contract, the firm's indirect utility function becomes:

$$U\left(\hat{\theta}|\theta\right) \equiv t\left(\hat{\theta}\right) - \psi\left(\theta - c\left(\hat{\theta}\right)\right)$$

Suppose the regulator is interested in maximizing a weighted average of strictly concave consumer surplus $CS(q)$ less costs and transfers and producer surplus U:

$$\max_{q,c,t} \mathbb{E}_\theta \left[CS(q(\theta)) - c(\theta)q(\theta) - t(\theta) + \gamma U(\theta)\right]$$
$$\text{s.t. implementability}$$

Using the expression f or $t(\theta)$, we can rewrite the above as:

$$\max_{q,c,t} \mathbb{E}_\theta \left[CS(q(\theta)) - c(\theta)q(\theta) - \psi(\theta - c(\theta)) - (1-\gamma)U(\theta)\right]$$
$$\text{s.t. implementability}$$

Now we go through the steps as before:

1. Implementability:

$$(1) : c(\theta) \uparrow$$
$$(2) : U(\theta) = U\left(\bar{\theta}\right) + \int_\theta^{\bar{\theta}} \psi'(s - c(s))\, ds$$

2. Expression for $\mathbb{E}[U(\theta)]$:

$$\mathbb{E}[U(\theta)] = U(\bar{\theta}) + \mathbb{E}\left[\psi'(\theta - c(\theta))\frac{F(\theta)}{f(\theta)}\right]$$

3. Regulator's program:

$$\max_{q,c}\mathbb{E}\left[CS(q(\theta)) - c(\theta)q(\theta) - \psi(\theta - c(\theta)) - (1-\gamma)U(\bar{\theta}) - (1-\gamma)\psi'(\theta - c(\theta))\frac{F(\theta)}{f(\theta)}\right]$$
$$\text{s.t. } c(\theta) \text{ is non-decreasing}$$
$$U(\bar{\theta}) \geq 0$$

Taking the FOCs yields:

$$[q] : CS'(q(\theta)) = c(\theta)$$
$$[c] : q(\theta) - \psi'(e(\theta)) = (1-\gamma)\frac{F(\theta)}{f(\theta)}\psi''(e(\theta)) \geq 0$$

▷ $[c]$ tells us that the condition for the optimal choice of effort. Recall that the first-best effort, conditional on q, is:

$$q(\theta) - \psi'(e(\theta)) = 0$$

As a consequence, suboptimal effort is provided for every type except the lowest person. We always have no distortion on the bottom.

▷ $[q]$ corresponds to the full-information efficient production level conditional on the marginal cost $c(\theta) \equiv \theta - e(\theta)$.

The optimal non-linear contract can be implemented using a realistic menu of two-part tariffs (cost-sharing contracts).

6.5 Random Mechanisms

We must address whether or not the seller gave up expected profit by restricting attention to deterministic mechanisms. A clever paper by Starusz (JET, 2006) clarified much of the confusion and settled the question once and for all in the context of a discrete-type model with quasi-linear preferences. He found that if the screening environment is regular, then a random scheme cannot improve upon a deterministic one. The key insight here is that in a quasi-linear world with risk-neutral principal and agent, a random scheme can only be valuable if it helps relax a binding monotonicity constraint.

6.5.1 Setup

We will consider the following class of random mechanisms: $\{\phi(q|\theta), t(\theta)\}$ as opposed to the previous deterministic setup $\{q(\theta), t(\theta)\}$. Note that we only focus on randomizations for q but not for t, since principal is risk-averse while agents are risk-neutral.

▷ General form for utility is given as

$$U(\hat{\theta}|\theta) = \int_{\underline{q}}^{\bar{q}} \phi(q|\hat{\theta})\, u(q,\theta)\, dq - t(\hat{\theta})$$

6.5.2 Characterizing the Random Mechanism

1. First, establish that $\{\phi, t\}$ is IC if and only if :

   ▷ $\int_{\underline{q}}^{\bar{q}} \phi\left(q|\hat{\theta}\right) u_\theta\left(q,\theta\right) dq|_{\hat{\theta}=\theta}$ is increasing in $\hat{\theta}$ i.e. expectation of $U\left(\hat{\theta}|\theta\right)$ is increasing in $\hat{\theta}$.

   * It's weaker than monotonicity or FOSD. We can find it by writing out the inequalities for telling the truth vs. not telling the truth.

   ▷ $U\left(\theta\right)$ is absolutely continuous, and

   $$U\left(\theta\right) - U\left(\theta'\right) = \int_{\theta}^{\theta'} \int_{\underline{q}}^{\bar{q}} \phi\left(q|\tilde{\theta}\right) u_\theta\left(q, \hat{\theta}\right) dq d\tilde{\theta}$$

   * The idea is to take the equation

   $$u\left(\hat{\theta}|\theta\right) = \int_{\underline{q}}^{\bar{q}} \phi\left(q|\hat{\theta}\right) u\left(q,\theta\right) dq - t\left(\hat{\theta}\right)$$

   and differentiate each side respect to θ while applying the Envelope theorem. If you do the same arguments using the limit, you will still get the same results.

2. Second, write:

$$\mathbb{E}\left[U\left(\theta\right)\right] = u\left(\underline{\theta}\right) + \mathbb{E}_\theta\left[\left(\int_{\underline{q}}^{\bar{q}} \phi\left(q|\theta\right) u_\theta\left(q,\theta\right) dq\right)\left(\frac{1-F\left(\theta\right)}{f\left(\theta\right)}\right)\right]$$

3. Third, stick the above expression into the objective function for the principal:

$$\max_{\phi(\cdot)} \mathbb{E}_\theta\left[\int_{\underline{q}}^{\bar{q}} \phi\left(q|\theta\right) \underbrace{\left(u\left(q,\theta\right) - c\left(q\right) - u_\theta\left(q,\theta\right)\frac{1-F\left(\theta\right)}{f\left(\theta\right)}\right)}_{\Lambda(q,\theta)}\right] - U\left(\underline{\theta}\right)$$

   subject to $\int_{\underline{q}}^{\bar{q}} \phi\left(q|\hat{\theta}\right) u_\theta\left(q,\theta\right) dq|_{\hat{\theta}=\theta}$ is increasing in $\hat{\theta}$.

   ▷ Since this is difficult to solve, let's find ϕ pointwise. Suppose Λ is regular (it has peak) so we can put all mass on the peak (unique *q*) that maximizes:

   $$u\left(q,\theta\right) - C\left(q\right) - u_\theta\left(q,\theta\right)\frac{`1-F\left(\theta\right)}{f\left(\theta\right)}$$

   ▷ Now, if you do this for each value of θ, then you just get a peak for each outcome. which leads to a deterministic ϕ.

In other words, we have the result that if $\Lambda\left(q,\theta\right)$ is regular, then

$$\phi\left(q|\theta\right) = \begin{cases} \theta_1 & \text{if } q = q^{MR}\left(\theta\right) \\ 0 & \text{if } q \neq q^{MR}\left(\theta\right) \end{cases}$$

6.6 Countervailing Incentives

In some settings, the relevant outside option may be increasing in θ so that once we incorporate the participation constraint, we may be left with a situation where u_θ changes sign endogenously as a function of $q(\cdot)$. This is referred to as "countervailing incentive" and was first explored by Lewis and Sappington (JET 1989; AER 1989)

6.6.1 Setup

Consider $u(q,\theta) = \theta q$ with $u_\theta \geq 0$. Suppose the consumers also have the option to buy a market substitute with fixed output q_0 at price p_0:

$$\underline{U}(\theta) = \max\{0, \theta v(q_0) - p_0\}$$

i.e. a not-very-good substitute. If you don't buy from the monopolist, you will buy this crappy substitute. Therefore:

$$\tilde{u}(q,\theta) = \theta q - (\theta q_0 - p_0) = \theta(q - q_0) + p_0$$

but notice that the sign of $\tilde{u}_\theta$ is ambiguous – it changes at $q = q_0$. Therefore, while it is still the case that any IC mechanism will satisfy:

$$U(\theta) = U(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} v(q(s))\, ds$$

setting $U(\underline{\theta}) = 0$ will possibly violate the IR constraint for a higher type.

6.6.2 Solving Model a la Lewis-Sappington

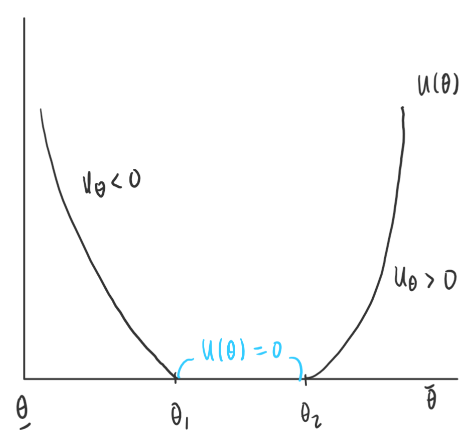
The maximization problem can be represented as

$$U(\theta) = \max_{\hat{\theta}} \theta\left(q\left(\hat{\theta}\right) - q_0\right) + p_0 - t\left(\hat{\theta}\right)$$

The logic for solving this problem is as follows.

1. Argue that $U(\theta)$ is convex and that $[IR]$ will bind over an interval.

   In this case, the agent's utility is linear in θ, and it therefore follows that $U(\theta)$ will be convex. This implies that the $[IR]$ constraint will bind over an interval in which $U'(\theta) = 0$ with $q(\theta) = q_0$ and $t(\theta) = p_0$. We know that the set is non-empty since we can always raise the price and shift around the graph. Denote the interval as $[\theta_1, \theta_2]$.



2. Over $[\underline{\theta}, \theta_1]$, we have $u_\theta \leq 0$ so set $U(\theta_1) = 0$ to satisfy the $[IR]$ constraint.

   Here the problem looks very much like the Baron-Myerson (BM) setting of regulating a firm with unknown marginal cost, so we can show that

$$\int_{\underline{\theta}}^{\theta_1} U(\theta) f(\theta) d\theta = F(\theta_1) U(\theta_1) - \int_{\underline{\theta}}^{\theta_1} (v(q(\theta)) - v(q_0)) \frac{F(\theta)}{f(\theta)} f(\theta) d\theta$$

   Solving this pointwise, optimal $q(\cdot)$ satisfies:

$$\theta v'(q(\theta)) - C'(q(\theta)) = -\frac{F(\theta)}{f(\theta)} v'(q(\theta)) < 0$$

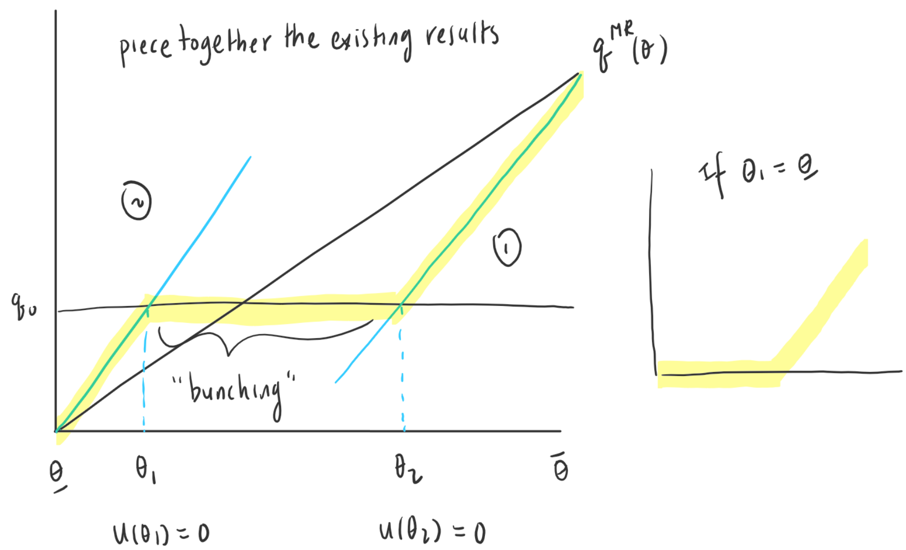
3. Over $[\theta_2, \bar{\theta}]$, we have $u_\theta \geq 0$ and thus find $q(\cdot)$ to solve $\Lambda_q(q(\theta), \theta) = 0$.

   The right region can be solved analogously.

4. Having solved both left and right problems, the firm now chooses θ_1 and θ_2 optimally.

   The result is kind of a smooth-pasting property which ensures that the three components of $q(\cdot)$ comes together continuously.

The intuition is as following. Suppose someone told you θ_1 and θ_2. How would you solve this problem? We can piece together existing results to obtain something like this:



6.7 Dynamic Screening via Courty-Li (2000)

We consider an extension of the static model to dynamic screening.

6.7.1 General Setup

We revisit our monopoly model with the simple setting of linear utility but with two periods. Let $u = \theta_1 q_1 + \theta_2 q_2$ where $\theta_1 \sim F(\theta_1)$ and $\theta_2 \sim G(\theta_2 | \theta_1)$. Alternatively, we can write:

$$\theta_2 = \lambda\theta_1 + (1 - \lambda)\epsilon, \quad \epsilon \sim H()$$

The timing is as follows:

▷ Nature chooses $\theta_1 \sim F(\theta_1)$ for the buyer

▷ Seller offers a mechanism (long-term contract which governs consumption not only today but also tomorrow) to the buyer. Using the revelation principle, we can write the direct mechanism as the following:

$$\left\{q_1\left(\hat{\theta}_1\right), q_2\left(\hat{\theta}_1, \hat{\theta}_2\right), t\left(\hat{\theta}_1, \hat{\theta}_2\right)\right\}$$

Note that the transfer is single since we don't have any discounting in this economy. If buyer accepts, he reports $\hat{\theta}_1$ and $q_1\left(\hat{\theta}_1\right)$ is consumed.

▷ Nature chooses $\theta_2 \sim G(\theta_2|\theta_1)$ for the buyer.

▷ Buyer reports $\hat{\theta}_2$, obtaining $q_2\left(\hat{\theta}\right)$ while incurring total cost of $t\left(\hat{\theta}_1, \hat{\theta}_2\right)$.

Recall that in the static world where there is only one period and F satisfies $MHRC$, the optimal mechanism $q^*(\theta)$ satisfies:

$$\theta - C'(q^*(\theta)) = \frac{1 - F(\theta)}{f(\theta)}$$

6.7.2 Case 1: $\theta_1 = \theta_2 = \theta$.

Baron and Besanko (1984) show that if $\theta_1 = \theta_2$, then the optimal dynamic screening allocation is the replication of the optimal one-period allocation: $q_1(\theta) = q_2(\theta) = q^*(\theta)$.

▷ First, note that repeating the static mechniams twice is both IC and IR.

▷ Second, suppose that the optimal IC and IR mechanism is $q_1(\theta) \neq q_2(\theta)$ and consider forming a new mechanism that implements $\bar{q}(\theta) = 0.5(q_1(\theta) + q_2(\theta))$ each period with the original transfer functino. This is also IC and IR, and it has higher profits because costs are strictly convex, which is a contradiction.

Note that this result is only true under full commitment. It is entirely possible that after observing the agent's report in the full period, the principal would like to change the mechanism. Without full commitment, both the buyer and the seller would anticipate future renegotiations, which would destroy the IC of the original full-commitment mechanism.

6.7.3 Case 2: $\theta_1 \perp \theta_2$.

If the types are independent, the firm is in a position to implement the first best on q_2 by selling an option contract to the buyer (i.e. buy as much as you like in period 2 for $C(q)$) and charging a first-period price of

$$p = \mathbb{E}\left[\max_q \theta_2 q - C(q)\right]$$

The determination of q_1 follows the standard monopoly screening problem because the agent is privately informed about θ_1 at the time of contracting.

▷ In the second period, θ_2 has no relation to θ_1 (they're booking two different flights). So we can break this into two separate problems, since there are no connections between the θs.

▷ We know how to solve each problem separately. In the second period, the principal can choose a price so that there is zero consumer surplus, i.e.

$$P_2(q_2) = \mathbb{E}\left[\max_q \theta_2 q - C(q)\right] + C(q_2)$$

akin to a two-part tariff. This implies that

$$q_2 = q_2^{FB}(\theta), \quad \forall\theta$$

▷ In the first period, we have the standard problem:

$$\theta_1 - C'(q_1(\theta_1)) = \frac{1 - F(\theta_1)}{f(\theta_1)}$$

which is the $q^{MR}(\theta)$.

The key here is that information which has not yet been revealed at the contracting stage can be acquired at no cost in information rents. Because θ_1 contains no information about θ_2, the second-period allocation can be made efficient and all of the buyer's surplus over q_2 consumption can be captured.

6.7.4 Case 3: $\theta_1 \not\perp \theta_2$ a la Courty and Li (2000)

We first introduce the dynamic version of the revelation principle.

Proposition 6.4. *(Dynamic Revelation Principle) For every dynamic mechanism Γ and every buyer-optimal strategy σ^* in Γ, there is a direct mechanism $\tilde{\Gamma}$ and an optimal buyer strategy $\tilde{\sigma}^* = (\tilde{\sigma}_1^*, \tilde{\sigma}_2^*)$ such that*

1. $\tilde{\sigma}_1^*(\theta_1) = \theta_1, \forall\theta_1 \in \Theta$
2. $\tilde{\sigma}_2^*\left(\theta_1, \theta_2, \hat{\theta}_1\right)_{\hat{\theta}_1 = \theta_1} = \theta_2, \forall\theta_1, \theta_2 \in \Theta$
3. *For every (θ_1, θ_2), $\tilde{q}^*(\theta_1, \theta_2) = q^*(\theta_1, \theta_2)$ and $\tilde{t}^*(\theta_1, \theta_2) = t^*(\theta_1, \theta_2)$.*

We skip the proof. This principle allows us to look at direct mechanisms $q_1\left(\hat{\theta}_1\right)$, $q_2\left(\hat{\theta}_1, \hat{\theta}_2\right)$ and $t\left(\hat{\theta}_1, \hat{\theta}_2\right)$ which are IC with respect to both θ_1 and θ_2. Now we characterize the equilibrium.

$$U^2\left(\hat{\theta}_1 | \theta_1\right) \equiv \mathbb{E}_{\theta_2}\left[\theta_2 q_2\left(\hat{\theta}_1, \theta_2\right) - t\left(\hat{\theta}_1, \theta_2\right) | \theta_1\right]$$

1. $[IC2]$ **constraint in the second-period**

   Define the second-period utility of agent of type θ_2 who reported $\hat{\theta}_1$ in the first period:

   $$U^2\left(\hat{\theta}_2 | \hat{\theta}_1, \theta_1\right) = \theta_2 q_2\left(\hat{\theta}_1, \theta_2\right) - t\left(\hat{\theta}_1, \hat{\theta}_2\right)$$

   and also define the period-2 truthfully utility $\hat{\theta}_2 = \theta_2$ given the first-period report $\hat{\theta}_1$:

   $$U^2\left(\theta_2 | \hat{\theta}_1, \theta_2\right) = \theta_2 q_2\left(\hat{\theta}_1, \theta_2\right) - t\left(\hat{\theta}_1, \theta_2\right)$$

   Then the $[IC2]$ constraint requires:

   $$U^2\left(\hat{\theta}_1, \theta_2\right) \equiv U^2\left(\theta_2 | \hat{\theta}_1, \theta_2\right) \geq U^2\left(\hat{\theta}_2 | \hat{\theta}_1, \theta_1\right), \forall\theta_2, \hat{\theta}_2$$

   which is equivalent to:

- $\triangleright$ $q_2\left(\hat{\theta}_1, \theta_2\right)$ is non-decreasing in θ_2;
- $\triangleright$ $U^2\left(\hat{\theta}_1, \theta_2\right) = U^2\left(\hat{\theta}_1, \underline{\theta}\right) + \int_{\underline{\theta}}^{\theta_2} q_2\left(\hat{\theta}_1, s\right)$

Therefore, plugging in and using integration by parts:

$$U^2\left(\hat{\theta}_1|\theta_1\right) \equiv \mathbb{E}_{\theta_2}\left[\overbrace{\theta_2 q_2\left(\hat{\theta}_1, \theta_2\right) - t\left(\hat{\theta}_1, \theta_2\right)}^{U^2\left(\hat{\theta}_1, \theta_2\right)}|\theta_1\right]$$

$$= U^2\left(\hat{\theta}_1, \underline{\theta}\right) + \mathbb{E}_{\theta_2}\left[q_2\left(\hat{\theta}_1, \theta_2\right)\frac{1 - G\left(\theta_2|\theta_1\right)}{g\left(\theta_2|\theta_1\right)}|\theta_1\right]$$

2. $[IC1]$ **constraint in the first period**

The actual constraint is:

$$\theta_1 q_1\left(\theta_1\right) + U^2\left(\theta_1|\theta_1\right) \geq \theta_1 q_1\left(\hat{\theta}_1\right) + U^2\left(\hat{\theta}_1|\theta_1\right), \quad \forall \hat{\theta}_1, \theta_1$$

Given the result for the second-period, the expected utility for θ_1 when reporting $\hat{\theta}_1$ is:

$$U\left(\hat{\theta}_1|\theta_1\right) = \theta_1 q_1\left(\hat{\theta}_1\right) + U^2\left(\hat{\theta}_1, \underline{\theta}\right) + \mathbb{E}_{\theta_2}\left[q_2\left(\hat{\theta}_1, \theta_2\right)\frac{1 - G\left(\theta_2|\theta_1\right)}{g\left(\theta_2|\theta_1\right)}|\theta_1\right]$$

A familiar envelope result emerges for first-period utility. To do so, figure out the place shere true θ_1 shows up. If write out the integral, $g\left(\theta_2|\theta_1\right)$ will cancel, so we have θ_1 will appear in the first term and $G\left(\theta_2|\theta_1\right)$.

$$U'\left(\theta\right) = q_1\left(\theta_1\right) - \mathbb{E}_{\theta_2}\left[q_2\left(\theta_1, \theta_2\right)\frac{\frac{\partial}{\partial \theta_1}G\left(\theta_2|\theta_1\right)}{g\left(\theta_2|\theta_1\right)}|\theta_1\right]$$

which yields:

$$\mathbb{E}\left[U\left(\theta_1\right)\right] = \mathbb{E}_{\theta_1}\left[\mathbb{E}_{\theta_2|\theta_1}\left[q_1\left(\theta_1\right) - q_2\left(\theta_1, \theta_2\right)\frac{\frac{\partial}{\partial \theta_1}G\left(\theta_2|\theta_1\right)}{g\left(\theta_2|\theta_1\right)}\right]\frac{1 - F\left(\theta_1\right)}{f\left(\theta_1\right)}\right] + U\left(\underline{\theta}\right)$$

3. $[IR]$ constraint:

$$U\left(\theta_1\right) \equiv \theta_1 q_1\left(\theta_1\right) + U^2\left(\theta_1|\theta_1\right) \geq 0, \quad \forall \theta_1$$

Now the Airline's relaxed program:

$$\max_{q_1(\cdot), q_2(\cdot)} \mathbb{E}\left[\theta_1 q_1\left(\cdot\right) - C\left(q_1\left(\cdot\right)\right) + \theta_2 q_2\left(\cdot\right) - C\left(q_2\left(\cdot\right)\right) - \left(q_1\left(\cdot\right) - q_2\left(\cdot\right)\frac{G_{\theta_1}\left(\theta_2|\theta_1\right)}{g\left(\theta_2|\theta_1\right)}\right)\left(\frac{1 - F\left(\theta_1\right)}{f\left(\theta_1\right)}\right)\right] + U\left(\underline{\theta}\right)$$

The first-order conditions are:

$$[q_1] : \theta_1 - C'\left(q_1\left(\theta_1\right)\right) = \frac{1 - F\left(\theta_1\right)}{f\left(\theta_1\right)}$$

$$[q_2] : \theta_2 - C'\left(q_2\left(\theta_1, \theta_2\right)\right) = \underbrace{\frac{-G_{\theta_1}\left(\theta_2|\theta_1\right)}{g\left(\theta_2|\theta_1\right)}}_{[A]}\left(\frac{1 - F\left(\theta_1\right)}{f\left(\theta_1\right)}\right)$$

- $[q_1]$: This says that you are not going to distort first-period consumption when compared to the static model.
- $[q_2]$: The second-period quantity is distorted downwards as a function of the informativeness of the first-period signal. $[A]$ tells you the information content of the first-period. Trying $\lambda = 0$ and $\lambda = 1$ yields results consistent with the special cases. Why is this case? Recall that
$$\theta_2 = \lambda\theta_1 + (1 - \lambda)\,\epsilon$$
in which case
$$\frac{G_{\theta_1}(\theta_2|\theta_1)}{g(\theta_2|\theta_1)} = -\lambda$$

We want q_1 and q_2 to satisfy desired monotonicity properties. To do this we need to make the **dynamic regularity** assumption, i.e. the function
$$J(\theta_1, \theta_2) = \theta_2 + \left(\frac{\frac{\partial}{\partial\theta_1}G(\theta_2|\theta_1)}{g(\theta_2|\theta_1)}\right)\left(\frac{1 - F(\theta_1)}{f(\theta_1)}\right)$$
is non-decreasing in θ_1 and θ_2.

Interpretation In the case of the airline example, business travelers have high θ_1 (high variance abut their travel needs) while tourists have low θ_1. In $[q_2]$, for below-mean realizations of θ_2, it will be the case that $G_{\theta_1} > 0$ in which lower θ_1 types will be induced to consume an excess amount of q_2 relative to the first-best. This corresponds to offering tourists a cheaper ticket with a no or low refund policy. If the refund is lower than the airline's marginal cost, the tourist will take the flight, even when it is ineffficient to do so. Also see problem set for a more specific example.

7 Auctions

7.1 Four Standard Auctions

We explore four standard auctions. We characterize the Nash Equilibrium and the revenue.

1. First-price, Sealed Bid
2. Second-price, Sealed Bid (Vickrey)
3. Ascending Bid (English)
4. Clock (Dutch)

It turns out that (2) and (3) are very similar, and (1) and (4) are very similar. In fact, the solutions coincide.

7.1.1 Independent Private Values (IPV) Environment

We have risk-neutral bidders (to be relaxed later) with unit demands and a single unit for sale.

▷ Utility of the bidder: $v_i(\theta_1, ..., \theta_n)\,\phi_i - t_i$ where $(\theta_1, ..., \theta_n) \sim f(\theta_1, ..., \theta_n)$

In the independent private values environment, we have $v_i(\theta_1, ..., \theta_n) = \theta_i$ i.e. "private" and

$$f(\theta_1, ..., \theta_n) = f(\theta_1) \cdots f(\theta_n)$$

which is the "independent" part. In other words, I don't need to know whether you like or hate this wine. This is opposed to a common values environment where

$$v_i(\theta_1, ..., \theta_n) = \frac{1}{n} \sum_{i=1}^{n} \theta_i$$

An example would be the geologist's estimate of the oil in the ground.

7.1.2 First-price, Sealed Bid

Each person has a private information θ_i to figure out their bid b_i. We want to find an equilibrium mapping between θ_i and b_i. We will impose symmetry here, i.e.

$$F_i = F_j, \theta \in [0,1]$$

We will look for a symmetric equilibrium i.e. bidding function $\bar{b}(\theta)$ such that everyone wants to use this bidding function assuming everyone else does. We will solve this model in two different ways. The first way is the one involving differential equations; the second way is the one using Envelope theorem.

A. Differential-equation Approach

1. **Derive expression for $\bar{b}$ imposing symmetry**

   Assume for now that $\bar{b}$ is strictly increasing and differentiable. We know that it is weakly increasing by a revealed preference argument, so these are not strong assumptions.

   Consider the utility of the form $U_i(b_i, \theta_i)$:

$$U_i(b_i, \theta_i) = P\left(b_i > \max_{j \neq i} \bar{b}(\theta_j)\right)(\theta_i - b_i)$$

Noting that the cdf of $\max_{j\neq i}\theta_j$ is $G(\theta) = F(\theta)^{n-1}$, which means that when you differentiate G, you will have $(n-1)$ will pop up. this allows us to write the maximization problem as

$$\max_{b_i} G\left(\bar{b}^{-1}(b_i)\right)(\theta_i - b_i)$$

First-order condition:

$$-G\left(\bar{b}^{-1}(b_i)\right) + g\left(\bar{b}^{-1}(b_i)\right)\frac{d\bar{b}^{-1}(b_i)}{db_i}(\theta_i - b_i) = 0$$

In a symmetric equilibrium, player i also wants to use the same function, so we can replace b_i with $\bar{b}(\theta_i)$:

$$-G(\theta_i) + g(\theta_i)\frac{1}{\bar{b}'(\theta_i)}(\theta_i - b_i) = 0$$

Rewriting yields:

$$\begin{aligned} g(\theta_i)\,\theta_i &= G(\theta_i)\,\bar{b}'(\theta_i) + g(\theta_i)\,\bar{b}(\theta_i) \\ &= \frac{d}{d\theta_i}G(\theta_i)\,\bar{b}(\theta_i) \end{aligned}$$

Solving this differential equation yields:

$$\int_0^\theta g(\theta_i)\,\theta d\theta_i = \int dG(\theta_i)\,\bar{b}(\theta_i) = G(\theta)\,\bar{b}(\theta) + \text{ constant}$$

Using the initial condition that $b(0) = 0$ pins down the condition, so we have:

$$\begin{aligned} [1]: \bar{b}(\theta) &= \frac{1}{G(\theta)}\left(\int_{\underline{\theta}=0}^{\theta} xg(x)\,dx\right) \\ &= \frac{1}{G(\theta)}\left(xG(x)\,|_0^\theta - \int_0^\theta G(x)\,dx\right) \end{aligned}$$

which yields:

$$[2]: \bar{b}(\theta) = \theta - \int_0^\theta \left(\frac{G(x)}{G(\theta)}\right)dx = \theta - \int_0^\theta \left(\frac{F(x)}{F(\theta)}\right)^{n-1} dx$$

2. **Verify that $\bar{b}(\theta)$ is an equilibrium. To do this, we need to check:**

    ▷ (1) The bid should be in the set; $b_i \in \left[0, \bar{b}(1)\right]$

    ▷ (2) Since θ uniquely pins down b i.e. $\exists!\theta$ such that $b_i = \bar{b}(\theta)$, it suffices to show that truth-reporting is optimal given that others are truth-reporting. In other words, defining:

    $$U\left(\hat{\theta}|\theta\right) \equiv \max_{\theta} G\left(\hat{\theta}\right)\left(\theta_i - \bar{b}\left(\hat{\theta}\right)\right)$$

    then $U(\theta|\theta) \geq U\left(\hat{\theta}|\theta\right)$. To show (2), recall that we arrived at the following:

    $$\bar{b}(\theta) = \theta - \int_0^\theta \left(\frac{G(x)}{G(\theta)}\right)dx = \theta - \int_0^\theta \left(\frac{F(x)}{F(\theta)}\right)^{n-1} dx$$

So we can write:

$$
\begin{aligned}
U\left(\hat{\theta}|\theta\right) &= G\left(\hat{\theta}\right)\theta - G\left(\hat{\theta}\right)\bar{b}\left(\hat{\theta}\right) \\
&= G\left(\hat{\theta}\right)\theta - \hat{\theta}G\left(\hat{\theta}\right) + \int_0^{\hat{\theta}} G\left(x\right)dx \\
&= G\left(\hat{\theta}\right)\left(\theta - \hat{\theta}\right) + \int_0^{\hat{\theta}} G\left(x\right)dx
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
U\left(\theta|\theta\right) - U\left(\hat{\theta}|\theta\right) &= \int_0^{\theta} G\left(x\right)dx - G\left(\hat{\theta}\right)\left(\theta - \hat{\theta}\right) - \int_0^{\hat{\theta}} G\left(x\right)dx \\
&= -G\left(\hat{\theta}\right)\left(\theta - \hat{\theta}\right) + \int_{\hat{\theta}}^{\theta} G\left(x\right)dx \\
&= \int_{\hat{\theta}}^{\theta}\left[G\left(x\right) - G\left(\hat{\theta}\right)\right]dx \geq 0
\end{aligned}
$$

and thus the verification is complete.

3. **Expected Revenue**

The expected revenue for player θ_i is defined as:

$$
R_i\left(\theta_i\right) = G\left(\theta_i\right)\bar{b}\left(\theta_i\right)
$$

Recall that [1] gives us:

$$
\begin{aligned}
\bar{b}\left(\theta\right) &= \frac{1}{G\left(\theta\right)}\left(\int_{\underline{\theta}=0}^{\theta} xg\left(x\right)dx\right) \\
\Rightarrow \bar{b}\left(\theta_i\right) &= \mathbb{E}\left[\max_{j\neq i}\theta_j | \max_{j\neq i}\theta_j \leq \theta_i\right]
\end{aligned}
$$

Therefore:

$$
\begin{aligned}
R_i\left(\theta_i\right) &= G\left(\theta_i\right)\mathbb{E}\left[\max_{j\neq i}\theta_j | \max_{j\neq i}\theta_j \leq \theta_i\right] \\
&= \int_0^{\theta_i} xg\left(x\right)dx
\end{aligned}
$$

The total expected revenue is then

$$
\begin{aligned}
\mathbb{E}\left[\sum_i R_i\left(\theta_i\right)\right] &= n\cdot\mathbb{E}\left[\int_0^{\theta} xg\left(x\right)dx\right] \\
&= n\int_0^1\left(\int_0^{\theta} xg\left(x\right)dx\right)f\left(\theta\right)d\theta
\end{aligned}
$$

Integration by parts:

$$
\begin{aligned}
=& n\left[\left(\int_0^{\theta} xg\left(x\right)dx\right)\left(F\left(\theta\right)-1\right)|_0^1 - \int_0^1 \theta g\left(\theta\right)\left(F\left(\theta\right)-1\right)d\theta\right] \\
=& n\int_0^1 \theta g\left(\theta\right)\left(F\left(\theta\right)-1\right)d\theta
\end{aligned}
$$

Since $g(\theta) = (n-1) F(\theta)^{n-2} f(\theta)$:

$$= n \int_0^1 \theta \underbrace{n(n-1) f(\theta) F(\theta)^{n-2} (1 - F(\theta))}_{[A]} d\theta$$

It turns out that $[A]$ is the density of the 2nd-order statistic.

▷ This will be different from the Wikipedia since it lists the 2nd lowest order statistic.

This tells us that

$$\mathbb{E}\left[\sum_i R_i(\theta_i)\right] = \mathbb{E}[\theta | \theta = \text{ 2nd highest type}]$$

B. Envelope Theorem Approach Recall that bidder i's expected utility is:

$$\begin{aligned}&[\text{Probability of } \theta_i \text{ Winning }](\theta_i - b_i)\\&= G\left(\bar{b}^{-1}(b_i)\right)(\theta_i - b_i)\end{aligned}$$

In equilibrium, $b_i = \bar{b}(\theta_i)$, so bidder i's expected utility is then

$$G(\theta_i)\left(\theta_i - \bar{b}(\theta_i)\right)$$

From the Envelope theorem, we know this has to equal $\int_0^\theta G(s)\,ds$ so we have:

$$G(\theta_i)\left(\theta_i - \bar{b}(\theta_i)\right) = \int_0^\theta G(s)\,ds$$

Solving for $\bar{b}$ yields:

$$\bar{b}(\theta_i) = \theta_i - \int_0^{\theta_i} \frac{G(x)}{G(\theta_i)} dx$$

Note that this technique isn't always helpful. It works for the all-pay auction but not in general.

Intuition We see that each bidder bids the expectation of the second highest bidder's value conditional on his own value being the highest. This is because $F^{N-1}(\cdot)$ is the distribution function of the highest value among a bidder's $N-1$ competitors, so

$$\hat{b}(v) = \frac{1}{F^{N-1}(v)} \int_0^v x dF^{n-1}(x)$$

yields the intuition.

7.1.3 First-price = Clock (Dutch) Auction

In a Dutch auction, each bidder has a single decision to make: at what price should I raise my hand to signal that I am willing to buy the good at that price? Moreover, the bidder who chooses the highest price wins the auction and pays the price. Replacing the word "price" by "bid," we see that this auction is equivalent to a first-price auction. Thus, we have the conclusion that it yields the same equilibrium as the first-price auction.

7.1.4 Second-price, Sealed Bid

Ex ante, we expect the bidders to bid more aggressively in a second-price auction than they would in a first-price auction. We will **claim** that it is a dominant strategy for player i to bid $b_i = \theta_i$. In fact, if N bidders have independent private values, then bidding one's value is the unique weakly dominant bidding strategy for each bidder in a second-price, sealed-bid auction.

Proof. Define $b_i^* = \max_{j \neq i} b_j$ i.e. the most competitive outside bid.

- ▷ Suppose your type $\theta_i > b_i^*$.
  - * If you bid $b_i = \theta_i$ then you win and earn $\theta_i - b_i^* > 0$. In fact, any $b_i > b_i^*$ will lead to the same return.
- ▷ Suppose your type $\theta_i \leq b_i^*$.
  - * If you bid $b_i^* > b_i > \theta_i$, then you lose and payoff is zero.
  - * If you bid $b_i > b_i^* > \theta_i$, then you win but payoff $\theta_i - b_i^* < 0$.

Note that in this case, the expected revenue is then $\mathbb{E}\left[\theta | \theta = \text{ 2nd highest type}\right]$. ■

7.1.5 Second-price = English Auction

As in the second-price auction, it turns out to be a dominant strategy for a bidder to drop out when the price reaches his value, regardless of which bidders remain active.

7.2 Revenue Equivalence Theorem

To explain the equivalence of revenue in the four standard auctions, we must find a way to fit all of these auctions into a single framework.

7.2.1 Direct Selling Mechanism (DSM)

A direct selling mechanism (DSM) is a collection of N probability assignment functions:

$$p_1\left(v_1, ..., v_N\right), ..., p_N\left(v_1, ..., v_N\right)$$

and N cost functions:

$$c_1\left(v_1, ..., v_N\right), ..., c_N\left(v_1, ..., v_N\right)$$

where p_i denotes the probability of winning for agent i. Essentially, the seller asks the bidders to report their values (which need not be truthful), and he assigns the object to one of the bidders according to $p_i\left(\boldsymbol{v}\right)$ and secures the payment $c_i\left(\boldsymbol{v}\right)$ from each bidder.

- ▷ The sum of the probabilities need not equal unity.

7.2.2 Incentive-Compatible Direct Selling Mechanism (IC-DSM)

As a refinement, we introduce an **incentive-compatible direct selling mechanism (IC-DSM)**. Specifically, a DSM is IC if when other bidders always report their values truthfully and each bidder's expected payoff is maximized by always reporting the value truthfully. It only says that a bidder can do no better than to report truthfully so long as all other bidders report truthfully. Given this framework, it is easy to see that the four standard auctions can be equivalently viewed as an IC-DSM.

▷ First-price auction with symmetric bidders

Instead of the bidders submitting bids computed by plugging their values into the equilibrium bidding function, the bidders will be asked to submit their values and the seller will then compute the equilibrium bids for them. Thus the corresponding IC-DSM is

$$p_i(\boldsymbol{v}) = \begin{cases} 1 & \text{if } v_i > v_j, \forall j \neq i \\ 0 & \text{otherwise} \end{cases}$$
$$c_i(\boldsymbol{v}) = \begin{cases} \hat{b}(v_i) & \text{if } v_i > v_j, \forall j \neq i \\ 0 & \text{otherwise} \end{cases}$$

To show that this incentive-compatible, we need to show that truth-telling is a Nash equilibrium. To see this, note that the bidder's expected payoff from reporting value r when his true value is v is

$$F^{N-1}(r)\left(v - \hat{b}(r)\right)$$

But since $\hat{b}$ is an equilibrium bidding function, it is maximized at $r = v$.

7.2.3 Revenue Equivalence Theorem (Light)

The first-price and Dutch auctions raise the same ex-post revenue, and the second-price and English auctions raise the same ex-post revenue. Because our analysis of the first-price auction involved symmetric bidders, we must assume symmetry here to compare the expected revenue generated by a first-price versus a second-price auction. We showed earlier that the revenues of the two types of auctions are actually the same.

Theorem 7.1. *Assume $F_i = F_j, \forall i, j$. If two auctions both (1) awarded the good to the highest-type bidder and (2) yields zero payoff to the lowest-type bidder, then the expected revenue is the same. Since all four previously described auctions above share this characteristic, they have the same revenue.*

Proof. Define the expected utility of agent i's participation with type θ_i:

$$U_i(\theta_i) = \max_{b_i} \mathbb{E}_{\theta_{-i}}\left[\phi_i\left(b_i, \bar{b}_{-i}(\theta_{-i})\right)\right]\theta_i - \mathbb{E}_{\theta_{-i}}\left[t_i\left(b_i, \bar{b}_{-i}(\theta_{-i})\right)\right]$$

By the Envelope theorem which implies that $\partial U_i(\theta_i)/\partial b_i(\theta_i) = 0$, we know that

$$U_i'(\theta_i) = \mathbb{E}_{\theta_{-i}}\left[\phi_i\left(b_i, \bar{b}_{-i}(\theta_{-i})\right)\right] = G(\theta_i)$$

where $G(\theta_i)$ is the probability of agent i winning the good. It is equal to $G(\theta_i)$ since in equilibrium, $b_i = \bar{b}(\theta_i)$. Using this result:

$$\begin{aligned} U_i(\theta_i) &= C + \int_0^{\theta_i} U_i'(x)\,dx \\ &= C + \int_0^{\theta_i} G(\theta_i)\,dx \end{aligned}$$

Pin down the constant using $0 = U_i(0)$ yields:

$$U_i(\theta_i) = \int_0^{\theta_i} F(s)^{n-1}\,ds$$

Now compute the expected revenue of the auctioneer, which is the expected valuation of the person who gets the good minus the total expected surplus of all agents:

$$
\begin{aligned}
&\mathbb{E}\left[\max\left\{\theta_1, ..., \theta_n\right\} - \sum_{i=1}^{n} U_i\left(\theta_i\right)\right] \\
=&\mathbb{E}\left[\max\left\{\theta_1, ..., \theta_n\right\} - \sum_{i=1}^{n} \int_0^{\theta_i} G\left(x\right) dx\right] \\
=&\mathbb{E}\left[\max\left\{\theta_1, ..., \theta_n\right\} - \sum_{i=1}^{n} \int_0^{\theta_i} F\left(x\right)^{N-1} dx\right]
\end{aligned}
$$

which is only about the characteristics of the agents and does not involve any auction-specific terms ϕ or t. ■

As we saw above, in the first-price auction, we have:

$$U_i\left(\theta_i\right) = G\left(\theta_i\right)\left(\theta_i - \bar{b}\left(\theta_i\right)\right) = \int_0^{\theta_i} G\left(s\right) ds$$

where

$$\bar{b}_i\left(\theta_i\right) = \theta_i - \int_0^{\theta_i} \frac{G\left(s\right)}{G\left(\theta_i\right)} ds$$

Remarks:

- ▷ Suppose you have one guy FOSD-ing the other guy. In the second-price auction, bidding the truth would still be optimal; in the first-price auction, weird things will happen. Generally, you will find that the weak bidders will sometimes win. In this case, $\phi_i^I\left(\theta\right) \neq \phi_i^{II}\left(\theta\right), \forall\theta, \forall i$.
- ▷ We can extend the revenue equivalence theorem to more complicated auctions.
- ▷ Suppose you have a symmetric distribution of bidders such that we can apply the light version of the revenue equivalence theorem. The distribution of the bids WILL NOT be the same.

7.2.4 Revenue Equivalence Theorem (Full)

Define

$$\bar{\phi}_i\left(\theta_i\right) = \mathbb{E}_{\theta_{-i}}\left[\phi_i\left(\theta_i, \theta_{-i}\right)\right]$$

i.e. the expected probability that the good goes to person i when he/she reports θ_i, and:

$$\bar{t}_i\left(\theta_i\right) \equiv \mathbb{E}_{\theta_{-i}}\left[t_i\left(\theta_i, \theta_{-i}\right)\right]$$

which is the expected payment when he reports θ_i. This allows us to write bidder i's expected payoff when he reports $\hat{\theta}_i$ when all others report truthfully:

$$
\begin{aligned}
U_i\left(\theta_i\right) &\equiv U_i\left(\theta_i|\theta_i\right) = \bar{\phi}_i\left(\theta_i\right)\theta_i - \bar{t}_i\left(\theta_i\right) \\
U_i\left(\hat{\theta}_i\right) &\equiv U_i\left(\hat{\theta}_i|\theta_i\right) = \bar{\phi}_i\left(\hat{\theta}_i\right)\theta_i - \bar{t}_i\left(\hat{\theta}_i\right)
\end{aligned}
$$

Theorem 7.2. $\left\{\phi_1, ..., \phi_n, t_1, ..., t_n\right\}$ *is incentive-compatible if and only if*

1. *$\forall i, \bar{\phi}_i(\theta_i)$ is increasing in θ_i .*
2. *$\forall i$, we have*

$$U_i(\theta_i) - U_i(\underline{\theta}_i) = \int_{\underline{\theta}_i}^{\theta_i} \bar{\phi}_i(s)\, ds$$

Proof. Mechanism is IC $\Rightarrow$ (1) and (2) holds.

▷ To get (1), recall that from incentive compatibility, we have

$$\bar{\phi}_i(\theta_i)\theta_i - \bar{t}_i(\theta_i) = U_i(\theta_i|\theta_i) \geq U_i\left(\hat{\theta}_i|\theta_i\right) = \bar{\phi}_i\left(\hat{\theta}_i\right)\theta_i - \bar{t}_i\left(\hat{\theta}_i\right)$$

Adding and subtracting $\bar{\phi}_i(\theta_i)\hat{\theta}_i$ on the LHS:

$$\bar{\phi}_i(\theta_i)\left(\theta_i - \hat{\theta}_i\right) + \underbrace{\bar{\phi}_i(\theta_i)\hat{\theta}_i - \bar{t}_i(\theta_i)}_{\leq U_i(\hat{\theta}_i) = \bar{\phi}_i(\hat{\theta}_i)\hat{\theta}_i - \bar{t}_i(\hat{\theta}_i)} \geq \bar{\phi}_i\left(\hat{\theta}_i\right)\theta_i - \bar{t}_i\left(\hat{\theta}_i\right)$$

which yields:

$$\bar{\phi}_i\left(\hat{\theta}_i\right)\hat{\theta}_i - \bar{t}_i\left(\hat{\theta}_i\right) + \bar{\phi}_i(\theta_i)\left(\theta_i - \hat{\theta}_i\right) \geq \bar{\phi}_i\left(\hat{\theta}_i\right)\theta_i - \bar{t}_i\left(\hat{\theta}_i\right)$$

which yields

$$\left[\bar{\phi}_i(\theta_i) - \bar{\phi}_i\left(\hat{\theta}_i\right)\right]\left(\theta_i - \hat{\theta}_i\right) \geq 0$$

▷ To get (2), use the Envelope Theorem since U_i is absolute continuous and is defined as

$$U_i\left(\hat{\theta}_i|\theta_i\right) \equiv \bar{\phi}_i\left(\hat{\theta}_i\right)\theta_i - \bar{t}_i\left(\hat{\theta}_i\right)$$

so the derivative is given as

$$\begin{aligned} U_i'(\theta_i) \equiv U_i'(\theta_i|\theta_i) &= U_1(\theta_i|\theta_i) + U_2(\theta_i|\theta_i) \\ &= \bar{\phi}_i\left(\hat{\theta}_i\right) + \left[\bar{\phi}_i'\left(\hat{\theta}_i\right)\theta_i - \bar{t}_i'\left(\hat{\theta}_i\right)\right]_{\hat{\theta}_i = \theta_i} \\ &= \bar{\phi}_i\left(\hat{\theta}_i\right) \end{aligned}$$

and thus the proof is complete.

(1) and (2) holds $\Rightarrow$ Mechanism is IC.

▷ Mechanism being IC is saying that

$$\begin{aligned} U_i(\theta_i) &\geq U_i\left(\hat{\theta}_i|\theta_i\right) \\ &= \bar{\phi}_i\left(\hat{\theta}_i\right)\theta_i - \bar{t}_i\left(\hat{\theta}_i\right) \\ &= \bar{\phi}_i\left(\hat{\theta}_i\right)\left(\theta_i - \hat{\theta}_i\right) - U_i\left(\hat{\theta}_i\right) \end{aligned}$$

or;

$$U_i(\theta_i) - U_i\left(\hat{\theta}_i\right) \geq \bar{\phi}_i\left(\hat{\theta}_i\right)\left(\theta_i - \hat{\theta}_i\right), \quad \forall \hat{\theta}_i, \theta_i$$

This is equivalent to (2)

$$\int_{\hat{\theta}_i}^{\theta_i} \bar{\phi}_i(s)\, ds \geq \int_{\hat{\theta}_i}^{\theta_i} \bar{\phi}_i\left(\hat{\theta}_i\right) ds$$

which is true only if (1) is true.

■

Given the result above it has to be that in equilibrium, when everyone reports truthfully, bidder i's expected payoff can be written as:

$$U_i(\theta_i) = \bar{\phi}_i(\theta_i)\,\theta_i - \bar{t}_i(\theta_i)$$

where

$$\bar{\phi}_i(\theta_i) \equiv \mathbb{E}_{\theta_{-i}}\left[\phi_i(\theta_i, \theta_{-i})\right]$$
$$\bar{t}_i(\theta_i) \equiv \mathbb{E}_{\theta_{-i}}\left[t_i(\theta_i, \theta_{-i})\right]$$

We also just showed that

$$U_i(\theta_i) - U_i(\underline{\theta}_i) = \int_{\underline{\theta}_i}^{\theta_i} \bar{\phi}_i(s)\,ds$$

So plugging in, we obtain:

$$\underline{\bar{t}_i(\theta_i) = \bar{\phi}_i(\theta_i)\,\theta_i - U_i(\underline{\theta}_i) - \int_{\underline{\theta}_i}^{\theta_i} \bar{\phi}_i(s)\,ds}$$

i.e. expected payment i makes given that the type is θ_i. It depends on the lowest possible payment and ϕ. Consequently, the expected revenue for the full game is

$$\mathbb{E}\left[\sum_{i=1}^{n} \bar{t}_i(\theta_i)\right]$$

With the preceding results, we arrive at the following theorem:

Theorem 7.3. *Revenue Equivalence Theorem (Full): Two incentive-compatible DRM with the same probability assignment functions $\phi_i^I(\theta) = \phi_i^{II}(\theta), \forall\theta, \forall i$ and every bidder with value zero is indifferent between the two mechanisms*

$$U_i^I(\underline{\theta}_i) = U_i^{II}(\underline{\theta}_{-i}), \forall\theta, \forall i$$

then the expected revenues are the same.

Proof. Write out explicitly the seller's expected revenue:

$$\begin{aligned}
R &= \mathbb{E}\left[\sum_{i=1}^{n} \bar{t}_i(\theta_i)\right] \\
&= \sum_{i=1}^{n} \mathbb{E}\left[\bar{t}_i(\theta_i)\right] \\
(\text{expression for } \bar{t}_i(\theta_i)) &= \sum_{i=1}^{n} \mathbb{E}\left[\mathbb{E}_{\theta_{-i}}\left[t_i(\theta_i, \theta_{-i})\right]\right] \\
&= \sum_{i=1}^{n} \mathbb{E}_{\theta_i}\left[\bar{t}_i(\theta_i)\right] \\
(\text{expression for } \bar{t}_i(\theta_i)) &= \sum_{i=1}^{N} \mathbb{E}_{\theta_i}\left[\bar{\phi}_i(\theta_i)\,\theta_i - U_i(\underline{\theta}_i) - \int_{\underline{\theta}_i}^{\theta_i} \bar{\phi}_i(s)\,ds\right] \\
&= \sum_{i=1}^{N} \mathbb{E}_{\theta_i}\left[\bar{\phi}_i(\theta_i)\,\theta_i - \int_{\underline{\theta}_i}^{\theta_i} \bar{\phi}_i(s)\,ds\right] - \sum_{i=1}^{N} U_i(\underline{\theta}_i)
\end{aligned}$$

Consequently, the seller's expected revenue depends only on the probability assignment functions and the amount bidders expect to pay when their values are zero. This result is very general and it allows us to add additional auctions to the list of those yielding the same expected revenue as the four standard ones. For example, a first-price all-pay auction, in which the highest among all sealed bids wins but *every* bidder pays an amount equal to his bid, yields the same expected revenue. ■

7.2.5 Efficiency

Each of these auctions allocates the object to the bidder who values it the most – each of these auctions is efficient.

7.3 Design of Revenue-Maximizing Mechanisms

We saw that the four standard auctions generate the same expected revenue (under bidder symmetry), but do they maximize the seller's expected revenue? Or is it possible that there is a better selling mechanism for the seller?

7.3.1 Setup

Let $\theta_i \sim F_i(\theta)$ with $\left[\underline{\theta}_i, \bar{\theta}_i\right]$ and $i = 1, ..., n$. The utility is $u_i(\phi_i, \theta_i) = \phi_i \theta_i$. The auction (indirect) mechanism is defined as

$$\Gamma = \{\phi_1(\mathbf{s}), ..., \phi_n(\mathbf{s}), t_1(\mathbf{s}), ..., t_n(\mathbf{s}), S_1, ..., S_n\}$$

where $\phi_i(\mathbf{s})$ is the probability that the good goes to agent i given that vector $\mathbf{s}$ is played.

7.3.2 Revelation Principle

Given an arbitrary selling procedure and a Nash equilibrium in which each bidder employs a strategy mapping his value into payoff-maximizing behavior under the selling procedure, we can construct an equivalent incentive-compatible direct selling mechanism. Consequently, this will yield the same expected revenue for the seller. This implies that we can restrict our **search for a revenue-maximizing selling procedure to a set of incentive-compatible direct selling mechanisms.** This technique is an instance of the revelation principle.

Proposition 7.1. *For any* $(s_1^*, ..., s_n^*)$ *BNE of* Γ *defined above, there exists a direct mechanism* $S_i = \Theta, \tilde{\Gamma} = \left\{\tilde{\phi}, \tilde{t}, \Theta\right\}$ *and an equilibrium* $\{\tilde{s}_1^*, ..., \tilde{s}_n^*\}$ *such that*

1. $\tilde{s}_i^*(\theta_i) = \theta_i, \forall i$
2. $\phi_i(\mathbf{s}^*(\boldsymbol{\theta})) = \tilde{\phi}_i(\boldsymbol{\theta}), \forall \boldsymbol{\theta}, \forall i$
3. $t_i(\mathbf{s}^*(\boldsymbol{\theta})) = \tilde{t}_i(\boldsymbol{\theta}), \forall \boldsymbol{\theta}, \forall i$

Proof. Define $\tilde{\phi}_i(\boldsymbol{\theta}) \equiv \phi_i(s_1^*(\theta_1), ..., s_n^*(\theta_n))$ and $\tilde{t}_i(\boldsymbol{\theta}) \equiv t_i(s_1^*(\theta_1), ..., s_n^*(\theta_n))$. Given this construction, it suffices to show that truth-telling is optimal given that everyone else is telling the truth.

▷ Notice that in the original game, player i can choose $s_i \in S_i$. In the new game, the person has to choose $s_i \in s_i^*(\Theta_i)$ i.e. take all possible bids that I might have set in the original mechanism, and this is the only set of strategies I can choose. Thus:

$$s_i^*(\Theta_i) \leq S_i$$

i.e. I have less options under the direct mechanism.

- ▷ In the original game, the BNE of Γ for $s_1^*, \ldots, s_n^*$ implies that

$$\begin{aligned}&\mathbb{E}_{\theta_{-i}}\left[\phi_i\left(s_i^*\left(\theta_i\right), s_{-i}^*\left(\theta_{-i}\right)\right)\right]-\mathbb{E}_{\theta_{-i}}\left[t_i\left(s_i^*\left(\theta_i\right), s_{-i}^*\left(\theta_{-i}\right)\right)\right]\\ \geq &\mathbb{E}_{\theta_{-i}}\left[\phi_i\left(s_i, s_{-i}^*\left(\theta_{-i}\right)\right)\right]-\mathbb{E}_{\theta_{-i}}\left[t_i\left(s_i, s_{-i}^*\left(\theta_{-i}\right)\right)\right]\end{aligned}$$

  for all i, θ_i and $\forall s_i \in S_i$.

- ▷ Using the fact that $s_i^*\left(\Theta_i\right) \leq S_i$, we have $\tilde{s}_i^*\left(\theta_i\right)=\theta_i$.

■

7.3.3 Individual Rationality

We also need to consider an additional restriction. Because participation by the bidders is entirely voluntary, no bidder's expected payoff can be negative given his value:

$$U\left(\theta_i\right)=\bar{\phi}_i\left(\theta_i\right)\theta_i-\bar{t}_i\left(\theta_i\right) \geq 0, \forall \theta_i$$

From the previous characterization of incentive compatibility, we also have:

$$\bar{t}_i\left(\theta_i\right)=\bar{\phi}_i\left(\theta_i\right)\theta_i-U_i\left(\underline{\theta}_i\right)-\int_{\underline{\theta}_i}^{\theta_i} \bar{\phi}_i(s)\, ds$$

Combining the two equations yields:

$$U\left(\theta_i\right)=U_i\left(\underline{\theta}_i\right)+\int_{\underline{\theta}_i}^{\theta_i} \bar{\phi}_i(s)\, ds \geq 0$$

which holds if $U_i\left(\underline{\theta}_i\right) \geq 0$

7.3.4 Characterizing the Optimal Selling Mechanism

We have now reduced the task of finding an optimal selling mechanism to maximizing the seller's expected revenue among all individually rational, incentive-compatible DSMs. Given our previous results, our task has now been reduced to solving the following problem:

$$\max_{\{\phi_i, t_i\}_{i=1}^N} \mathbb{E}_\theta\left[\sum_{i=1}^n \bar{t}_i\left(\theta_i\right)\right]=\mathbb{E}_\theta\left[\underbrace{\sum_{i=1}^n \phi_i(\theta)\left(\theta_i-\theta_0\right)+\theta_0-}_{\text{total surplus}} \underbrace{\sum_{i=1}^n U_i\left(\theta_i\right)}_{\text{consumer surplus}}\right] \text{ subject to } [IC],[IR]$$

where θ_0 is the opportunity cost of the seller. Recall the familiar step-by-step approach:

1. Replace the $[IC]$ constraint with
   (a) $\forall i, \bar{\phi}_i\left(\theta_i\right)$ is increasing in θ_i .
   (b) $\forall i$, we have

$$U_i\left(\theta_i\right)-U_i\left(\underline{\theta}_i\right)=\int_{\underline{\theta}_i}^{\theta_i} \bar{\phi}_i(s)\, ds$$

2. Write expected utility as the following:

$$\begin{aligned}\mathbb{E}\left[U_i\left(\theta_i\right)\right] &= U_i\left(\underline{\theta}_i\right) + \mathbb{E}\left[\int_{\underline{\theta}_i}^{\theta_i} \bar{\phi}_i\left(s\right) ds\right] \\ &= U_i\left(\underline{\theta}_i\right) + \int_{\underline{\theta}_i}^{\bar{\theta}_i} \underbrace{\left(\int_{\underline{\theta}_i}^{\theta_i} \bar{\phi}_i\left(s\right) ds\right)}_{u} \underbrace{f_i\left(\theta_i\right) d\theta_i}_{dv}\end{aligned}$$

Using integration by parts:

$$\begin{aligned}&= U_i\left(\underline{\theta}_i\right) + \left[\left(\int_{\underline{\theta}_i}^{\theta_i} \bar{\phi}_i\left(s\right) ds\right)\left(F_i\left(\theta_i\right) + K\right)\right]_{\underline{\theta}_i}^{\bar{\theta}_i} - \int_{\underline{\theta}_i}^{\bar{\theta}_i} \left[F_i\left(\theta_i\right) + K\right] \bar{\phi}_i\left(\theta\right) d\theta_i \\ &= U_i\left(\underline{\theta}_i\right) + \left(\int_{\underline{\theta}_i}^{\bar{\theta}_i} \bar{\phi}_i\left(s\right) ds\right)\left(1 + K\right) - \int_{\underline{\theta}_i}^{\bar{\theta}_i} \left[F_i\left(\theta_i\right)\right] \bar{\phi}_i\left(\theta\right) d\theta_i - \int_{\underline{\theta}_i}^{\bar{\theta}_i} K\bar{\phi}_i\left(\theta\right) d\theta_i \\ &= U_i\left(\underline{\theta}_i\right) + \left(\int_{\underline{\theta}_i}^{\bar{\theta}_i} \bar{\phi}_i\left(s\right) ds\right) - \int_{\underline{\theta}_i}^{\bar{\theta}_i} \left[F_i\left(\theta_i\right)\right] \bar{\phi}_i\left(\theta\right) d\theta_i\end{aligned}$$

Rearranging, we have

$$\begin{aligned}\mathbb{E}_\theta\left[U_i\left(\theta_i\right)\right] &= U_i\left(\underline{\theta}_i\right) + \int_{\underline{\theta}_i}^{\bar{\theta}_i} \bar{\phi}_i\left(\theta_i\right) \frac{1 - F_i\left(\theta_i\right)}{f_i\left(\theta_i\right)} f_i\left(\theta_i\right) d\theta_i \\ &= U_i\left(\underline{\theta}_i\right) + \mathbb{E}_{\theta_i}\left[\bar{\phi}_i\left(\theta_i\right) \frac{1 - F_i\left(\theta_i\right)}{f_i\left(\theta_i\right)}\right] \\ \left(\because\ \bar{\phi}_i\left(\theta_i\right) \text{ is already an expectation}\right) &= U_i\left(\underline{\theta}_i\right) + \mathbb{E}_\theta\left[\phi_i\left(\theta\right) \frac{1 - F_i\left(\theta_i\right)}{f_i\left(\theta_i\right)}\right]\end{aligned}$$

3. Rewrite the problem:

$$\begin{aligned}&\max_{\phi, U(\underline{\theta})} \mathbb{E}\left[\sum_{i=1}^n \phi_i\left(\theta\right)\left(\theta_i - \theta_0\right) + \theta_0 - \sum_{i=1}^n U_i\left(\underline{\theta}_i\right) - \sum_{i=1}^n \phi_i\left(\theta\right)\left(\frac{1 - F_i\left(\theta_i\right)}{f_i\left(\theta_i\right)}\right)\right] \\ &\text{s.t. } \bar{\phi}_i \uparrow\end{aligned}$$

Rewriting:

$$\begin{aligned}&\max_{\phi, U(\underline{\theta})} \mathbb{E}\left[\sum_{i=1}^n \phi_i\left(\theta\right)\left(\underbrace{\theta_i - \frac{1 - F_i\left(\theta_i\right)}{f_i\left(\theta_i\right)}}_{J_i(\theta_i)} - \theta_0\right) - \sum_{i=1}^n U_i\left(\underline{\theta}_i\right)\right] \\ &\text{s.t. } \bar{\phi}_i \uparrow\end{aligned}$$

We can define

$$J_i\left(\theta_i\right) = \theta_i - \frac{1 - F_i\left(\theta_i\right)}{f_i\left(\theta_i\right)}$$

as the "virtual type" of bidder i where the extra term is the inverse hazard rate that represents the information rent to the bidder. The regularity assumption is that $J_i\left(\theta_i\right) \uparrow$.

This says that given the reported values $\theta_1, ..., \theta_N$, the object is given to the bidder i whose $J_i(\theta_i)$ is strictly highest and positive. Otherwise, the seller keeps the object. The following result summarizes this allocation rule:

Theorem 7.4. *(Myerson, 1981) Optimal auction satisfies:*

1. $U_i(\underline{\theta}_i) = 0, \forall i$
2. *We have:*

$$\phi_i(\theta) = \begin{cases} 1 & \text{if } J_i(\theta_i) > \max_{j \neq i} J_j(\theta_j) \text{ and } J_i(\theta_i) \geq \theta_0 \\ 0 & \text{if } J_i(\theta_i) < \max_{j \neq i} J_j(\theta_j) \text{ or } J_i(\theta_i) \leq \theta_0 \\ \frac{1}{k} & \text{if } J_i(\theta_i) = \max_{j \neq i} J_j(\theta_j) \geq \theta_0 \text{ with } k \text{ ties} \end{cases}$$

Note that the last case is a probability-zero event.

Remark 7.1. $MR_i(\theta_i) = J_i(\theta_i)$. This is an intuitive reason why "virtual type" pops up so frequently. This is the marginal revenue that the seller obtains from increasing the probability that the object is assigned to bidder i when his value is θ.

▷ Example: Suppose there is one bidder ($N = 1$). The revenue function is $R(p) = p(1 - F(p))$ and $D(p) = 1 - F(p)$.

  * We can say that

$$\frac{dq}{dp} = -f(p)$$

  * Explicitly writing the MR:

$$MR = \frac{dR(p)}{dp}\frac{dp}{dq} = (-pf(p) + (1 - F(p)))\frac{1}{-f(p)} = p - \frac{1 - F(p)}{f(p)}$$

  Since we are selling to one person, to profit maximize, we want to set it equal to marginal cost θ_0.

▷ Intuition: Consider the effect of increasing the probability that the the object is assigned to bidder i when his value is θ_i.

  * This enables the seller to increase the cost to θ_i so as to leave his utility unchanged, so the seller's revenue increases at the rate $\theta_i f_i(\theta_i)$ as a result of this change.
  * On the other hand, increasing the probability for θ_i implies that for the mass of $1 - F_i(\theta_i)$ values above θ_i, we have a reduction in revenue.
  * Therefore, altogether the seller's revenue increases by

$$\theta_i f_i(\theta_i) - (1 - F_i(\theta_i))$$

  which is due to the density $f_i(v_i)$ of values equal to v_i. Therefore, we have

$$\frac{\theta_i f_i(\theta_i) - (1 - F_i(\theta_i))}{f_i(\theta_i)} = \theta_i - \frac{1 - F_i(\theta_i)}{f_i(\theta_i)} \equiv J_i(\theta_i)$$

Remark 7.2. Everything so far can be generalized to more general forms of utility i.e. $u(q, \theta) \neq \theta q$. In fact, as long as we have quasi-linearity in money and single crossing ($u_{q\theta} > 0, u_\theta \geq 0$), the results follow through.

Remark 7.3. Risk aversion.

- ▷ Nothing so far will work under risk aversion i.e. non quasi-linear utilities. Because we know risk aversion is important, there is a body of literature that looks at the auction styles in the weird and writes theorems about which is better.
- ▷ One paper is about risk-averse bidders still bid their type in the second-price auction. In the first-price auction, you have this additional effect.
  - * On the margin, bidding aggressively ensures you against risk so you have this additional kick in aggressiveness.

Remark 7.4. Correlation in signals between first-price vs. second-price.

- ▷ Which is better if the signals are correlated with each other? The result is that the signals are correlated in a very strong way ("affiliated"), then you get that second price does better than the first price.
- ▷ The logic is called the "Linkage Principle." Anytime you incorporate more information than just the bidder's information, it's always better. First-price auction only uses one piece of information, so revenue is lower.

7.3.5 Efficiency

In the optimal selling mechanism, the object is not always allocated efficiently. There are two ways in which this inefficiency can occur.

1. Not enough sold i.e. seller keeps the good too often even though his value for it is zero and all bidders have positive value. This can happen when every bidder's value θ_i is such that $\theta_i \leq J_i(\theta_i)$.

$$r_i^* : J_i(r_i^*) = \theta_0 \Rightarrow r_i^* - \theta_0 = \frac{1 - F(r_1^*)}{f(r_1^*)} > 0$$

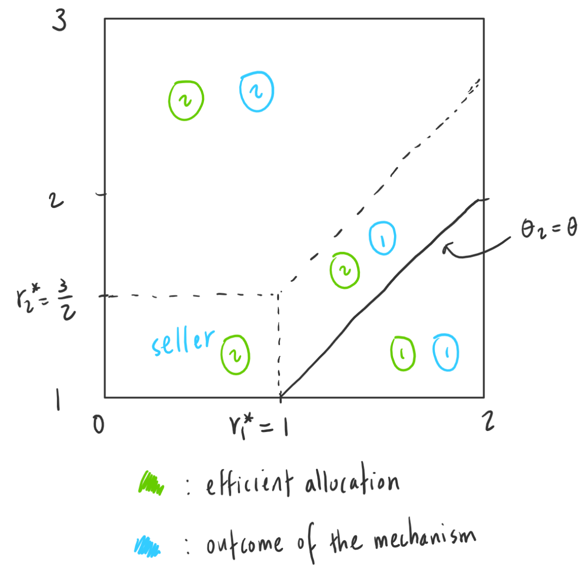
2. Wrong bidder gets good:

$$\theta_i > \max_{j \neq i} \theta_j \Leftrightarrow J_i(\theta_i) < \max_{j \neq i} J_j(\theta_j)$$

   To see this in asymmetric distributions, let $\theta_0 = 0, \theta_1 \sim U[0,2]$ and $\theta_2 \sim U[1,3]$. Then:

$$J_1(\theta_1) = 2\theta_1 - 2 = 2(\theta_1 - 1) \Rightarrow r_1^* = 1$$
$$J_2(\theta_2) = \theta_2 - (3 - \theta_2) = 2\theta_2 - 3 \Rightarrow r_2^* = 3/2$$

   where the r_i^* was found to be the largest value he could have reported, given the other bidder's reported values without winning the object. Setting $J_1(\theta_1) = J_2(\theta_2)$, we have $\theta_2 = 0.5 + \theta_1$. Graphically:



- ▷ θ_1 bids more aggressively, and θ_2 bids less aggressively.

The presence of inefficiency is not surprising – after all, the seller is a monopolist seeking maximal profits. The first source of inefficiency is the same as before – since there is only one unit of an indivisible object for sale, the seller here restricts supply by sometimes keeping the object. But there is also a **second** source of inefficiency. Here the seller can distinguish bidder i from bidder j and knows the distribution of values. This allows the monopolist to discriminate between the bidders, which leads to higher profits.

7.3.6 Connection to Standard Auctions.

Let's first re-visit the first-price and the second-price auctions. We are essentially eliminating the second source of inefficiency by imposing symmetry. In this case, the optimal selling mechanism is as follows: the bidder with highest $J_i(\theta_i)$ receives the object and pays the seller r_i^* the largest value he could have reported, given the other bidder's reported values without winning the object.

- ▷ Recall that we had $F_i(\theta) = F_j(\theta), \forall i.j$. This implies that $J_i(\theta) = J_j(\theta)$. From the regularity condition:

$$J_i(\theta_i) > J_j(\theta_j) \Leftrightarrow \theta_i > \theta_j$$

which yields:

$$\phi(\theta) = \begin{cases} 1 & \text{if } \theta_i > \max_{j \neq i} \theta_j \text{ and } \theta_i \geq r^* \\ 0 & \text{if } \theta_i < \max_{j \neq i} \theta_j \text{ and } \theta_i < r^* \end{cases}$$

where $J(r^*) = \theta_0$.

- ▷ Is the first-price auction optimal? It is optimal under symmetric distribution. We can set the reserve price to be equal to r^*.

In sum, the four standard auctions – first-price, second-price, Dutch, and English – all yield the same revenue under symmetry. Moreover, by supplementing each by an approriate reserve price, the seller maximizes his expected revenue.

7.3.7 Extensions

We can consider the types are common values with conditionally independent signals.

1. **Myerson's Extension**:

$$v_i(\theta) = \theta_i + \sum_{j \neq i} e_j(\theta_j), \quad \mathbb{E}[e_j(\theta_j)] = 0$$

$$v_0(\theta) = \theta_0 + \sum_{j=1}^{n} e_j(\theta_j)$$

Repeating the same procedure as before, we obtain the same virtual types:

$$J_i(\theta_i) = \theta_i - \frac{1 - F_i(\theta_i)}{f(\theta_i)} - e_i(\theta_i)$$

2. **Mineral–Rights Model**

We talked about a setup in which

$$v(\theta) = \frac{1}{n} \sum_{j=1}^{n} \theta_j$$

and $\theta_i = x + \epsilon_i, \quad \epsilon_i \sim N\left(0, \sigma_\epsilon^2\right)$ where x is commonly observed by all bidders. The moment your signal is correlated, you can take all the surplus in the bidders. One is Crember-Mclean (1988) mechanism.

7.4 Auctions vs. Negotiations

This section is based on the paper "Auctions vs. Negotiations" by Buloew-Klenadder in 1996, AER. The setup is once again an IPV environment.

7.4.1 Setup

The model consists of two ingredients.

1. Symmetric distributions: $\theta_i \sim F$ on $\left[\underline{\theta}, \overline{\theta}\right]$ with previously defined

$$MR(\theta_i) \equiv J(\theta_i) = \theta_i - \frac{1 - F(\theta_i)}{f(\theta_i)}$$

2. $\underline{\theta} \geq \theta_0$.

Then the expected revenue of an optimal auction with n bidders (ER_n^*) can be defined as

$$\begin{aligned} ER_n^* &\equiv \max_{\phi} \mathbb{E}\left[\sum_{i=1}^{n} \phi_i(\boldsymbol{\theta})\left(MR(\theta_i) - \theta_0\right) + \theta_0\right] \\ &= \mathbb{E}\left[\max\left\{MR(\theta_1), ..., MR(\theta_n), \theta_0\right\}\right] \end{aligned}$$

Also define ER_n^{SPA} as the expected revenue in the 2nd price auction without a reserve price. If there is reserve, we know that $ER_n^{SPA} = ER_n^*$.

7.4.2 Results

Theorem 7.5. $ER_{n+1}^{SPA} \geq ER_n^*$ *i.e. the value of the additional bidder is at least good as getting the reserve price right. In other words, suppose you are trying to find a buyer. Just make sure two people show up and let them bid against each other. Then you don't have to worry about f or F.*

Proof. Writing out the expression for ER_{n+1}^{SPA} :

$$ER_{n+1}^{SPA} = \mathbb{E}\left[\max\left\{MR\left(\theta_1\right), ..., MR\left(\theta_n\right), MR\left(\theta_{n+1}\right)\right\}\right]$$

Notice that θ_0 is NOT floating around. By Jensen's inequality:

$$> \mathbb{E}\left[\max\left\{MR\left(\theta_1\right), ..., MR\left(\theta_n\right), \mathbb{E}\left[MR\left(\theta_{n+1}\right)\right]\right\}\right]$$

Lemma 7.1. $\mathbb{E}\left[MR\left(\theta_i\right)\right] = \underline{\theta}$.

Proof. Since

$$\begin{aligned}\mathbb{E}\left[MR\left(\theta_i\right)\right] &= \int_{\underline{\theta}}^{\bar{\theta}} \left(\theta_i - \frac{1-F\left(\theta_i\right)}{f\left(\theta_i\right)}\right) f\left(\theta_i\right) d\theta_i \\ &= \int_{\underline{\theta}}^{\bar{\theta}} \left\{\underbrace{\theta_i f\left(\theta_i\right) - \left(1 - F\left(\theta_i\right)\right)}_{=-\frac{d}{d\theta_i}\theta_i\left(1-F\left(\theta_i\right)\right)}\right\} d\theta_i = \underline{\theta}\end{aligned}$$

■

■

Using this lemma, we then have

$$\begin{aligned}&= \mathbb{E}\left[\max\left\{MR\left(\theta_1\right), ..., MR\left(\theta_n\right), \underline{\theta}\right\}\right] \\ &\geq \mathbb{E}\left[\max\left\{MR\left(\theta_1\right), ..., MR\left(\theta_n\right), \theta_0\right\}\right] \quad \left(\because \underline{\theta} \geq \theta_0\right)\end{aligned}$$

Note that this theorem breaks if we no longer have symmetric distributions.

Theorem 7.6. If we have one data point $\tilde{\theta} \sim F$, then we can guarantee a profit of half the reserve price.

Proof. Set the reserve price to be $r = \tilde{\theta}$. Then we have that

$$ER = \left(\frac{1}{2}\right) ER_2^{SPA} \geq \left(\frac{1}{2}\right) ER_1^*$$

■

8 Bilateral Trade

8.1 The General Problem

Now consider a trade between a buyer and a seller, instead of the setting with one seller and multiple buyers.

8.1.1 Setup

The buyer's value is

$$\theta_B \sim F_B \text{ on } \left[\underline{\theta}_B, \bar{\theta}_B\right]$$

and the seller's value is

$$\theta_S \sim F_S \text{ on } \left[\underline{\theta}_S, \bar{\theta}_S\right]$$

Denote $\phi\left(\hat{\theta}_B, \hat{\theta}_S\right)$ as the probability that the buyer gets the good, and $t\left(\hat{\theta}_B, \hat{\theta}_S\right)$ is the associated payment from the buyer to the seller. We also assume **ex post budget balancing** i.e. if we add up all the transfer payments, they should equal zero. This important because later when we talk abut socially efficient choices, it requires some injection of money to get to the socially optimal choice.

$$\begin{aligned}
\bar{\theta}_B\left(\hat{\theta}_B\right) &\equiv \mathbb{E}_{\theta_S}\left[\phi\left(\hat{\theta}_B, \theta_S\right)\right] \\
\bar{\theta}_S\left(\hat{\theta}_S\right) &\equiv \mathbb{E}_{\theta_B}\left[\phi\left(\theta_B, \hat{\theta}_S\right)\right] \\
\bar{t}_B\left(\hat{\theta}_B\right) &\equiv \mathbb{E}_{\theta_S}\left[t\left(\hat{\theta}_B, \theta_S\right)\right] \\
\bar{t}_s\left(\hat{\theta}_S\right) &\equiv \mathbb{E}_{\theta_B}\left[t\left(\theta_B, \hat{\theta}_S\right)\right]
\end{aligned}$$

Furthermore, denote

$$\begin{aligned}
U_B\left(\hat{\theta}_B|\theta_B\right) &\equiv \bar{\phi}_B\left(\hat{\theta}_B\right)\theta_B - \bar{t}_B\left(\hat{\theta}_B\right) \\
U_S\left(\hat{\theta}_S|\theta_S\right) &\equiv \bar{t}_S\left(\hat{\theta}_S\right) - \bar{\phi}_S\left(\hat{\theta}_S\right)\theta_S \\
U_B\left(\theta_B\right) &\equiv U_B\left(\theta_B|\theta_B\right) \\
U_S\left(\theta_S\right) &\equiv U_S\left(\theta_S|\theta_S\right)
\end{aligned}$$

We say a mechanism $\phi\left(\cdot,\cdot\right)$ is **ex-post efficient** if and only if

$$\phi\left(\theta_B, \theta_S\right) = \begin{cases} 1 & \text{if } \theta_B \geq \theta_S \\ 0 & \text{if } \theta_B < \theta_S \end{cases}$$

We say that (ϕ, t) is **incentive-compatible** if and only if

$$\begin{aligned}
U_B\left(\theta_B\right) &\geq U_B\left(\hat{\theta}_B|\theta_B\right), \forall \hat{\theta}_B, \theta_B \\
U_S\left(\theta_S\right) &\geq U_s\left(\hat{\theta}_S|\theta_S\right), \forall \hat{\theta}_S, \theta_S
\end{aligned}$$

We say that (ϕ, t) is (interim) individually rational if and only if $U_B\left(\theta_B\right) \geq 0, U_S\left(\theta_S\right) \geq 0$.

8.1.2 Characterizing the Equilibrium

We solve essentially similarly as before.

1. (ϕ, t) is IC if and only if

$$[1] : \bar{\phi}_B(\cdot) \uparrow, \quad \bar{\phi}_S(\cdot) \downarrow$$
$$[2] : U_B(\theta_B) = U_B(\underline{\theta}_B) + \int_{\underline{\theta}_B}^{\theta_B} \bar{\phi}_B(x)\, dx$$
$$U_S(\theta_S) = U_S(\bar{\theta}_S) + \int_{\theta}^{\bar{\theta}_S} \bar{\phi}_S(x)\, dx$$

   Note that [2] can be replaced with

$$[2^*] : \bar{t}_B(\theta_B) = \bar{\phi}_B(\theta_B)\theta_B - U_B(\underline{\theta}_B) - \int_{\underline{\theta}_B}^{\theta_B} \bar{\phi}_B(x)\, dx$$
$$\bar{t}_S(\theta_S) = U_S(\bar{\theta}_S) + \int_{\theta_S}^{\bar{\theta}_S} \bar{\phi}_S(x)\, dx - \bar{\phi}_S(\theta_S)\theta_S$$

2. Write the expected utilities of the buyer and the seller:

$$\mathbb{E}_{\theta_B}[U_B(\theta_B)] = U_B(\underline{\theta}_B) + \mathbb{E}\left[\phi(\theta_B, \theta_S)\frac{1 - F_B(\theta_B)}{f_B(\theta_B)}\right]$$
$$\mathbb{E}_{\theta_S}[U_S(\theta_S)] = U_S(\bar{\theta}_S) + \mathbb{E}\left[\phi(\theta_B, \theta_S)\frac{F_S(\theta_S)}{f_S(\theta_S)}\right]$$

Theorem 8.1. *(ϕ, t) is IC if and only if*

$$(*) : \mathbb{E}\left[\phi(\theta_B, \theta_S)\left(\left(\theta_B - \frac{1 - F_B(\theta_B)}{f_B(\theta_B)}\right) - \left(\theta_S + \frac{F_S(\theta_S)}{f_S(\theta_S)}\right)\right)\right] = U_B(\underline{\theta}_B) + U_S(\bar{\theta}_S)$$

and for any ϕ such that $\bar{\phi}_B \uparrow$ and $\bar{\phi}_S \downarrow$, if

$$(**) : \mathbb{E}\left[\phi(\theta_B, \theta_S)\left(\left(\theta_B - \frac{1 - F_B(\theta_B)}{f_B(\theta_B)}\right) - \left(\theta_S + \frac{F_S(\theta_S)}{f_S(\theta_S)}\right)\right)\right] \geq 0$$

then there exists a t such that (ϕ, t) is IC and IR.

Proof. We know that the total surplus generated by this mechanism, by definition, is

$$\mathbb{E}[\phi(\theta_B, \theta_S)(\theta_B - \theta_S)]$$

and by the budget balance, we have that total surplus is equal to sum of expected utilities for the two agents, $\mathbb{E}[U_B(\theta_B) + U_S(\theta_S)]$, which is equal to

$$= U_B(\underline{\theta}_B) + U_S(\bar{\theta}_S) + \mathbb{E}\left[\phi(\theta_B, \theta_S)\left(\frac{1 - F_B(\theta_B)}{f_B(\theta_B)} + \frac{F_S(\theta_S)}{f_S(\theta_S)}\right)\right]$$

which gives us $(*)$. For the second part, we will construct a transfer:

$$t(\theta_B, \theta_S) = \bar{\phi}_B(\theta_B)\theta_B + \bar{\phi}_S(\theta_S)\theta_S - \int_{\underline{\theta}_B}^{\theta_B} \bar{\phi}_B(x)\,dx + \int_{\theta_S}^{\bar{\theta}_S} \bar{\phi}_S(x)\,dx - \mathbb{E}_{\theta_S}\left[\bar{\phi}_S(\theta_S)\theta_S\right] - \mathbb{E}_{\theta_S}\left[\int_{\theta_S}^{\bar{\theta}_S} \bar{\phi}_S(x)\,dx\right]$$

With this expression, note that

$$\bar{t}_B(\theta_B) = \bar{\phi}_B(\theta_B)\theta_B - \int_{\underline{\theta}_B}^{\theta_B} \bar{\theta}_B(x)\,dx \Rightarrow IC \text{ by } [2^*] \text{ and } U_B(\underline{\theta}_B) = 0$$

and

$$\bar{t}_S(\theta_S) = \bar{\phi}_S(\theta_S)\theta_S + \int_{\theta_S}^{\bar{\theta}_S} \bar{\phi}_S(x)\,dx + \text{constant} \Rightarrow IC \text{ by } [2^*] \text{ and } U_S(\bar{\theta}_S) = \text{ constant}$$

Note that $(*)$ and $(**)$ implies $U_S(\bar{\theta}_S) \geq 0$. ■

Theorem 8.2. *(Impossibility Theorem) Suppose $\bar{\theta}_S > \underline{\theta}_B$ and $\underline{\theta}_S < \bar{\theta}_B$. Any ex-post efficient mechanism that is budget balanced and IC is not IR.*

Proof. We can show that LHS of $(*)$

$$(*) : \mathbb{E}\left[\phi(\theta_B, \theta_S)\left(\left(\theta_B - \frac{1 - F_B(\theta_B)}{f_B(\theta_B)}\right) - \left(\theta_S + \frac{F_S(\theta_S)}{f_S(\theta_S)}\right)\right)\right] = U_B(\underline{\theta}_B) + U_S(\bar{\theta}_S)$$

is negative at the efficient mechanism:

$$\begin{aligned}(p.272) : &\int_{\underline{\theta}_B}^{\theta_B} \int_{\underline{\theta}_S}^{\min\{\theta_B, \bar{\theta}_S\}} \left[\theta_B f_B(\theta_B) - (1 - F_B(\theta_B))\right] f_S(\theta_S)\, d\theta_S d\theta_B \\ &+ \int_{\underline{\theta}_B}^{\bar{\theta}_B} \int_{\underline{\theta}_S}^{\min\{\theta_B, \bar{\theta}_S\}} \left[\theta_S f_S(\theta_S) - F_S(\theta_S)\right] f_B(\theta_B)\, d\theta_S d\theta_B \\ &= \cdots \\ &= -\int_{\underline{\theta}_B}^{\bar{\theta}_S} (1 - F_B(\theta_B)) F_S(\theta_B)\, d\theta_B < 0\end{aligned}$$

■

From now on, we define:

$$\psi[\phi] \equiv \mathbb{E}\left[\phi(\theta_B, \theta_S)\left(\left(\theta_B - \frac{1 - F_B(\theta_B)}{f_B(\theta_B)}\right) - \left(\theta_S + \frac{F_S(\theta_S)}{f_S(\theta_S)}\right)\right)\right], \quad \phi^*(\theta_B, \theta_S) \equiv \begin{cases} 1 & \text{if } \theta_B \geq \theta_S \\ 0 & \text{otherwise} \end{cases}$$

8.1.3 Extension #1: Welfare Maximization

The welfare maximization problem becomes:

$$\max_{\phi(\cdot)} \mathbb{E}\left[\phi\left(\cdot\right)\left(\theta_B-\theta_S\right)\right] \text{ s.t. } \bar{\phi}_B \uparrow, \bar{\phi}_S \downarrow, \psi\left[\phi\right] \geq 0$$

The impossibility theorem tells us that $\psi\left[\phi\right]=0$ is binding in the welfare maximization problem.

- ▷ We are choosing a trading rule $\phi\left(\cdot\right)$ to maximize the social surplus. The associated Lagrangian:

$$\max_{\phi(\cdot)} \mathbb{E}\left[\phi\left(\cdot\right)\left(\theta_B-\theta_S\right)\right]+\lambda\psi\left[\phi\right]$$

  Note that both terms is linear in ϕ. We will solve point-wise and later check the monotonicity conditions.

- ▷ Plugging in the expression for $\psi\left[\phi\right]$:

$$\max_{\phi(\cdot)} \mathbb{E}\left[\phi\left(\theta_B,\theta_S\right)\left\{\theta_B\left(1+\lambda\right)-\lambda\frac{1-F_B\left(\theta_B\right)}{f_B\left(\theta_B\right)}-\left(\theta_S\left(1+\lambda\right)-\lambda\frac{F_S\left(\theta_S\right)}{f_S\left(\theta_S\right)}\right)\right\}\right]$$

- ▷ Dividing by $\left(1+\lambda\right)$:

$$\max_{\phi(\cdot)} \mathbb{E}\left[\phi\left(\theta_B,\theta_S\right)\left\{\left(\theta_B-\frac{\lambda}{1+\lambda}\frac{1-F_B\left(\theta_B\right)}{f_B\left(\theta_B\right)}\right)-\left(\theta_S-\frac{\lambda}{1+\lambda}\frac{F_S\left(\theta_S\right)}{f_S\left(\theta_S\right)}\right)\right\}\right]$$

- ▷ Thus, we have the optimal policy:

$$\phi_\lambda\left(\theta_B,\theta_S\right)=\begin{cases}1 & \text{if } \theta_B-\theta_S\geq\frac{\lambda}{1+\lambda}\left[\frac{1-F_B(\theta_B)}{f_B(\theta_B)}+\frac{F_S(\theta_S)}{f_S(\theta_S)}\right]\\ 0 & \text{otherwise}\end{cases}$$

  i.e. trade if the value generated $\theta_B-\theta_S$ is greater than some fraction of the information rent.

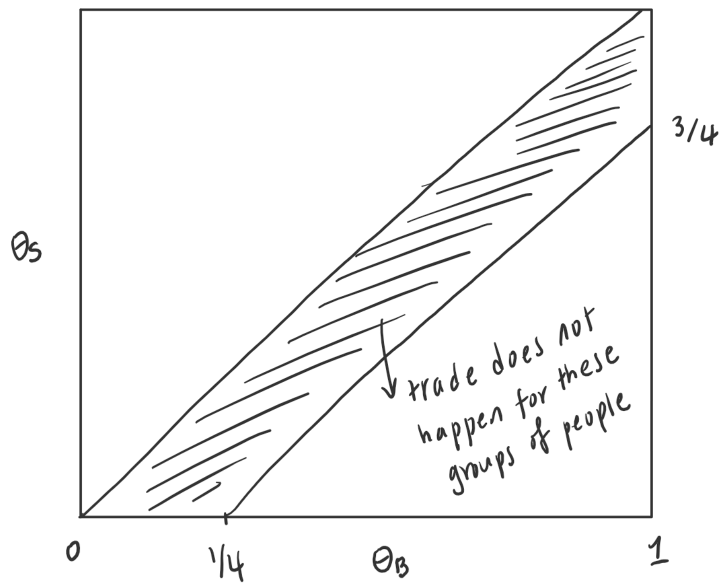
- ▷ We know that λ is determined by setting $\psi\left[\phi_\lambda\right]=0$. This is because know that the efficient $\psi\left[\phi^*\right]<0$, so when we solve the optimization problem with the constraint $\psi\left[\phi\right]\geq 0$, the constraint will try to become as negative as possible which gets it to zero.

As an example, let $\theta_B,\theta_S\sim u\left[0,1\right]$. Then

$$\phi_\lambda=\begin{cases}1 & \text{if } \theta_B-\theta_S\geq\left(\frac{\lambda}{1+\lambda}\right)\left(1-\theta_B+\theta_S\right)\Leftrightarrow\left(\theta_B-\theta_S\right)\geq\frac{\lambda}{1+2\lambda}=\text{gap}\end{cases}$$

Then what is λ? Setting $\psi\left[\phi_\lambda\right]=0$, we have that

$$\frac{\lambda}{1+2\lambda}=\frac{1}{4}$$



8.1.4 Extension #2: Platform-Profit Maximization

The buyer has preferences $U_B \equiv \phi\theta_B - t_B$ and $U_S = t_s - \phi\theta_S$.

$$\max_{\phi(\cdot)} \mathbb{E}\left[t_B\left(\theta_B, \theta_S\right) - t_S\left(\theta_B, \theta_S\right)\right]$$

With budget balance, we had $t = t_B = t_S$ but now we distinguish. Define:

$$\begin{aligned}
\bar{\phi}_B\left(\hat{\theta}_B\right) &= \mathbb{E}_{\theta_S}\left[\phi\left(\hat{\theta}_B, \theta_S\right)\right] \\
\bar{\phi}_S\left(\hat{\theta}_S\right) &= \mathbb{E}_{\theta_B}\left[\phi\left(\theta_B, \hat{\theta}_S\right)\right] \\
\bar{t}_B\left(\hat{\theta}_B\right) &= \mathbb{E}_{\theta_S}\left[t_B\left(\hat{\theta}_B, \theta_S\right)\right] \\
\bar{t}_S\left(\hat{\theta}_S\right) &= \mathbb{E}_{\theta_B}\left[t_S\left(\theta_B, \hat{\theta}_S\right)\right]
\end{aligned}$$

Then we can define:

$$\begin{aligned}
U_B\left(\hat{\theta}_B | \theta_B\right) &\equiv \bar{\phi}_B\left(\hat{\theta}_B\right)\theta_B - \bar{t}_B\left(\hat{\theta}_B\right) \\
U_S\left(\hat{\theta}_S | \theta_S\right) &\equiv \bar{t}_S\left(\hat{\theta}_S\right) - \bar{\phi}_S\left(\hat{\theta}_S\right)\theta_S \\
U_B\left(\theta_B\right) &\equiv U_B\left(\theta_B | \theta_B\right) \\
U_S\left(\theta_S\right) &\equiv U_S\left(\theta_S | \theta_S\right)
\end{aligned}$$

The platform wants to maximize:

$$\max_{\phi} \mathbb{E}_{\theta_B, \theta_S}\left[\bar{t}_B\left(\theta_B\right) - \bar{t}_S\left(\theta_S\right)\right] = \max_{\phi} \mathbb{E}_{\theta_B, \theta_S}\left[\phi\left(\cdot\right)\left(\theta_B - \theta_S\right) - U_B\left(\theta_B\right) - U_S\left(\theta_S\right)\right]$$

subject to $[IC]$ and $[IR]$.

1. First, note that (ϕ, t_B, t_S) is IC if and only if

$$[1] : \bar{\phi}_B(\cdot) \uparrow, \quad \bar{\phi}_S(\cdot) \downarrow$$
$$[2] : U_B(\theta_B) = U_B(\underline{\theta}_B) + \int_{\underline{\theta}_B}^{\theta_B} \bar{\phi}_B(x)\, dx$$
$$U_S(\theta_S) = U_S(\bar{\theta}_S) + \int_{\theta}^{\bar{\theta}_S} \bar{\phi}_S(x)\, dx$$

Note that [2] can be replaced with

$$[2^*] : \bar{t}_B(\theta_B) = \bar{\phi}_B(\theta_B)\,\theta_B - U_B(\underline{\theta}_B) - \int_{\underline{\theta}_B}^{\theta_B} \bar{\phi}_B(x)\, dx$$
$$\bar{t}_S(\theta_S) = U_S(\bar{\theta}_S) + \int_{\theta_S}^{\bar{\theta}_S} \bar{\phi}_S(x)\, dx - \bar{\phi}_S(\theta_S)\,\theta_S$$

2. Express

$$\mathbb{E}\left[U_B(\theta_B)\right] = U_B(\underline{\theta}_B) + \mathbb{E}\left[\phi(\cdot)\left(\frac{1 - F_B(\theta_B)}{f_B(\theta_B)}\right)\right]$$
$$\mathbb{E}\left[U_S(\theta_S)\right] = U_S(\bar{\theta}_S) + \mathbb{E}\left[\phi(\cdot)\left(\frac{F_S(\theta_S)}{f_S(\theta_S)}\right)\right]$$

3. The maximization problem then becomes:

$$\max_{\phi(\cdot)} \mathbb{E}\left[\phi(\theta_B, \theta_S)\left(\left(\theta_B - \frac{1 - F_B(\theta_B)}{f_B(\theta_B)}\right) - \left(\theta_S + \frac{F_S(\theta_S)}{f_S(\theta_S)}\right)\right)\right] \equiv \psi[\phi] \text{ s.t. } \bar{\phi}_B \uparrow, \bar{\phi}_S \downarrow$$

The solution is then

$$\phi^*(\theta_B, \theta_S) = \begin{cases} 1 & \text{if } \theta_B - \theta_S \geq \left(\frac{1 - F_B(\theta_B)}{f_B(\theta_B)} + \frac{F_S(\theta_S)}{f_S(\theta_S)}\right) \\ 0 & \text{otherwise} \end{cases}$$

Note that the solution that satisfies the constraints exists because $\psi[\phi] \geq 0$ by Theorem 1.

Turns out that this line lies below the welfare-maximizing line. You are raising your brokerage fees (you will have less efficiency in your market), but you will make more money from each trade. This is the exact analog of the monopoly solution.

8.1.5 Few Remarks

Here are some useful observations.

1. What's driving the impossibility theorem is the interim $[IR]$ constraints i.e. you have people who know their types and they have some cost being involved in the trade.You could alternatively imposes an ex-ante $[IR]$ constraints (when you don't know your types) where you play if the expected value is non-negative. It turns out that it's very easy to construct a mechanism that is efficient, $[IR]$, $[IC]$that satisfies ex-ante $[IR]$. The basic idea is that we need the bad types (both buyers and sellers) to have negative utility.

2. One way to think about weakening the $[IR]$ constraint to get rid of impossibility result is a famous one (it's actually a problem in Phil's textbook) from Cramdon, Gibbons, and Klemperer (1987). They are looking at what's called a "Dissolving a Partnership." (See Reny's 9.29)

- ▷ The value of the partnership to i is $\theta_i \sim u\,[0,1]$. Suppose currently it is $\theta_1/2$ for θ_1 and $\theta_2/2$ for θ_2. If $\theta_2 > \theta_1$, then it's efficient for $[1]$ to sell $1/2$ shares to 2 and vice versa.
- ▷ The difference is that the starting point is that you already own half the good. In the original starting point is that one person owns one and the other guy owns zero.
- ▷ Consider the policy where if 1 buys shares then pays $\frac{1}{3}\hat{\theta}_1$ and if 2 buys out 1, then 2 pays $\frac{1}{3}\hat{\theta}_2$. The $[IR]$ constraint is slack.

The key idea is that if you have a more dispersed ownership of the underlying assets, the impossibility theorem does not hold.

9 Design of Efficient Mechanisms

1. Can we implement ex-post efficiency with DSIC? $\Rightarrow$ budget balance (BB) will not be met.
2. Can we implement ex-post efficiency with BIC? $\Rightarrow$ budget balance (BB) will be achieved.
3. Can we implement BIC, IR, and BB? $\Rightarrow$ VCG will be useful in answering this question.

9.1 General Framework

We talked about profit-maximization before – now we ask the question of efficiency. Our main focus is in implementing efficient allocations when agents have private information that is relevant for efficiency.

9.1.1 Setup

Preferences are independent (to prevent achieving first-best using a side-bet). Consider a government i.e. the mechanism designer. Define

$$U_i = u_i\left(x, \theta_i\right) - t_i, \quad x \in X = \{x_1, ..., x_k\}$$

where X is the set of social outcomes that the government chooses. Also, $\theta_i \in \Theta_i$. The direct mechanism will be the following:

$$\left\{\phi\left(x|\hat{\boldsymbol{\theta}}\right), t_1\left(\hat{\boldsymbol{\theta}}\right), ..., t_n\left(\hat{\boldsymbol{\theta}}\right)\right\}$$

We already have the revelation principle from before – if we want to look at BNE, we can restrict our attention to direct mechanisms. We also want to define a notion of **ex-post efficiency**:

$$\hat{x}\left(\theta\right) \equiv \arg\max_{x \in X} \sum_{i=1}^{n} u_i\left(x, \theta_i\right)$$

We are assuming that it's unique. Since we have quasi-linear utility, then this is Pareto-optimal.

9.1.2 Two Notions of Incentive Compatibility

Throughout we will consider DRM of the form

$$\left\{\phi\left(\cdot|\cdot\right), t_1, .., t_n\right\}$$

where $\phi\left(x|\hat{\theta}\right)$ gives the probability that social state $x \in \mathcal{X}$ is chosen when agents report $\hat{\theta} = \left(\hat{\theta}_1, ..., \hat{\theta}_n\right)$. Also define

$$\bar{t}_i\left(\hat{\theta}_i\right) \equiv \mathbb{E}\left[t_i\left(\hat{\theta}_i, \theta_{-i}\right)\right]$$
$$\bar{\phi}_i\left(x|\hat{\theta}_i\right) \equiv \mathbb{E}\left[\phi\left(x|\hat{\theta}_i, \theta_{-i}\right)\right]$$

and

$$U_i\left(\hat{\theta}_i|\theta_i\right) \equiv \sum_{x \in \mathcal{X}} \bar{\phi}_i\left(x|\hat{\theta}_i\right) u_i\left(x, \theta_i\right) - \bar{t}_i\left(\hat{\theta}_i\right)$$

Bayesian IC BIC requires that a player is willing to tell the truth given her expected payoffs when other agents' are also telling the truth. Thus, a direct mechanism $\{\phi, t_1, .., t_n\}$is BIC if and only if for all i,

$$U_i(\theta) \geq U_i\left(\hat{\theta}_i|\theta_i\right), \forall \theta_i, \hat{\theta}_i$$

Proposition 9.1. *(BIC Revelation Principle) Statement omitted.*

Proof. If all players other than player i report their types truthfully, then player i can achieve the same equilibrium payoffs as in the original by also reporting truthfully. Moreover, any non-truthful report by player i will correspond to choosing the equilibrium strategy of another type of player i in the original game. Because the original allocation is a BNE, it must be that choosing a different strategy is not preferred to reporting truthfully. ■

Dominant-Strategy IC There is a stronger notion of incentive compatibility than BIC, which is the dominant-strategy IC.

▷ Compare the first-price auction to the second-price auction. In the second-price auction, it did not require common knowledge of the distribution of types, whereas the first-price auction did.

We say that a direct mechanism $\{\phi, t_1, ..., t_n\}$ is **dominant strategy IC**: if and only if :

$$\sum_{x\in X} \phi\left(x|\theta_i, \hat{\theta}_{-i}\right) u_i(x|\theta_i) - t_i\left(\theta_i, \hat{\theta}_{-i}\right) \geq \sum_{x\in X} \phi\left(x|\hat{\theta}_i, \hat{\theta}_{-i}\right) u_i(x_i|\theta_i) - t_i\left(\hat{\theta}_i, \hat{\theta}_{-i}\right), \quad \forall \hat{\theta}_i, \theta_i, \hat{\theta}_{-i}$$

▷ In the Bayesian, we took the expectation to get rid of $\hat{\theta}_{-i}$ and replaced it with $\bar{\phi}$. This implies Bayesian $[IC]$ – this is stronger. BIC requires that i prefers to tell the truth after taking expectations using a commonly known probability distribution.

▷ In the monopoly-screening environment with a single agent of unknown type, BIC and DSIC are equivalent.

Ex-Post IC This is slightly weaker than DSIC but generally much stronger than BIC. This requires that it is a dominant-strategy *in equilibrium* for i to report truthfully for any θ_{-i}. That is, after all the reports are revealed, if $\hat{\theta}_{-i} = \theta_{-i}$, the player has no regret about reporting $\hat{\theta}_i = \theta_i$.

9.2 Vickrey-Clarke-Groves (VCG) Mechanism

The VCG Mechanism plays a central role in solving the ex-post efficient allocation. It can be thought of as a generalization of a second-price auction. The question we seek to address is the following:

Can we implement $\hat{x}(\theta)$ for any $\theta \in \Theta$ using a DISC mechanism?

We will get around the Gibbard-Satterthwaite Theorem by limiting our preferences to quasi-linear utility, since the original theorem's conditions included all possible preferences.

9.2.1 Groves Mechanism (1973)

We are interested in the following question:

Can we implement $\hat{x}(\theta)$ for any $\theta \in \Theta$ using a DISC mechanism?

The answer is "yes." The idea is to construct transfers in a way that each agent pays an amount equal to the impact of the agent's report on social welfare, i.e. agent i pays her social externality, where the externality is evaluated using the reports of the other agents.

- ▷ In the second-price auction, for example, the highest bidder won and paid the second-highest bid. Her the highest bidder essentially "paid his externality." The reason is that if the winner were not present, the bidder with second highest value would have won. Thus, the winning bidder, by virtue of his presence, precludes the second-highest value from being realized and imposes an externality. Of course, he pays for the good precisely the amount of the externality he imposes, and the end result is efficient.

Definition 9.1. (Groves Mechanism) $\left\{\hat{x}\left(\cdot\right), t_1^g, \ldots, t_n^t\right\}$ where

$$t_i^g\left(\hat{\theta}\right) \equiv \left(-\sum_{j\neq i}^{n} u_j\left(\hat{x}\left(\hat{\theta}_i, \hat{\theta}_{-i}\right), \hat{\theta}_j\right)\right) + h_i\left(\hat{\theta}_{-i}\right)$$

Note that this is two-part.

- ▷ First term: You are given the social surplus of all other players when they play $\hat{x}$
- ▷ Second term: Some constant, which does not depend on $\hat{\theta}_i$

Theorem 9.1. *The Groves Mechanism, defined above, is DSIC.*

Proof. We need to prove that for any player i, your weak best response is to say θ_i regardless of what other people say.

- ▷ The utility from reporting $\hat{\theta}_i$ is

$$U_i\left(\hat{\theta}_i|\theta_i, \hat{\theta}_{-i}\right) = \underbrace{U_i\left(\hat{x}\left(\hat{\theta}_i, \hat{\theta}_{-i}\right), \theta_i\right) + \sum_{j\neq i} U_j\left(\hat{x}\left(\hat{\theta}_i, \hat{\theta}_{-i}\right), \theta_i\right)}_{\text{total surplus evaluated at } \left(\theta_i, \hat{\theta}_{-i}\right)} - h_i\left(\hat{\theta}_{-i}\right)$$

- ▷ We want to find a report that maximizes the term above, but we know that the maximizing choice is $\hat{x}\left(\theta_i, \hat{\theta}_{-i}\right)$ so the best response is to report truthfully.

■

Remark from Green-Laffont (1979): Two neat results come out from the Groves mechanism:

1. If preferences $u_i \in \mathcal{U}$ are "rich," then any DSIC ex-post efficiency mechanism is a Groves mechanism.
2. Under the same assumptions,

$$\sum_{i=1}^{n} t_i\left(\boldsymbol{\theta}\right) \neq 0, \forall \boldsymbol{\theta}$$

The key idea here is to define individual costs so that each individual internalizes the externality that, through his report, he imposes on the rest of the society.

9.2.2 VCG Mechanism

Define:

$$\hat{x}_{-i}\left(\hat{\theta}_{-i}\right) = \arg\max_{x\in X} \sum_{j\neq i}^{n} U_j\left(\hat{x}_{-i}\left(\hat{\theta}_{-i}\right), \hat{\theta}_j\right)$$

Then essentially, we are picking $h_i\left(\hat{\theta}_{-i}\right)$ such that

$$h_i\left(\hat{\theta}_i\right) = \sum_{j\neq i} u_j\left(\hat{x}_{-i}\left(\hat{\theta}_{-i}\right), \hat{\theta}_j\right)$$

Definition 9.2. (VCG Mechanism) Each individual simultaneously reports his type to the designer, and the social state $\hat{x}\left(t\right)$ is chosen. In addition, each individual is assessed a monetary cost equal to

$$t_i^{VCG}\left(\theta\right) = \sum_{j\neq i} U_j\left(\hat{x}_{-i}\left(\hat{\theta}_{-i}\right), \hat{\theta}_j\right) - \sum_{j\neq i} U_j\left(\hat{x}\left(\hat{\theta}_i, \hat{\theta}_{-i}\right), \hat{\theta}_j\right) \geq 0$$

i.e. each individual pays his externality based on the reported types.

- ▷ The first piece is how much fun everyone had when I was not around
- ▷ The second piece is how much fun everyone has when I am around. This is because it changes $\hat{x}$.

Definition 9.3. ((Pivot Mechanism) A pivot mechanism is one where $t_i = 0$ if i is not pivotal and $t_i > 0$ if it is pivotal, where i's report being pivotal is defined to be

$$\hat{x}_{-i}\left(\hat{\theta}_{-i}\right) \neq \hat{x}\left(\hat{\theta}_i, \hat{\theta}_{-i}\right)$$

Since VCG is an example of Groves mechanism, it is DSIC.

1. **VCG transfers are non-negative.** Whenever agents are pivotal, $\sum_i t_i^{VCG}\left(\theta\right) > 0$.

   This can be easily seen by noting that $\hat{x}_{-i}\left(\hat{\theta}_{-i}\right)$ maximizes $\sum_{j\neq i} u_j$ and $\hat{x}\left(\hat{\theta}\right)$ does not necessarily.

   - ▷ This implies that a VCG mechanism will generally run a budget surplus.
   - ▷ From the point of view of the n agents, this is inefficient since the mechanism has taken away $\sum_j t_j^{VCG}$ units of utility. Pareto efficiency here not only requires that $x = \hat{x}\left(\theta\right)$ but also $\sum_i t_i = 0$ (budget balance).

   Therefore, you will have a budget surplurs, and VCG auctions run budget surpluses.

2. **Players will voluntarily participate.**

   Suppose that if player i does not participate in the VCG mechanism, then the VCG mechanism is played for the remaining $n-1$ players, and player i pays nothing. There are, however, utility consequences. In this case, it is an equilibrium for all agents to voluntarily participate in a VCG mechanism.

   - ▷ If i participates, then

$$U_i\left(\hat{x}\left(\theta\right), \theta_i\right) + \left[\sum_{j\neq i} U_j\left(\hat{x}\left(\theta\right), \theta\right) - \sum_{j\neq i} U_j\left(\hat{x}_{-i}\left(\theta_{-i}\right), \theta_j\right)\right]$$

▷ If i does not participate, then:

$$U_i\left(\hat{x}_{-i}\left(\theta_{-i}\right),\theta_i\right)$$

▷ The difference between these two utilities:

$$\begin{aligned}\Delta U_i &\equiv U_i\left(\hat{x}\left(\theta\right),\theta_i\right)+\left[\sum_{j\neq i}U_j\left(\hat{x}\left(\theta\right),\theta\right)-\sum_{j\neq i}U_j\left(\hat{x}_{-i}\left(\theta_{-i}\right),\theta_j\right)\right]-U_i\left(\hat{x}_{-i}\left(\theta_{-i}\right),\theta_i\right)\\&=\sum_{j=1}^{n}U_j\left(\hat{x}\left(\theta\right),\theta_j\right)-\sum_{j=1}^{n}U_j\left(\hat{x}_{-i}\left(\theta_{-i}\right),\theta_j\right)\end{aligned}$$

Since $\hat{x}\left(\boldsymbol{\theta}\right)$ maximizes the sum while $\hat{x}_{-i}\left(\theta_{-i}\right)$ does not, it follows that the above expression is ≥ 0, assuming agents do not control $x\in X$.

3. Truth-telling is weakly dominant.

4. Computing the transfer, albeit tedious, is actually not that bad.

9.2.3 VCG in Bilateral Trade

9.3 BIC Implementation

We now ask whether or not we can achieve budget balance (BB) by replacing the DSIC requirement with the weaker notion of BIC. The answer is "yes."

9.3.1 Existence

The only option to achieve budget balance is to redistribute the revenue among the N individuals, but this causes problems. The cost functions now overstates the actual cost since they don't take into account the redistributed revenue. Thus, it is not clear that once individuals take into account their share of the revenue that is generated, it remains a dominant strategy to report their types truthfully.

Theorem 9.2. *Let $\{\phi,t_1,...,t_n\}$ that is BIC, and such that the expected budget surplus is non-negative:*

$$\mathbb{E}\left[\sum_{i=1}^{n}t_i\left(\theta\right)\right]\geq 0$$

Then there exists a new mechanism $\left\{\phi,\tilde{t}_1,...,\tilde{t}_n\right\}$ such that $\sum_{i=1}^{n}\tilde{t}_i\left(\theta\right)=0$ (budget-balanced) and the utilities of the players

$$\tilde{U}_i\left(\theta_i\right)\geq U_i\left(\theta_i\right),\forall i,\forall\theta_i$$

In other words, give me a mechanism that has budget surplus, I can find a budget-balanced mechanism that is weakly better.

Proof. We prove by construction. Suppose people are sitting in a circle.

▷ Denote $\bar{t}_i=\mathbb{E}_{\theta_i}\left[\bar{t}_i\left(\theta_i\right)\right]$ and consider:

$$\tilde{t}_i\left(\theta\right)\equiv\bar{t}_i\left(\theta_i\right)+\left(\bar{t}_{i+1}-\bar{t}_{i+1}\left(\theta_{t+1}\right)\right)-\frac{1}{n}\sum_{j=1}^{n}\bar{t}_j$$

i.e. give everybody their (1) interim payments, (2) something that is on average zero, minus (3) your share of the surplus.

▷ To see budget balance, take the sum of all these new transfers, we have

$$\sum_{i=1}^{n} \tilde{t}_i(\theta) = 0$$

▷ To see that everyone is weakly better off, write out the relevant objects:

$$\begin{aligned} U_i\left(\hat{\theta}_i|\theta_i\right) &\equiv \sum_{x\in X} \bar{\phi}_i\left(x|\hat{\theta}_i\right) U_i(x,\theta_i) - \bar{t}_i(\theta_i) \\ \tilde{U}_i\left(\hat{\theta}|\theta_i\right) &= \sum \bar{\phi}_i\left(x|\hat{\theta}_i\right) U_i(x,\theta_i) - \mathbb{E}_{\theta_{-i}}\left[\tilde{t}_i(\theta_i,\theta_{-i})\right] \\ &= U_i\left(\hat{\theta}_i|\theta_i\right) + \bar{t}_i(\theta_i) - \bar{t}_i(\theta_i) + \frac{1}{n}\sum_{j=1}^{n}\bar{t}_j \\ &= U_i\left(\hat{\theta}_i|\theta_i\right) + \frac{1}{n}\sum_{j=1}^{n}\bar{t}_j \end{aligned}$$

which shows that $\left(\phi, \tilde{t}_1, ..., \tilde{t}_n\right)$ is BIC.

A **corollary** is the following: $\exists$ a .ex-post efficient, BIC, BB mechanism. The proof is that

$$\tilde{t}_i(\theta) \equiv \bar{t}_i^{VCG}(\theta_i) + \left(\bar{t}_{i+1}^{VCG} - \bar{t}_{i+1}^{VCG}(\theta_{t+1})\right) - \frac{1}{n}\sum_{j=1}^{n}\bar{t}_j^{VCG}$$

■

9.3.2 Budget-Balanced Expected Externality Mechanism

The agent i's expected externality when his type is t_i is defined to be

$$\begin{aligned} \bar{c}_i^{VCG}(t_i) &= \sum_{t_{-i}\in T_{-i}} q_{-i}(t_{-i})\, c_i^{VCG}(t_i,t_{-i}) \\ &= \sum_{t_{-i}\in T_{-i}} q_{-i}(t_{-i}) \left(\sum_{j\neq i} v_j\left(\tilde{x}^j(t_{-i}), t_j\right) - \sum_{j\neq i} v_j\left(\hat{x}(t), t_j\right)\right) \end{aligned}$$

These expected externalities can then be used to define costs in a way that delivers ex-post efficiency and a balanced budget.

Definition 9.4. The Expected Externality (EE) mechanism is the ex-post efficient, budget-balanced, BIC mechanism with payments

$$t_i^{EE}(\theta) \equiv \bar{t}_i^{VCG}(\theta_i) - \left(\bar{t}_{i+1}^{VCG}(\theta_{i+1}) - \bar{t}_{i+1}^{VCG}\right) - \frac{1}{n}\sum_{j=1}^{n}\bar{t}_j^{VCG}$$

where

$$\bar{t}_i^{VCG}(\theta_i) \equiv \mathbb{E}_{\theta_{-i}}\left[t_i^{VCG}(\theta_i,\theta_{-i})\right], \quad \bar{t}_i^{VCG} \equiv \mathbb{E}_{\theta}\left[t_i^{VCG}(\theta)\right]$$

This is also referred to as an AGV mechanism. This says that player i now pays her expected externality on the remaining $(n-1)$ agents. EE mechanisms can solve a lot of interesting economic problems.

▷ Re-visit the bilateral trade example where agents can commit to the mechanism before learning their types. In this case, they can implement the first-best by constructing an expected-externality mechanism.

▷ Specifically:

9.4 Individually-Rational Mechanisms

Now we continue to explore BIC mechanisms, but we now suppose that each agent i has a type-dependent interim IR constraint given by $\underline{U}_i(\theta_i)$. This IR constraint may capture an exogenous requirement for agent payoffs, or it may capture the value the agent could obtain by exercising control over some components of $\mathcal{X}$.

9.4.1 Individual Rationality

We are interest

9.4.2 IR-VCG Mechanism

We are interest