Mechanisms of Mechanisms Design

ECON 8104 note

Jakub Pawelczak*

May 2021

Contents

*These notes are intended to summarize the main concepts, definitions and results covered in the first year of micro sequence for the Economics PhD of the University of Minnesota. The material is not my own. Please let me know of any errors that persist in the document. E-mail: pawel042@umn.edu .

1 Voting and Power

1.1 Simple Games

Let $N = \{1, \ldots, n\}$ be a finite set of n individuals, or players.

Definition 1 (Coalition). *A subset $D \subseteq N$ is called a coalition*

N is the grand coalition.

Given a coalition D, write $D' = N \backslash D$ for its complement. The collection of all coalitions (i.e., the power set of N) is denoted by 2^N.

Definition 2 (Simple game). *(or simply a game) on N is a family $\mathcal{D}$ of nonempty coalitions, called winning or strong, assumed to be monotone:*

$$D \in \mathcal{D} \text{ and } D \subset E \subset N \text{ implies } E \in \mathcal{D}$$

Think of a simple game in terms of voting. Consider a binary voting problem before the electorate: every individual in society must choose whether to vote for the status quo, say, or for change. Some players, collected in the coalition D, will vote for change. If $D \in \mathcal{D}$ then we say that D, or D 's issue wins. Simple majority with n voters can be expressed as a simple game whose collection of winning coalitions is:

Definition 3 (Winning coalitions). *satisfy*

$$M_n = \{D \subset \{1, \ldots, n\} : |D| > n/2\}$$

Definition 4 (Proper). *A game $\mathcal{D}$ is called proper when*

$$D \in \mathcal{D} \text{ implies } D' \notin \mathcal{D}$$

Definition 5 (Strong). *A game $\mathcal{D}$ is called strong if there does not exist a coalition D such that neither D nor D' belong to $\mathcal{D}$, that is*

$$D \notin \mathcal{D} \text{ implies } D' \in \mathcal{D}$$

In a binary voting context, a game that is proper and strong always renders an outcome. For interpretation, say each player i chooses from two alternatives, a and b. Let D_a be the coalition of players choosing a and $D_b = D_a'$ those choosing b. Let us say that a game D reveals that society prefers a to b denoted aRb - if $D_a \in D$, and **reveals that society strictly prefers** a **to** b -denoted aPb -if aRb but not bRa. The following exercise suggests that strong games reveal social preferences; proper and strong games reveal strict social preferences.

Definition 6 (Veto)**.** *A player i has a veto if he belongs to every winning coalition, that is, $D \in \mathcal{D}$ implies $i \in D$.*

Definition 7 (Dictator)**.** *Player i is called a dictator if i has a veto and every coalition to which i belongs wins, that is, $i \in D$ if and only if $D \in \mathcal{D}$.*

Thus, in the example above, player 1 is not a dictator despite being the only player with a veto. If player i is a dictator then $\{i\} \in \mathcal{D}$ therefore no other player can have a veto: if j has a veto, he belongs to every winning coalition, so in particular $j \in \{i\}$ and $i = j$. As a result, there can be only one dictator.

Definition 8 (Free game)**.** *Let us call a game free if no player has a veto.*

The presence of players with veto has immediate implications.

Theorem 1. ***A game is proper if it isn't free. If the game is strong, too, it has a dictator.***

Proof. Let player i have a veto. Every pair of complementary coalitions has one but not the other contain i. The one without i cannot win, so the game is proper. If the game is strong, too, then $(*)$ holds for every coalition, so $\{i\} \in \mathcal{D}$. By monotonicity, this implies $D \in \mathcal{D}$ whenever $i \in D$. Since i has a veto, $i \in D$ whenever $D \in \mathcal{D}$ too, so i is a dictator. □

Definition 9 (Generalized majority rule)**.** *The class of games that upholds symmetry is generalized majority rule. If $k, n \in \mathbb{N}$ and $k \leq n$*

$$M_{n,k} = \{D \subset \{1, \ldots, n\} : |D| \geq k\}$$

is the game with n players whose winning coalitions have at least k members.

If $k > (n+1)/2$ then $M_{n,k}$ is proper but not strong, and if $k < (n-1)/2$ then $M_{n,k}$ is strong but not proper. If $k = (n+1)/2$ then n is odd and $M_{n,k}$ is both proper and strong, that is, it reveals strict preference.

Definition 10 (Weighted majority)**.** *Given "voting strength" weights w_i for each player i and an overall quota q, a weighted majority game determines a coalition D to be winning* *whenever*

$$\sum_{i \in D} w_i \geq q$$

Let us write $W(q;w)$ for an n -person weighted majority game with quota q and weights vector $w = (w_1, \ldots, w_n)$.

Measuring power

This follows Shapley and Shubik (1954) and Straffin (1977).

Definition 11 (The Shapley-Shubik index). *For each $i \in N$, the Shapley-Shubik power index of player i, denoted by φ_i, is calculated as follows:*

$$\varphi_i = \sum_{S \ni i} \frac{(s-1)!(n-s)!}{n!} [v(S) - v(S \backslash \{i\})], \quad \textit{where } s = |S|$$

Player i is pivotal if he swings an election, that is, if $v(S) = 1$ but $v(S \backslash \{i\}) = 0$ for some coalition S. One by one, players cast their votes. Player i is pivotal for coalition S if, when all other members of S have voted before him for alternative a, say, i 's vote for a renders a winning but his vote for b renders a losing, when everyone after i votes for b. In other words, S needs i to win.

According to the Shapley-Shubik index, player i 's pivotality in coalition S must be weighted by the number of orderings that keep all of S voting before everyone else and i the last member of S to vote. There are $(s-1)!$ different orderings of the players before i and $(n-s)!$ orderings of those behind i. In total, the players in N can be ordered in $n!$ different ways. This explains the weight on player i 's pivotality in coalition S above. Shapley and Shubik suggest that the different voting orders may be interpreted more loosely - for instance, as different degrees of enthusiasm for an alternative.

Definition 12 (Penrose-Banzhaf index). *denoted by π_i for every player i and expressed as:*

$$\pi_i = \sum_{S \ni i} \frac{v(S) - v(S \backslash \{i\})}{2^{n-1}}$$

This index simply counts the number of times a player is pivotal. The parameter 2^{n-1} corresponds to the number of coalitions that a player could swing. Normalizations are arbitrary, of course, so the ratio of power indices is arguably more meaningful than individual values.

Exercise 1. *To illustrate how the two indices above are calculated, consider the following simple example: the weighted majority game $W(3;2,1,1)$. This game has three players. Those* *that are pivotal in each of the six possible orderings (123,132,213,231,312,321) are (2,3,1,1,1,1) , respectively; therefore, $\varphi_1 = 2/3$ and $\varphi_2 = \varphi_3 = 1/6$. On the other hand, player 1 is pivotal in $\overline{12}, \overline{13}$ and $\overline{123}$, whereas players 2 and 3 are only pivotal in $\overline{12}$ and $\overline{13}$, respectively. Therefore, $\pi_1 = 3/4$ and $\pi_2 = \pi_3 = 1/4$*

Exercise 2. *Say the electorate is to vote between two alternatives, a and b. Assume that players vote independently and let $p_i \in [0,1]$ be the probability with which player i votes for alternative a. These probabilities, together with the simple game, determine the extent to which a player's vote matters. Thus, one way to measure power is by making assumptions on the population distribution of voting probabilities. Another is to assume that every player votes with the same probability. These two assumptions characterize the power indices above.*

To see this, define i 's power as

$$P_i(p_1, \ldots, p_n) = \sum_{S \ni i} \left[\prod_{j \in S \setminus \{i\}} p_j \prod_{k \in N \setminus S} (1 - p_k) \right] \cdot [v(S) - v(S \setminus \{i\})]$$

the probability that player i is pivotal, that is, everyone in $S \setminus \{i\}$ votes for a, everyone outside of S votes for b, and i 's vote determines the electoral outcome. Treating the individual voting probabilities $\{p_k : k \in N\}$ as random variables, let $\mathbb{E}[P_i]$ be the expectation of P_i.

Theorem 2. ***If** i **is any player then** $\mathbb{E}[P_i] = \pi_i$ **whenever** $p_k \sim U[0,1]$ **is IID for every** k **and** $\mathbb{E}[P_i] = \sigma_i$ **whenever** $p_k = p \sim U[0,1]$ **for every** k.*

Proof. If $p_k \sim U[0,1]$ is IID then $\mathbb{E}[p_k] = \frac{1}{2}$ for all k and the p_k 's are independent, so

$$\mathbb{E}[P_i] = \sum_{S \ni i} \left[\prod_{j \in S \setminus \{i\}} \mathbb{E}[p_j] \prod_{k \in N \setminus S} \mathbb{E}[1 - p_k] \right] \cdot [v(S) - v(S \setminus \{i\})] = \sum_{S \ni i} \frac{v(S) - v(S \setminus \{i\})}{2^{n-1}} = \pi_i$$

By the well-known beta-function identity

$$\int_0^1 x^\alpha (1-x)^\beta dx = \frac{\alpha! \beta!}{(\alpha + \beta + 1)!} \quad \alpha, \beta \in \mathbb{N},$$

$$\mathbb{E}[P_i] = \sum_{S \ni i} \int p^{s-1} (1-p)^{n-s} dp [v(S) - v(S \setminus \{i\})] = \sigma_i \quad \text{when} \quad p_k = p \sim U[0,1] \quad k$$

□

1.2 Condorcet Theorem

Elections may be interpreted as a way of aggregating objective information about candidates, rather than subjective preferences over them. This was Condorcet's view when he offered the following argument in 1785.

- Consider an electorate consisting of n voters where every voter i makes the 'better' choice for society $(v_i = +1)$ among two competing alternatives (± 1) with probability p, independently of others.

- One might wish to relax this assumption and permit differing opinions for what constitutes a better choice, but at least this seems like a natural starting point for a normative evaluation of voting systems.

- For simplicity, assume that the number of voters is odd.

- If P_n is the probability that society elects the better candidate then

$$P_n = \Pr\left[\sum_{j=1}^{n} v_j > 0\right] = \sum_{k=\ell}^{n} \binom{n}{k} p^k (1-p)^{n-k}$$

  where $\ell = (n+1)/2$, since we are assuming that n is odd.

Theorem 3 (Condorcet). *If $p > \frac{1}{2}$ then $P_n \to 1$ as $n \to \infty$. If $p < \frac{1}{2}$ then $P_n \to 0$ as $n \to \infty$. If $p = \frac{1}{2}$ then $P_n = \frac{1}{2}$ for every odd number n.*

Theorem 4. *If $\bar{p} > \frac{1}{2}$ then $P_n \to 1$ as $n \to \infty$. If $\bar{p} < \frac{1}{2}$ then $P_n \to 0$ as $n \to \infty$. If $\bar{p} = \frac{1}{2}$ then $P_n \to \frac{1}{2}$ as $n \to \infty$.*

Statistics of Voting Power

How is power distributed in large elections? To begin to address this question, consider two alternatives before the electorate, denoted by ± 1.

Let $v_i \in \{\pm 1\}$ be voter i 's vote and $v = (v_1, \ldots, v_n)$ collect all votes. The result of the election is denoted by $R(v)$. If $w = (w_1, \ldots, w_n) \geq 0$ is a vector voting weights, and $v \cdot w = \sum_k v_k w_k$ the weighted sum votes, simple weighted majority, denoted by $W_n(w)$, is determined by

$$R(v) = \begin{cases} +1 & \text{if } v \cdot w > 0 \\ -1 & \text{if } v \cdot w < 0 \\ \frac{1}{2}[+1] + \frac{1}{2}[-1] & \text{if } v \cdot w = 0 \end{cases}$$

that is, if $v \cdot w > 0$ then alternative +1 is enacted, if $v \cdot w < 0$ then -1 is chosen, and if $v \cdot w = 0$ then society flips a fair coin to choose between ± 1 . This breaking of ties does not exactly fit into the format of simple games, but as we shall see, the cost of subscribing to symmetry in this case will be negligible. Let $V_{-i} = R(v) - v_i w_i$ denote the weighted votes of everyone but voter i, and $\|w\| = \sqrt{w \cdot w}$ the 'length' of the weights vector. With this notation, i 's voting power equals

$$P_i = \Pr\left(|V_{-i}| < w_i\right) + \frac{1}{2} \Pr\left(|V_{-i}| = w_i\right)$$

Of course, optimal weights vectors are not unique, but they can be chosen uniquely in a way that is independent of others' discernment, as follows. Consider a benevolent planner who is able to aggregate the vector of society's votes $v = (v_1, \dots, v_n)$ into a single decision that maximizes society's chance of making the right choice. The planner's prior belief that +1 is the better choice is summarized in the probability p_0. For every $i \in N \cup \{0\}$, let $p_i \in (0,1)$ and

$$w_i = \ln\left(\frac{p_i}{1 - p_i}\right)$$

denote i 's likelihood ratio of +1 relative to -1 being society's better choice.

Theorem 5. ***The planner's optimal rule is given by***

$$R^*(v) = \begin{cases} +1 & \text{if } v \cdot w + w_0 > 0 \\ -1 & \text{if } v \cdot w + w_0 < 0 \end{cases}$$

both candidates are optimal when $v \cdot w + w_0 = 0$.

Proof. Let $D = \{i \in N : v_i = +1\}$ be the coalition that voted $+1$, with $D' = N \backslash D$ the coalition that voted for -1 . Since p_i is the probability that voter i votes correctly, the probability of observing the votes vector v conditional on +1 being the right choice is given by

$$\Pr(v \mid +1) = \prod_{k \in D} p_k \prod_{\ell \in D'} (1 - p_\ell)$$

The probability $\Pr(v \mid -1)$ of v when the right choice is -1 follows from this formula by noticing that $-v$ preserves who voted correctly and incorrectly as the right choice changes from +1 to -1 , therefore $\Pr(v \mid -1) = \Pr(-v \mid +1)$. Using Bayes' Rule, the joint probability of v and +1 being the right choice is given by $\Pr(v \wedge +1) = \Pr(v \mid +1)\Pr(+1)$, where $\Pr(+1)$ is the planner's prior probability p_0 of +1 being the right choice. Of course, the same formula $\Pr(v \wedge -1) = \Pr(v \mid -1)\Pr(-1)$ obtains for -1 . Choosing +1 is a best response for the planner to v if and only if $\Pr(+1 \mid v) \geq \Pr(-1 \mid v)$. By Bayes' Rule, $\Pr(+1 \mid v) = \Pr(v \wedge +1)/\Pr(v)$ and $\Pr(-1 \mid v) = \Pr(v \wedge +1)/\Pr(v)$. Substituting:

$$\Pr(+1 \mid v) \geq \Pr(-1 \mid v) \Leftrightarrow \Pr(v \wedge +1) \geq \Pr(v \wedge -1) \Leftrightarrow \Pr(v \mid +1)\Pr(+1) \geq \Pr(v \mid -1)\Pr(-1)$$

$$\Leftrightarrow p_0 \prod_{k \in D} p_k \prod_{\ell \in D'} (1 - p_\ell) \geq (1 - p_0) \prod_{k \in D} (1 - p_k) \prod_{\ell \in D'} p_\ell$$

Taking logs and rearranging, the result follows since the opposite inequality applies to -1 □

2 Social Choice Theory

2.1 Arrow's Theorem

- N individuals
- A alternatives
- individual preferences $R_i \subset A \times A$ are complete and transitive

Preferences satisfy following axioms

Definition 13 (Social Welfare Function). *Function f aggregates preferences of agents, $R = f(R_1, \ldots R_N) = f(\{R_i\})$.*

Definition 14 (Universal). *Every $\{R_i\}$ is plausible*

Definition 15 (Rational). *R is rational (complete and transitive)*

Definition 16 (Unanimous).

$$\forall a, b \in A \quad aP_ib \quad \forall i \in N \Rightarrow aPb$$

Definition 17 (Independent (IIA)). *Given $\{R_i\}$ $\{R_i^{'}\}$ and $a, b \in A$*

$$R_i|_{\{a,b\}} = R_i^{'}|_{\{a,b\}} \Rightarrow R|_{\{a,b\}} = R^{'}|_{\{a,b\}}$$

Definition 18 (Dictatorial). $\exists i \in N \quad aP_ib \Rightarrow aPb \quad \forall a, b \in A$

Theorem 6 (Arrow). ***If $|N| < \infty$, $|A| > 2$ then Universal, Rational, Unanimous and Independent $\Rightarrow$ Dictatorial***

Proof.

Definition 19. *Decisive A subset $D \subset N$ of individuals is decisive for a over b if aP_ib for all $i \in D$ implies aPb. Let $D(a,b) = \{D \subset N : D \text{ is decisive for } a \text{ over } b\}$. Say that $D \subset N$ is decisive if $D \in D(a,b)$ for every $a, b \in A$. Let $D = \{D \subset N : D \text{ is decisive }\}$.*

By definition, if $D \in D(a,b)$ and $D \subset E$ then $E \in D(a,b)$, too. Hence, if $D \in D$ and $D \subset E$ then $E \in D$, too. Moreover, if $D \in D(a,b)$ then $N \backslash D \notin D(b,a)$. By unanimity, $N \in D$, therefore $D \neq \emptyset$. Moreover, $\{\emptyset\} \notin D$, since otherwise aPb and bPa.

Lemma 1. *If $D \subset N$ is decisive then, for every $S \subset D$, either $S \in D$ or $D \backslash S \in D$.*

Proof. The case where S or $D \backslash S$ is empty is trivial. If $\emptyset \neq S \neq D$ then $|D| > 1$. If $D \backslash S$ is decisive we are done. If not, there is a preference profile $\{R_i\}$ and $a, b \in A$ such that aRb yet bP_ia for all $i \in D \backslash S$, that is, $D \backslash S \notin D(b,a)$ □

Lemma 2. *If $D \backslash S \notin D(b,a)$ then $S \in D(a,c) \cap D(c,b)$ for all $c \in A$ with $a \neq c \neq b$.*

Proof. Since aRb and D is decisive, there must exist a nonempty subset of individuals $E \subset S$ such that aR_ib for every $i \in E$. (Otherwise, bP_ia for all $i \in D$ would imply bPa by virtue of $D \in D$.) Let $c \in A$ be an alternative different from both a and b. □

To prove that $S \in D(a,c)$, choose a preference profile $\{R'_i\}$ such that (i) bP'_ic for every $i \in D$, so $bP'c$ since D is decisive, and (ii) R'_i agrees with R_i over $\{a,b\}$ for every individual i, so $aR'b$ by independence. Look at figure By rationality, $aR'bP'c$ implies $aP'c$. Using independence again, we may change the preference profile to an arbitrary $\{R''_i\}$ subject to every individual i 's preferences over $\{a,c\}$ remaining the same in R'_i and R''_i while maintaining the social preference $aP''c$. But only the preferences of the members of E over $\{a,c\}$ were specified, therefore $E \in D(a,c)$. Since $E \subset S$, it follows that $S \in D(a,c)$. To prove that $S \in D(c,b)$, choose a preference profile $\{R'_i\}$ such that (i) cP'_ia for every $i \in D$, and (ii) R'_i agrees with R_i over $\{a,b\}$ for every individual $i.S \in D(c,b)$ now follows by the same argument as for $S \in D(a,c)$.

Lemma 3. *If $S \in D(a,b)$ for some $a, b \in A$ then $S \in D$.*

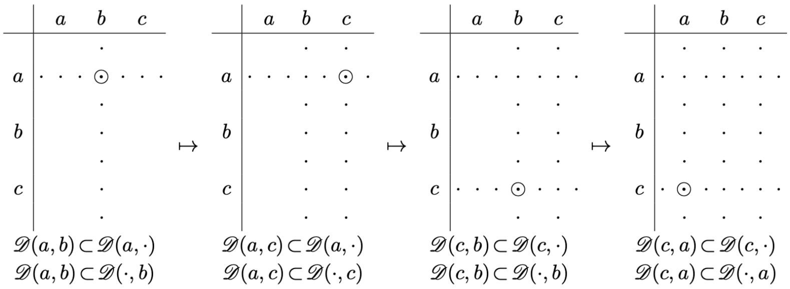

Figure 1:

Proof. If $S \in D(a,b)$ then $D \backslash S \notin D(b,a)$. By the previous lemma, $S \in D(a,c) \cap D(c,b)$ for all $c \in A$ with $a \neq c \neq b$. Applying the lemma again, $S \in D(a,d) \cap D(d,c)$ for d such that $a \neq d \neq c$ and $S \in D(c,e) \cap D(e,b)$ for e such that $b \neq e \neq c$. This implies that $S \in D(a,b)$ for every pair of alternatives a and b, that is, $S \in D$. □

The proposition now follows: if $D \backslash S \notin D(b,a)$ for some $a, b \in A$ then $S \in D(a,c)$ for some alternative c by first of abovementioned lemmas , which implies that $S \in D$ by Lemma above .

Finally we can prove Arrow's Theorem:

By unanimity, $N \in D$. Let $N_1 = N$ and for each $k \in \mathbb{N}$, if $|N_k| > 1$ pick a strict nonempty subset N_{k+1} of N_k such that $N_{k+1} \in D$. If $|N_k| = 1$, let $N_{k+1} = N_k$. Since there are finitely many individuals, $N_{|N|}$ is a singleton consisting of the dictator. □

2.2 Gibbard Satterthwaite Theorem

Assume $|N| < \infty$, $|A| < \infty$

Definition 20 (Social Choice Function (SCF)). *$f : \mathcal{L}^N \to A$ where $\mathcal{L}$ is set of strict linear orders.*

Definition 21. *Preference profile $P = (P_1, \dots P_N) \in \mathcal{L}^N$, $f(P) \in A$*

A SCF f is :

Definition 22 (Pareto Efficient (PE)). *If $f(P) = a$ whenever $a \in A$ is at the top $\forall i P_i$*

Definition 23 (Dictatorial). *If $\exists i$ s.t. $f(P) = a \iff a$ is at top of P_i*

Definition 24 (Strategy proof (SP)).

$$If \quad i \in N, P \in \mathcal{L}^N P_i' \mathcal{L} \quad f(P_i', P_{-i}) \neq f(P) \quad \Rightarrow f(P) P_i f(P_i', P_{-i})$$

Strategy proofness is our first exposure to incentive compatibility. Suppose that a SCFn is executed by first asking all individuals in society to report their preferences and then choosing an alternative according to the reported preference profile. Strategy proofness means that nobody has an incentive to lie about their preference given the preference profile that everyone else reported. In other words, the SCF delivers an alternative that is preferred by every individual under their true preference than that delivered by any other of their possible preferences.

Definition 25 (Maskin Monotone (MM)). *If whenever $f(P) = a$ and $\forall i, b$ P_i' ranks a above b if P_i does then $f(P') = a$. In other words, define $B(a, P_i) = \{b \in A | a P_i b\}$*

$$f(P_1, \dots P_N) = a \quad and \quad i \quad B(a, P_i) \subset B(a, P_i') \Rightarrow f(P_1', \dots P_N') = a$$

Theorem 7 (Gibbard Satterthwaite). ***If $|A| > 2$, $f : \mathcal{L}^N \to A$ is onto and SP $\Rightarrow f$ is Dictatorial***

Gibbard (1973), Satterthwaite (1975)

Corollary 1. *Dictatorial $\Rightarrow$ SP and onto*

Proof. PE $\Rightarrow$ onto

PR and SP $\Rightarrow$ dictatorial

TBD □

3 Mechanisms of Mechanism Design

In this section will deconstruct following powerful picture:

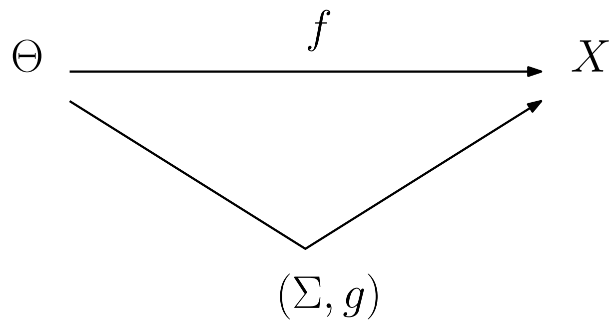

Figure 2: Hurwicz mechanism design diagram

In SCT we studied aggregation of preferences but here preferences are not publicly observable. Individuals must be relied upon to reveal private information. We study here how this information can be elicited and the extent to which the information revelation problem constraints the ways in which social decisions can respond to individual preferences. This is **mechanism design problem**.
The interpretation of x, θ and t (up to sign adjustments) in next examples are as follows

1. **Price discrimination** x is the consumer's purchase and t is the price paid to the monopolist; θ indexes the consumer's surplus from consumption

2. **Income tax** x is the agent's income, and t is the amount of tax paid by the agent; θ is a technological parameter indexing the cost function

3. **Public good provision** x is the amount of public good supplied and t_i is i's consumer monetary contribution to its financing;θ_i is consumer i surplus from the public good

4. **Auction** x_i is the probability that consumer i buys the good ($\sum_i x_i = 1$) and t_i is the amount paid by consumer i, θ_i is consumer i willingness to pay for the good that is auctioned off

5. **Bargaining** x is the quantity sold by a seller to buyer, t_1 is the transfer to the seller and t_2 is negative transfer to buyer, s.t. $t_1 + t_2 = 0$, $\theta_1 = c$ indexes the seller's cost of producing the good and $\theta_2 = v$ is buyer's willingness to pay

3.1 Three examples

In this section we present three canonical examples of mechanism design which show following issues

- Example 1: There are SCF which are not Nash implementable
- Example 2: Allocation rules are monotone. There are transfers which are not truthfully implementable.
- Example 3: in 2nd price auction SCF is truthfully implementable in DS by direct mechanism

Example 1. ***King Solomon's dilemma***
Two mothers Ann and Beth came to King Solomon with a baby and both claimed to be the baby's genuine mother. King Solomon faced the problem of finding out which of two women was the true mother of the baby.
He proposes to split the baby in half and give one half to each woman. This macabre proposal prompts one of the women to scream while the other remains silent. King Solomon decrees that the true mother would never stand by while her baby was murdered, and thus gives the baby to woman who screamed.

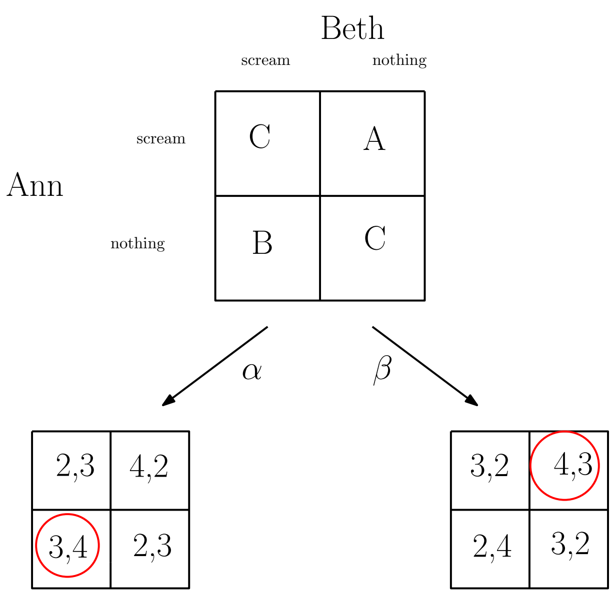

Figure 3: Γ_1

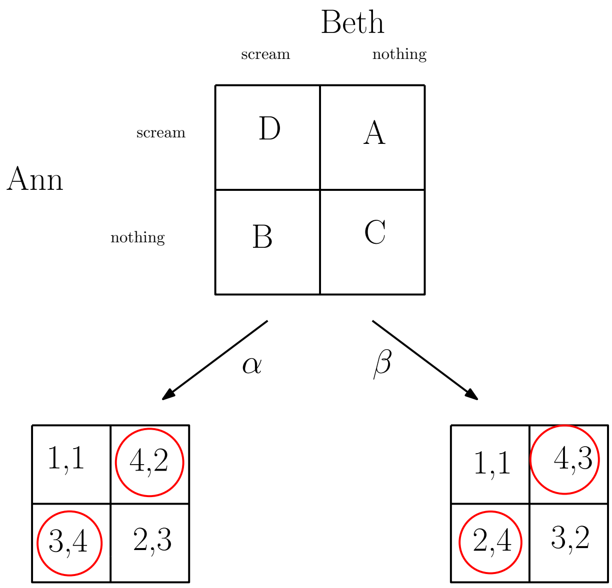

Figure 4: Γ_2

Furthermore he proposed two solutions (remember that it was 10th century BCE kings were cruel back then):

- Γ_1 *give baby to mother which cried, cut the kid in half when neither of them cried or if both of them screamed -Figure 3*
- Γ_2 *give baby to mother which cried, cut the kid in half when neither of them cried and if both screamed kill both women -Figure 4*

Let's define components of the problem:

- *agents* $N = \{Ann, Beth\}$
- *outcomes* $X = \{A, B, C, D\}$
  - $A =$ *give baby to Ann*
  - $B =$ *give baby to Beth*
  - $C =$ *cut the baby to halves*
  - $D =$ *death to everyone*
- *types* $\Theta = \{\alpha, \beta\}$
- *preference profiles* $\Omega = \{P^\alpha, P^\beta\}$

- *in state α their preferences are*

$$\begin{array}{ll} Ann & A \succ B \succ C \succ D \\ Beth & B \succ C \succ A \succ D \\ utils & 4 \succ 3 \succ 2 \succ 1 \end{array}$$

- *in state β their preferences are*

$$\begin{array}{ll} Ann & A \succ C \succ B \succ D \\ Beth & B \succ A \succ C \succ D \\ utils & 4 \succ 3 \succ 2 \succ 1 \end{array}$$

- Γ_i *are a sort of meta-game that induces two NFG corresponding to the states α and β.*

- *Let $NE(\Gamma, \theta)$ denote the set of Nash equilibria in the game induced by Γ when the state of the world is θ.*

- *We say that Γ fully implements f^* in Nash equilibrium (aka Nash implements) if $g(NE(\Gamma, \alpha)) = \{f^*(P^\alpha)\}$ and $g(NE(\Gamma, \beta)) = \{f^*(P^\beta)\}$.*

- *We say that Γ truthfully implements f^* in Nash equilibrium (aka truthful Nash implements) if $g(NE(\Gamma, \alpha)) \ni f^*(\alpha)\}$ and $g(NE(\Gamma, \beta)) \ni f^*(P^\beta)$.*

So which mechanism was played? *Let's take a look at Nash equilibria in all 4 games. Consider cases corresponding to types α and β:*

Case 1 *The game induced by Γ_1 when the state is α.*

In the game α, screaming strictly dominates doing nothing for the fake mother Beth and the true mother's best response when Beth screams is to do nothing. Thus $NE(\Gamma, \alpha) =$ (nothing,scream), so $g(NE(\Gamma, \alpha)) = B \neq f^*(P^\alpha) = A$. In other words, when Anna is the true mother, the mechanism designed by King Solomon causes him to end up allocating the baby to the fake mother Beth.

Similarly, in the game β, screaming strictly dominates doing nothing for the fake mother Anna and doing nothing is a best response for the true mother Beth. Thus $NE(\Gamma, \alpha) =$ (nothing, scream) and $g(NE(\Gamma, \alpha)) = A \neq f^*(P^\beta) = B$, i.e., the fake mother Anna gets the baby.

This means that King Solomon's mechanism does exactly the opposite of what he was intending.
In the parlance of mechanism design, Γ *does not Nash-implement* f^*.

Case 2
The game induced by Γ_2 *when the state is* α *we have two NE in pure strategies* $NE(\Gamma, \alpha) =$ {(*nothing,scream),(scream, nothing)*}, *so* $g(NE(\Gamma, \alpha)) = \{B, A\} \neq f^*(P^\alpha) = \{A\}$ *In other words, when Anna is the true mother, the mechanism designed by King Solomon causes him to end up allocating the baby either to true mother Ann or to the fake mother Beth.*

Similarly, in the game β, *we have two NE in pure strategies* $NE(\Gamma, \alpha) =$ {(*nothing,scream),(scream, nothing)*}, *so* $g(NE(\Gamma, \alpha)) = \{B, A\} \neq f^*(P^\beta) = \{B\}$. *Thus* Γ_2 *does not Nash implements, though it implements it in truthfully in Nash equilibrium.*

It must be something that Nash implements f. *It turns out that there is necessary condition for Nash implementability Let's look at this problem from Mechanism Design perspective.*

Formally problem consists of *(let's skip index i taking* $i = 1$*):*

- *agents* $N = \{Ann, Beth\}$
- *outcomes* $X = \{A, B, C, D\}$
- *types* $\Theta = \{\alpha, \beta\}$
- *type dependent preference profiles* $\Omega = \{P^\alpha, P^\beta\}$ *in short*
- *King Solomon has social choice function* $f^* : \Omega \to X$ *(but we will write in short* Θ *keeping in mind that we track preferences)*
- *Social Choice Function is such that* $f^* : \Theta \to X$ *such that* $f^*(P^\alpha) = A$ *and* $f^*(P^\beta) = B$.
- *To impose that SCF king Solomon introduces the mechanism* $\Gamma = (\Sigma, g)$
- *where* $\Sigma = \Sigma_A \times \Sigma_B$ *is an action (message) space*
- $g : \Sigma \to X$ *is outcome rule that determines which alternative in* X *is chosen based on the actions of the players.*

- *Notice that this is not really game- each induces 2 games!*
- *This is game form (aka **mechanism**) each becomes a game when coupled with a preference profile.*
- *mechanism induces 2 games* (Γ, P^{α}), (Γ, P^{β})
- $NE(\Gamma, P) =$ *set of pure strategy Nash equilirbia*
- *we look at Nash equilibria of* $NE(\Gamma, \alpha)$ *and* $NE(\Gamma, \beta)$
- $g(NE(\Gamma, P)) =$ *set of Nash equilibrium outcomes*
- *in example we saw:* $g(NE(\Gamma, P^{\alpha})) = \{b\}$, $g(NE(\Gamma, P^{\beta})) = \{a\}$
- *We say that* Γ *fully implements* f^* *in Nash equilirbium (aka Nash implements) if* $g(NE(\Gamma, \alpha)) = \{f^*(P^{\alpha})\}$ *and* $g(NE(\Gamma, \beta)) = \{f^*(P^{\beta})\}$.
- *neither seems to be what was played*

Precise definitions are presented later on.
It turns out that this SCF does not satisfy necessary condition for Nash implementability (which is Maskin monotonicity). As The Rolling Stones sang: You can't always get what you want!

Example 2. ***Provision of public good***

Society consists of agents decides about conducting public investment. It needs to set up transfers between its members.

Let's introduce environment and necessary definitions

Set-up

1. *N agents*

2. *each agent has private type* $\theta_i \in \Theta_i \quad \theta = (\theta_1, \ldots, \theta_N) \in \prod_{i=1}^N \Theta_i$

3. *agent i has utility* $u_i(k, \theta_i) = v_i(k, \theta_i) + t_i$

4. *we even specify it to* $v_i(k, \theta_i) = \theta_i k$ *so*

$$u_i(k, \theta_i) = \theta_i k + t_i$$

5. $k \in K$ *where K is the set of non-monetary allocation, let's take* $K = \{0, 1\}$

6. $t_i \in \mathbb{R}$ *monetary transfer*

7. *The allocation is* (k, t) *where* $t = (t_1, \ldots, t_N)$

8. *Set of feasible allocation* $X = \{(k, t) \quad | \, k \in K, t_i \in \mathbb{R}, \quad \sum_i t_i \leq 0\}$ *;*

9. *Let* $k : \Theta \to K$*, where* $\Theta = \Theta_1, \cdots \times \Theta_N = [0, 1]^N$

10. $t_i : \Theta \to \mathbb{R}$*,*

11. *The social choice function is* $f(\theta) = (k(\theta), t(\theta))$ *where* $t(\theta) = (t_i(\theta))_{i=1}^n$.

12. *we are looking for* $\Gamma = (M, g)$ *s.t.* Γ *implements f in Dominant Strategies*

To wrap it up f is dominant strategy incentive compatible (DSIC) or strategy-proof if $\forall i, \theta_i, \theta_i', \theta_{-i}$,

$$k(\theta)\theta_i + t_i(\theta) \geq k\left(\theta_i', \theta_{-i}\right)\theta_i + t\left(\theta_i', \theta_{-i}\right)$$

The first thing we want to know is what kind of allocation functions are implementable? We borrow from auction setting and we can think of K as

$$K = \left\{(y_1, \ldots, y_N) : \forall i, y_i \in \{0, 1\}, \sum_{i \in N} y_i = 1\right\}$$

Example. Consider $N = 2$

Suppose that $N = 2$ *and* k *is such that for some* θ_2 *and* $\theta_1 < \theta_1', y_1(\theta_1) = 1$ *and* $y_1(\theta_1') = 0$
Suppose that $k(\cdot)$ *is implementable. Then there exists* $t(\cdot)$ *such that*

$$\theta_1 + t_1(\theta_1, \theta_2) \geq 0 + t_1(\theta_1', \theta_2)$$

and

$$0 + t_1(\theta_1', \theta_2) \geq \theta_1' + t_1(\theta_1, \theta_2)$$

Adding these inequalities together, we get $\theta_1 \geq \theta_1'$ *which is a contradiction. This leads us to the observation that* $k(\cdot)$ *is implementable only if* $\forall i \in I$ *and* $\forall \theta_{-i} \in \Theta_{-i}, \exists \bar{\theta}_i$ *such that*

$$y_i(\theta_i, \theta_{-i}) = \begin{cases} 1 & \theta_i > \bar{\theta}_i \\ 0 & \theta_i < \bar{\theta}_i \end{cases}$$

We call this type of allocation function a monotone allocation function. This naturally leads one to ask if all monotone allocation functions are implementable. Answer to this is positive what will be presented later.

Finally we want to have some notion of efficiency, control over level of transfers:

Definition 26. *A social choice function* f *satisfies full ex-post efficiency if:*

1. *it is ex-post efficient:* $\sum_{i=1}^{n} v_i(k(\theta), \theta_i) \geq \sum_{i=1}^{n} v_i(k, \theta_i)\ \forall k, \theta;$

2. *it is balanced budget:* $\sum_{i=1}^{n} t_i(\theta) = 0$

Let's solve it in more general set up with N agents and with $0 < c < N$.
In this case condition for ex-post efficiency is equivalent to:

$$\sum_{i=1}^{N} k(\theta)\theta_i \geq \sum_{i=1}^{N} k \cdot \theta_i \quad \forall k \in \{0, 1\}, \theta$$

which is equivalent to

$$(k(\theta) - k) \cdot \sum_{i=1}^{N} \theta_i \geq 0 \quad \forall k \in \{0, 1\}, \theta$$

Consider following SCF $f(\theta) = (k(\theta), t_1(\theta), \ldots, t_I(\theta))\ \forall \theta$,

$$k(\theta) = \begin{cases} 1 & \textit{if } \sum_i \theta_i \geq c \\ 0 & \textit{otherwise} \end{cases}$$

with

$$\sum_i t_i(\theta) = \begin{cases} -c & \text{if } k(\theta) = 1 \\ 0 & \text{otherwise} \end{cases}$$

$k(\theta)$ is either one or zero.

Fix $\sum_i \theta_i \geq c$, in this case 1, then EPIC is satisfied. Analogous $\sum_i \theta_i < c$ with $k \in \{0, 1\}$ is obviously satisfied.

Example. Equal transfers

Suppose that the agents want to implement this ex post efficient SCF with average transfer,i.e.

$$t_i(\theta) = -ck(\theta)/n$$

Suppose that $\Theta_i = \{\bar{\theta}_i\}$ for $i \neq 1$ and $\Theta_1 = [0, +\infty)$ and assume $c > \sum_{i \neq 1} \bar{\theta}_i > c(n-1)/n$.

This implies that, firstly, with this social function agent 1's type is critical for whether the bridge is built (note that if $\theta_1 \geq c - \sum_{i \neq 1} \bar{\theta}_i$, it is. while if $\theta_1 < c - \sum_{i \neq 1} \bar{\theta}_i$, it is not), and, secondly, that the sum of the utilities of the agents $2, \ldots, n$ is strictly greater if the bridge is built than if it is not built (since $\sum_{i \neq 1} \bar{\theta}_i - c(n-1)/I > 0$)

What are the incentives of agent 1 to truthfully reveal her type when $\theta_1 = c - \sum_{i \neq 1} \bar{\theta}_i + \varepsilon$ for some $\varepsilon > 0$?

If agent 1 reveals her true preferences, then the bridge will be built because

$$\left(c - \sum_{i \neq 1} \bar{\theta}_i + \varepsilon\right) + \sum_{i \neq 1} \bar{\theta}_i > c$$

In this case agent 1 's utility is

$$\theta_1 - \frac{c}{n} = \left(c - \sum_{i \neq 1} \bar{\theta}_i + \varepsilon\right) + 0 - \frac{c}{n} = \left(\frac{c(n-1)}{n} - \sum_{i \neq 1} \bar{\theta}_i + \varepsilon\right) + 0$$

However, for $\varepsilon > 0$ small enough, the utility of agent 1 is less than 0 (in fact, this is her utility if she instead claims that $\theta_1 = 0$, a claim that results in the bridge not being built). Therefore, agent 1 will not truthfully reveal her type.

Under this allocation rule, when agent 1 causes the bridge to be built he has a positive externality on the other agents. Because he fails to internalize this effect, he has an incentive *to understate his benefit from this project.*

We have considered just a couple of examples which illustrate the two issues:

- *monotonicity of allocation in agent types*
- *not truthful revelation of private information*

The central question that we impose is the following: What social choice functions can be implemented when agents' types are private information?
In general, we need to consider not only the possibility of directly implementing social choice function by asking agents to reveal their types but also their indirect implementation through the design of institutions in which the agents interact.

Example 3. ***Allocation of single unit of indivisible private good***
There is a single unit of an indivisible private good to be allocated to one of N agents. Monetary transfers can also be made.

Here $\theta_i \in \mathbb{R}$ can be viewed as agent i 's valuation of the good, and we take the set of possible valuations for agent i to be $\Theta_i = [\underline{\theta}_i, \bar{\theta}_i] \subset \mathbb{R}$.

*Two special cases ubiquitous in the literature deserve mention. The first is the case of **bilateral trade**. In this case we have $N = 2$ agent 1 is the seller and agent 2 is the buyer. Consider cases*

- *When $\underline{\theta}_2 > \bar{\theta}_1$ there are certain to be gains from trade regardless of the realizations of θ_1 and θ_2*
- *when $\underline{\theta}_1 > \bar{\theta}_2$ there are certain to be no gains from trade*
- *finally, if $\underline{\theta}_2 < \bar{\theta}_1$ and $\underline{\theta}_1 < \bar{\theta}_2$ then there may or may not be gains from trade, depending on the realization of θ*

*The second special case is the **auction** setting. Here, one agent, whom we shall designate as agent 0 is interpreted as the seller of the good (the auctioneer) and is assumed to derive no value from it (more generally, the seller might have a known value $\theta_0 = \bar{\theta}_0$ different from zero). The other agents, $1, \ldots, I$, are potential buyers (the bidders). Let's present auction in form of 2nd price sealed auctioned aka Vickery auction.*

Definition 27 (2nd price auction)**.** *is such incomplete information game in which*

- *Bidders are asked to submit sealed bids $b_1(\theta_1), \ldots, b_N(\theta_N)$. The bidder who submits the highest bid is awarded the object, and pays the amount of the second highest bid.*
- $K = \{0, 1, \ldots N\}$
- *k- who gets the object if anyone*
- *utility*
$$u_i(k, \theta_i) = \theta_i k + t_i$$
$$v_i(k, \theta_i) = \begin{cases} \theta_i & \text{if } k = i \\ 0 & \text{otherwise} \end{cases}$$
- *utility of auctioneer $u_0 = t_1 + t_2$*

- *define transfers in following way*

$$t_1(\theta) = -\theta_2 k(\theta)$$

$$t_2(\theta) = -\theta_1 k(\theta)$$

$$t_0(\theta) = -\left(t_1(\theta) + t_2(\theta)\right)$$

- *SCF is $f(\theta) = (k(\theta), t(\theta))$*

Lemma 4. *In a second price auction, it is a weakly dominant strategy to bid one's value, $b_i(\theta_i) = \theta_i$*

Proof. Suppose i 's value is θ_i, and she considers bidding $b_i > \theta_i$. Let $\hat{b}$ denote the highest bid of the other bidders $j \neq i$ (from i 's perspective this is a random variable). Consider three possible outcomes from i's perspective:

1. $\hat{b} > b_i, \theta_i$
2. $b_i > \hat{b} > \theta_i$
3. $b_i, \theta_i > \hat{b}$

- In the event of the 1st or 3rd outcome i would have done equally well to bid θ_i rather than $b_i > \theta_i$
- In 1st won't win regardless, and in 2nd she will win, and will pay $\hat{b}$ anyway.
- However, 2nd case i will win and pay more than her value if she bids $\hat{b}$, something that won't happen if she bids θ_i. Thus, i does better to bid $s\theta_i$ than $b_i > \theta_i$.
- A similar argument shows that i also does better to bid θ_i than to bid $b_i < \theta_i$

□

Since everyone is bidding their true value, seller will receive second highest value. The truthfull equilibrium described above is unique symmetric Bayesian Nash equilibrium .

Thus, this social choice function is implementable even though the buyers' valuations arc private information: it suffices to simply ask each buyer to report his type, and then to choose $f(\theta)$ We will show later that :

- *$k^*(\theta)$ ex post efficient allocation give it to whoever values it most*

- *i is pivotal when she is the highest bidder. His tax is the 2nd highest bid.*

Now in general one can ask: What SCF can be implemented directly when agents types are private information? And given SCF which institutions (mechanisms) may guarantee indirect implementation?

3.2 Game theory and Social Choice Theory kicks in

In this section we provide a notation and results on intersection of mechanism design, social choice theory and game theory.

Definition 28 (Mechnism design problem)**.** *consists of:*

- *finite set of agents* $I = \{1, \dots N\}$
- *agents take collective choice from set of possible alternatives* X
- *each agent has private type* $\theta_i \in \Theta_i \quad \theta = (\theta_1, \dots, \theta_N) \in \prod_{i=1}^{N} \Theta_i$
- *types are privately observed before collective choice*
- *agent* i *has utility* $u_i(x, \theta_i)$
- *agents are assumed to be an expected utility maximizers*
- ϕ *pdf over* Θ*,* Θ *and* $u_i(\cdot, \theta_i)$ *are common knowledge but specific values of each agent* i *are observed only by* i
- *collective action depend on* θ

Definition 29 (Social choice function)**.** $f : \Theta_1 \times \dots \times \Theta_N \to X$ *chooses an outcome* $f(\theta) \in X$*, given types* $\theta = (\theta_1, \dots, \theta_N)$*.*

Definition 30 (Ex post efficiency)**.** *The social choice function* $f : \Theta_1 \times \cdots \times \Theta_N \to X$ *is* ***ex post efficient*** *(EPE or Paretian) if for no profile* $\theta = (\theta_1, \dots, \theta_N)$ *is there an* $x \in X$ *such that*

$$\forall i \quad u_i(x, \theta_i) \geq u_i(f(\theta), \theta_i)$$

$$\exists j \quad u_j(x, \theta_j) > u_j(f(\theta), \theta_j)$$

SCF is ex post efficient if it selects, for every profile θ an alternative $f(\theta) \in X$ that is Pareto optimal given the agents' utility functions $u_1(\cdot, \theta_1), \dots, u_I(\cdot, \theta_I)$ The problem is that the θ_i 's are not publicly observable, and so for the social choice f to be chosen when the agents' types are $(\theta_1, \dots, \theta_N)$, each agent i must be relied upon to disclose his type θ_i. However, for a given social choice function f an agent may not find it to be in his best interest to reveal this information truthfully. We illustrated this information revelation problem in Example 2. In Example 3 we showed that in auction setting agent reveal information truthfully.

The mechanism design problem is to implement rules of a game or meta game by defining possible strategies and the method used to select an outcome based on agent strategies, to implement the solution to the social choice function despite agent's selfinterest.

Definition 31 (Mechanism)**.** *$\mathcal{M} = (\Sigma, g)$ consists of*

- *set of strategies $\Sigma = \Sigma_1 \times \ldots \times \Sigma_N$*
- *an outcome rule $g : \Sigma_1 \times \ldots \times \Sigma_N \to X$*
- *such that $g(s)$ is the outcome implemented by the mechanism for strategy profile $s = (s_1, \ldots, s_N)$.*

In words, a mechanism defines the strategies available (e.g., bid at least the ask price, etc.) and the method used to select the final outcome based on agent strategies (e.g., the price increases until only one agent bids, then the item is sold to that agent for its bid price).

Let's define remaining parts of incomplete information game: strategies and equilibrium concepts:

Definition 32 (Strategy)**.** *A strategy of player i is a mapping $s_i : \Theta_i \to \Sigma_i$.*

Definition 33 (Preferences)**.** *of agent i are function of outcome and private type:*

$$u_i(s_i, \theta_i) : \Sigma_i \times \Theta \to \mathbb{R}$$

Formally the mechanism $\mathcal{M}$ induces Bayesian game of incomplete information

Definition 34 (Bayesian game of incomplete information)**.** *is $\{I, \{S_i, \bar{u}_i(\cdot)\}, \Theta, \phi\}$ where we have:*

- *finite set of agents $I = \{1, \ldots N\}$*
- *strategy sets Σ_i*
- *$\bar{u}_i(s_1, \ldots, s_N, \theta_i) = u_i(g(s_1, \ldots, s_N), \theta_i)]$*

- *expected payoff (expectation with ϕ as pdf):*

$$\hat{u}_i(s_1(\theta_1), \ldots, s_N(\theta_N)) = \mathbb{E}_\theta[\bar{u}_i(s_1(\theta_1), \ldots, s_N(\theta_N), \theta_i)]$$

Game theory is used to analyze the outcome of a mechanism. Given mechanism $\mathcal{M}$ with outcome function $g(\cdot)$, we say that a mechanism implements social choice function $f(\theta)$ if the outcome computed with equilibrium agent strategies is a solution to the social choice function for all possible agent preferences.

Definition 35 (Mechanism implementation). *$\mathcal{M} = (\Sigma_1, \ldots, \Sigma_N, g)$ implements social choice function $f(\theta)$ in Z -equilibrium if $\exists$ a Z -equilibrium s^* such that*

$$\forall \theta \in \Theta \quad \mathbb{P}(\theta) > 0, \quad g\left(s_1^*\left(\theta_1\right), \ldots, s_N^*\left(\theta_N\right)\right) = f(\theta)$$

where strategy profile $(s_1^*, \ldots, s_N^*)$ is an equilibrium solution of game. induced by $\mathcal{M}$.

As an Z- equilibrium concept we will consider one of following

- Nash
- Dominant Strategy
- Bayes-Nash

Note that we allow for multiple equilirbium strategy profiles from which at least one satisfy condition from definition above.

Definition 36. *Γ strongly implements f in Z -equilibrium there exists a Z -equilibrium, and for all Z -equilibrium s^* and for all $\theta \in \Theta$ with $\mathbb{P}(\theta) > 0, g\left(s^*\left(\theta_i\right)_{i \in I}\right) = f(\theta)$.*

Picking right mechanism is a daunting task. Looking at space of whole possible mechanism may be exhausting task. The mechanism asks agents to report their types, and then simply implements the solution to the social choice function that corresponds with their reports. Very naive mechanisms gives no good reason for self-interested to truthfully report their types.

Definition 37 (Direct Mechanism). *Given SCF f, the direct mechanism is Γ_{direct} where $\Sigma_i = \Theta_i$ and $g = f$.*

Note that all other mechanisms are indirect.

3.2.1 Incentive Compatibility

Definition 38 (Incentive Compatible)**.** *The SCF f is truthfully implementable (or* ***Incentive Compatible****) if the direct revelation mechanism $\Gamma = (\Theta_1, \ldots, \Theta_N, f)$ has an equilibrium $(s_1^*(\cdot), \ldots, s_N^*(\cdot))$ in which*

$$s_i^*(\theta_i) = \theta_i \quad \forall i, \theta_i \in \Theta_i$$

That is, if truth telling by each agent i constitutes an equilibrium of $\Gamma = (\Theta_1, \ldots, \Theta_N, f)$.

Notion of Incentive Compatibility was introduced to economics by Minnesota faculty member Leonid Hurwicz (2007 Nobel prize winner), check Hurwicz(1972, 1976).

Example 4. *By keeping allocation in Example 3 the same and changing only transfers to*

$$t_1(\theta) = -\theta_1 y_1(\theta)$$

$$t_2(\theta) = -\theta_2 y_2(\theta)$$

In this social choice function, the seller gives the good to the buyer with the highest valuation (to buyer 1 if there is a tie) and this buyer gives the seller a payment equal to his valuation (the other, low-valuation buyer makes no transfer payment to the seller). Note that $f(\cdot)$ is not only ex post efficient but also is very attractive for the seller: if $f(\cdot)$ can be implemented, the seller will capture all of the consumption benefits that are generated by the good.

Suppose we try to implement this social choice function. Assume that the buyers are expected utility maximizers. We now ask: If buyer 2 always announces his true value, will buyer 1 find it optimal to do the same? For each value of θ_1, buyer l's problem is to choose the valuation to announce, say $\hat{\theta}_1$, so as to solve

$$\text{Max}_{\hat{\theta}_1} \left(\theta_1 - \hat{\theta}_1\right) \text{Prob}\left(\theta_2 \leq \hat{\theta}_1\right)$$

or

$$\text{Max}_{\hat{\theta}_1} \left(\theta_1 - \hat{\theta}_1\right) \hat{\theta}_1$$

The solution to this problem has buyer 1 set $\hat{\theta}_1 = \theta_1/2$. We see then that if buyer 2 always tells the truth, truth telling is not optimal for buyer 1. A similar point applies to buyer 2. Intuitively, for this social choice function, a buyer has an incentive to understate his valuation so as to lower the transfer he must make in the event that he has the highest announced *valuation and gets the good. Thus, we again see that there may be a problem in implementing certain social choice functions in settings in which information is privately held.*

In a moment we will see revelation principle.Because of the revelation principle, the constraints that incomplete information about types puts on the set of implementable social choice functions, we will be able to restrict our analysis to identifying those SCF that can be truthfully implemented.

Finally, we note that, in some applications, participation in the mechanism may be voluntary, and so a social choice function must not only induce truthful revelation of information but must also satisfy certain participation (or individual rationality) constraints if it is to be successfully implemented.

3.2.2 Dominant Strategy Implementation

Recall that a strategy is weakly dominant strategy if gives a player at least as large a payoff as any of his other possible strategies for every possible strategy that his rival migh play.

Definition 39 (Dominant Strategy Equilibrium)**.** *s^* is a DS. equilibrium: $\forall i \in I, \forall \theta_i \in \Theta_i, \forall \sigma_i \in \Sigma_i, \forall \theta_{-i} \in \Theta_{-i}, \forall s_{-i}$*

$$u_i(g\left(s^*\left(\theta_i\right), s_{-i}\right), \theta_i) \geq u_i(g\left(\sigma_i, s_{-i}\right), \theta_i)$$

Condition above is equivalent to

$$\mathbb{E}_{\theta_{-i}}[u_i(g\left(s^*\left(\theta_i\right), s^*_{-i}\left(\theta_{-i}\right)\right), \theta_i)|\theta_i] \geq \mathbb{E}_{\theta_{-i}}[u_i(g\left(\sigma_i, s^*_{-i}\left(\theta_{-i}\right)\right), \theta_i)|\theta_i]$$

Definition 40 (Implementation in DS)**.** *Γ implements f in DS if $\exists s^*$ such that*

- *s^* is a DS equilibrium*
- *Implementation: $\forall \theta \in \Theta, g\left(s^*(\theta)\right) = f(\theta)$.*

The concept of dominant strategy implementation is of special interest because if we can find a mechanism $\Gamma = (\Sigma_1, \ldots, \Sigma_N, g)$ that implements f in dominant strategics, then this mechanism implements f in a very strong and robust way.
This implementation will be robust even if agents have incorrect, and perhaps even contradictory, beliefs about this distribution. In particular, agent i 's beliefs regarding

the distribution of θ_{-i} do not affect the dominance of his strategy s_i^* The same mechanism can be used to implement $f(\cdot)$ for any ϕ. One advantage of this is that if the mechanism designer is an outsider (say, the "government"), he need not know $\phi(\cdot)$ to successfully implement f

Definition 41 (Dominant Strategy Incentive Compatible (DSIC)). *The SCF f is truthfully implementable in dominant strategies (or dominant strategy incentive compatible, or strategy-proof, or* ***DSIC****) if $s_i^*(\theta_i) = \theta_i$ for all $\theta_i \in \Theta_i$ and $i = 1, \ldots, N$ is a dominant strategy equilibrium of the direct revelation mechanism $\Gamma = (\Theta_1, \ldots, \Theta_N, f)$.*
That is, if for all i and all $\theta_i \in \Theta_i$

$$u_i\left(f\left(\theta_i, \theta_{-i}\right), \theta_i\right) \geq u_i\left(f\left(\hat{\theta}_i, \theta_{-i}\right), \theta_i\right) \quad \forall \theta', \theta_{-i}$$

The ability to restrict our inquiry, without loss of generality, to the question of whether $f(\cdot)$ is truthfully implementable is a consequence of what is known as the revelation principle for dominant strategies.

Theorem 8 (Revelation principle for DS equilibrium). . ***If there exists a mechanism that implements f in DS-equilibrium, then f can be implemented in DS equilibrium with the direct mechanism, with truth-telling as the dominant strategy (i.e., f is strategy proof).***

Proof. Let $\Gamma = (\Sigma, g)$ be the mechanism that implements f is d.s. equilibrium. Let σ^* be the d.s. equilibrium such that $g(s^*(\theta)) = f(\theta), \forall \theta \in \Theta$. Then by definition of d.s. equilibrium, $\forall i \in I, \forall \theta_i \in \Theta_i, \forall \sigma_i \in \Sigma_i, \forall \theta_{-i} \in \Theta_{-i}, \forall s_{-i}$

$$u_i(g\left(s_i^*\left(\theta_i\right), s_{-i}\left(\theta_{-i}\right)\right)) \geq u_i(g\left(\sigma_i, s_{-i}\left(\theta_{-i}\right)\right))$$

Since this holds for every possible strategy of the other players, we can substitute s_{-i}^* for s_{-i}. Similarly, since $s_i^*\left(\theta_i'\right) \in \Sigma_i, \forall \theta_i' \in \Theta_i$, we can substitute $s_i^*\left(\theta_i'\right)$ for m_i. Thus we have $\forall i \in I, \forall \theta_i, \theta_i' \in \Theta_i, \forall \theta_{-i} \in \Theta_{-i}$

$$u_i(g\left(s_i^*\left(\theta_i\right), s_{-i}^*\left(\theta_{-i}\right)\right)) \geq u_i(g\left(s_i^*\left(\theta_i'\right), s_{-i}^*\left(\theta_{-i}\right)\right))$$

By definition of s^*, this implies that $\forall i \in I, \forall \theta_i, \theta_i' \in \Theta_i, \forall \theta_{-i} \in \Theta_{-i}$

$$u_i(f(\theta)) \geq u_i(f\left(\theta_i', \theta_{-i}\right))$$

Thus f is strategy-proof, so f can be implemented by the direct mechanism with truth-telling as the d.s. equilibrium. □

The intuitive idea behind the revelation principle for dominant strategies can be put as follows: Suppose that the indirect mechanism Γ implements f in DS, and that in this indirect mechanism each agent i finds playing $s_i^*(\theta_i)$ when his type is θ_i better than playing any other $s_i \in S_i$ for any choices $s_{-i} \in S_{-i}$ by agents $j \neq i$.

Now consider altering this mechanism simply by introducing a mediator who says to each agent i : "You tell me your type, and when you say your type is θ_i, I will play $s_i^*(\theta_i)$ for you." Clearly, if $s_i^*(\theta_i)$ is agent i 's optimal choice for each $\theta_i \in \Theta_i$ in the initial mechanism Γ for any strategies chosen by the other agents, then agent i will find telling the truth to be a dominant strategy in this new scheme. But this means that we have found a way to truthfully implement f. The implication of the revelation principle is that to identify the set of social choice functions that are implementable in dominant strategies, we need only identify those that are truthfully implementable. In principle, for any f, this is just a matter of checking the inequalities above.

Theorem 9 (Gibbard-Satterthwaite once again). ***If*** $|X| \geq 3, f$ ***is onto,*** $\{\succeq_i(\theta) : \theta \in \Theta\} = P$***, and*** f ***is truthfully implementable in DS , then*** f ***is dictatorial, i.e.,*** $\exists i \in I$ ***such that*** $\forall \theta \in \Theta$***,***

$$f(\theta) \in \{x \in X : x \succeq_i (\theta) y, \forall y \in X\}$$

Given this negative conclusion, if we are to have any hope of implementing desirable social choice functions, we must either weaken the demands of our implementation concept by accepting implementation by means of less robust equilibrium notions (such as Bayesian Nash equilibria) or we must focus on more restricted environments.After some results and properties on monotonicity of tax rules, we focus studying the possibilities for implementing desirable social choice functions in dominant strategics when preferences take a quasilinear form.

3.3 Properties of Mechanisms

Here we follow Example 2 and we outline a number of desirable properties for social choice functions $f = (k, t)$.

Theorem 10 (Taxation principle). . ***Suppose*** t ***implements*** k***. If*** $k(\theta_i, \theta_{-i}) = k(\theta_i', \theta_{-i})$,

then $t_i(\theta_i, \theta_{-i}) = t_i(\theta_i', \theta_{-i})$. ***Hence*** t_i ***can be written as***

$$t_i(\theta_i, \theta_{-i}) = \tau(k(\theta_i, \theta_{-i}), \theta_{-i})$$

Proof. Suppose not. Then player i will lie at either θ_i or either θ_i', i.e., he will just say whichever type will give him a higher transfer. Combined with the monotonicity condition, this implies that transfers must take the following form:

$$t_i(\theta_i, \theta_{-i}) = \begin{cases} \alpha & \theta_i < \bar{\theta}_i \\ \beta & \theta_i > \bar{\theta}_i \end{cases}$$

Further, it must be that $\alpha > \beta$. Otherwise player i will lie when his type is less than $\bar{\theta}_i$. Even more specifically, $\bar{\theta}_i = \alpha - \beta$ since player i must be indifferent between getting the object and not getting it at $\bar{\theta}_i$ (otherwise he will lie).

To summarize, suppose $k(\cdot)$ is montone and suppose $t(\cdot)$ implements $k(\cdot)$. Then $t(\cdot)$ takes the following form:

$$\forall i \in I, \forall \theta_{-i} \in \Theta_{-i}, t_i(\theta_i, \theta_{-i}) = \begin{cases} \alpha & \theta_i < \bar{\theta}_i \\ \beta = \alpha - \bar{\theta}_i & \theta_i > \bar{\theta}_i \end{cases}$$

where $\bar{\theta}$ is a function of θ_{-i}, i.e.

$$\bar{\theta}_i = \bar{\theta}_i(\theta_{-i}) = \inf\{\theta_i \in \Theta_i : y_i(\theta_i, \theta_{-i}) = 1\}$$

□

Theorem 11 (Monotone transfers). $k(\cdot)$ ***is implementable if and only if it is monotone. If it is monotone, then it can be implemented as above.***

Exercise 3. *Let* $k^*(\theta) = (y_1^*(\cdot), \ldots, y_n^*(\cdot))$, *where*

$$y_i^*(\theta) = \begin{cases} 0 & \theta_i < \max_{j \neq i} \theta_j \\ 1 & \theta_i > \max_{j \neq i} \theta_j \end{cases}$$

Note that this describes more than one k^* *since you can break ties in many different ways. We can implement this allocation function using the following transfers:*

$$t_i(\theta, \theta_{-i}) = \begin{cases} 0 & \theta_i < \max_{j \neq i} \theta_j \\ 0 - \max_{j \neq i} \theta_j & \theta_i > \max_{j \neq i} \theta_j \end{cases}$$

---

4 Vickery-Clarke-Groves Mechanisms

Following negative results of Arrow/GS Impossibility theorems. We restrict the domain of preferences to Transferable Utility or Quasi Linear.

We follow the story of public good provision from Example 2.

Set-up

- N agents, each agent has utility $u_i(k, \theta_i) = v_i(k, \theta_i) + t_i$
- $k \in K$ where K is the set of non-monetary allocation
- $t_i \in \mathbb{R}$ monetary transfer
- The allocation is (k, t) where $t = (t_1, \ldots, t_n)$
- $\theta_i \in \Theta_i$ type of agents, $\theta = (\theta_1, \ldots, \theta_N) \in \prod_{i=1}^{N} \Theta_i$
- Set of feasible allocation $X = \{(k, t) \mid k \in K, t_i \in \mathbb{R}, \sum_i t_i \leq 0\}$;
- Let $k : \Theta \to K$, where $\Theta = \Theta_1, \cdots \times \Theta_N = [0, 1]^N$, K
- $t_i : \Theta \to \mathbb{R}$,
- we even specify it to $v_i(k, \theta_i) = \theta_i k$
- The social choice function is $f(\theta) = (k(\theta), t(\theta))$ where $t(\theta) = (t_i(\theta))_{i=1}^{n}$.

Definition 42. *A social choice function f satisfies :*

1. *it is ex-post efficient:* $\sum_{i=1}^{n} v_i(k(\theta), \theta_i) \geq \sum_{i=1}^{n} v_i(k, \theta_i)\ \forall k, \theta$;
2. *it is balanced budget:* $\sum_{i=1}^{n} t_i(\theta) = 0$
3. *full ex-post efficiency: if it is ex-post efficient with balanced budget*
4. *dominant strategy incentive compatible (DSIC) or strategy-proof if* $\forall i, \theta_i, \theta_i', \theta_{-i}$,

$$v_i(k(\theta), \theta_i) + t_i(\theta) \geq v_i\left(k\left(\theta_i', \theta_{-i}\right), \theta_i\right) + t\left(\theta_i', \theta_{-i}\right)$$

Theorem 12. ***Let*** $k^* : \prod_{i=1}^n \Theta_i \to K$ ***be ex-post efficient. If*** $\forall i, \exists h_i : \forall \theta$,

$$t_i(\theta) = \sum_{j \neq i} v_j\left(k^*(\theta), \theta_j\right) + h_i\left(\theta_{-i}\right)$$

then, $f = (k^*, t)$ ***is*** *DSIC.*

Proof. Proof. Suppose, for contradiction, $f = (k^*, t)$ is not DSIC, then, $\exists i, \theta_i, \theta_i', \theta_{-i}$:

$$\begin{gathered} v_i\left(k(\theta), \theta_i\right) + t_i(\theta) < v_i\left(k\left(\theta_i', \theta_{-i}\right), \theta_i\right) + t\left(\theta_i', \theta_{-i}\right) \\ \Rightarrow v_i\left(k(\theta), \theta_i\right) + \sum_{j \neq i} v_j\left(k^*(\theta), \theta_j\right) + h_i\left(\theta_{-i}\right) < \\ < v_i\left(k\left(\theta_i', \theta_{-i}\right), \theta_i\right) + \sum_{j \neq i} v_j\left(k^*\left(\theta_i', \theta_{-i}\right), \theta_j\right) + h_i\left(\theta_{-i}\right) \\ \Rightarrow \sum_{i=1}^n v_i\left(k(\theta), \theta_i\right) < \sum_{i=1}^n v_i\left(k\left(\theta_i', \theta_{-i}\right), \theta_i\right) \end{gathered}$$

It contradicts with ex-post efficiency. □

Definition 43 (Clarke mechanism)**.** *Consider to construct* $h_i : \prod_{j \neq i} \Theta_j \to \mathbb{R}$

$$h_i\left(\theta_{-i}\right) = -\sum_{j \neq i} v_j\left(k_{-i}^*\left(\theta_{-i}\right), \theta_j\right)$$

where $k_{-i}^*\left(\theta_{-i}\right) \in \arg\max_{k \in K} \sum_{j \neq i} v_j\left(k, \theta_j\right) \forall \theta_{-i}$

k_{-i}^* maximizes welfare among everyone but i

Intuitively h_i can be thought of as follows:

- We have seen that when agents type is $\theta_1 \ldots \theta_N$ each agent payment is $\xi_i(\theta_i)$
- Now if each agent contributes an equal $\frac{1}{n-1}$ share of all of the other agent's payments, payments from a given agent i to each other $n-1$ will total $\frac{1}{n-1} \sum_{j \neq i} \xi_j(\theta_j)$ and agent i will receive from these agents in return payment that total to $\xi(\theta_i)$.
- Agent i's net transfer will therefore be $\xi_i(\theta_i) - \frac{1}{n-1} \sum_{j \neq i} \xi_j(\theta_j)$.

Example 5 (Pivotal mechanism)**.**

$$t_i(\theta) = \sum_{j \neq i} \left(v_j\left(k^*(\theta), \theta_j\right) - v_j\left(k_{-i}^*(\theta_{-i}), \theta_j\right)\right)$$

- If $k(\theta) = k_{-i}^*(\theta_{-i})$ then i pays nothing.
- If $k(\theta) \neq k_{-i}^*(\theta_{-i})$ then i pays a tax equal to his effect on others.

This direct revelation mechanism is known as the expected externality mechanism due to d'Aspremont adn Gerard Varret and Arrow.

In an auction setting, this is just like a Vickery auction (2nd price sealed action)

Example 6. *(Vickery auction)*

- $K = \{0, 1, \ldots n\}$
- *k- who gets the object if anyone*
- *utility*

$$v_i(k, \theta_i) = \begin{cases} \theta_i & \textit{if } k = i \\ 0 & \textit{otherwise} \end{cases}$$

- $k^*(\theta)$ *give it to whoever values it most*
- i *is pivotal when she is the highest bidder. His tax is the 2nd highest bid.*

Theorem 13. ***Suppose, for each*** $i, \{v_i(\cdot, \theta_i) \mid \theta_i \in \Theta_i\} = \mathbb{R}^K$. ***If*** $f = (k^*, t)$ ***satisfies ex-post efficiency and*** $DSIC$, ***then,*** $\forall i, \exists h_i : \forall \theta,$

$$t_i(\theta) = \sum_{j \neq i} v_j\left(k^*(\theta), \theta_j\right) + h_i\left(\theta_{-i}\right)$$

Proof. Let $f = (k^*, t)$ satisfies ex-post efficiency and DSIC, and define $h_i(\theta) = t_i(\theta) - \sum_{j \neq i} v_j\left(k^*(\theta), \theta_j\right)$. WTS $h_i(\theta) = h_i\left(\theta_{-i}\right) \forall \theta_i$

1. Suppose $k^*\left(\theta_i, \theta_{-i}\right) = k^*\left(\theta_i', \theta_{-i}\right)$. Then, by DSIC, $\forall i, \theta_i, \theta_i', \theta_{-i}$

$$v_i\left(k^*(\theta), \theta_i\right) + t_i(\theta) \geq v_i\left(k^*\left(\theta_i', \theta_{-i}\right), \theta_i\right) + t\left(\theta_i', \theta_{-i}\right) \Rightarrow t_i(\theta) \geq t_i\left(\theta_i', \theta_{-i}\right)$$
$$v_i\left(k^*\left(\theta_i', \theta_{-i}\right), \theta_i'\right) + t\left(\theta_i', \theta_{-i}\right) \geq v_i\left(k^*(\theta), \theta_i'\right) + t_i(\theta) \Rightarrow t_i\left(\theta_i', \theta_{-i}\right) \geq t_i(\theta)$$
$$\Rightarrow t_i\left(\theta_i, \theta_{-i}\right) = t_i\left(\theta_i', \theta_{-i}\right) = t_i\left(\theta_{-i}\right) \forall \theta_i, \theta_i'$$
$$\Rightarrow h_i(\theta) = h_i\left(\theta_{-i}\right)$$

where the last implication is due to $k^*\left(\theta_i, \theta_{-i}\right) = k^*\left(\theta_i', \theta_{-i}\right)$.

2. Suppose, $\exists i, \exists \theta_i, \theta_i' : k^*(\theta_i, \theta_{-i}) \neq k^*(\theta_i', \theta_{-i})$, and suppose, per contra, wlog, $h_i(\theta_i, \theta_{-i}) > h_i(\theta_i', \theta_{-i})$ Fix an $\epsilon > 0$, and let $\theta_i^\epsilon \in \Theta_i$ be such that:

$$v_i(k, \theta_i^\epsilon) = \begin{cases} -\sum_{j \neq i} v_j(k^*(\theta), \theta_j) & k = k^*(\theta) \\ -\sum_{j \neq i} v_j(k^*(\theta_i', \theta_{-i}), \theta_j) + \epsilon & k = k^*(\theta_i', \theta_{-i}) \\ -\infty & \text{otherwise} \end{cases}$$

First, notice that, $k^*(\theta_i^\epsilon, \theta_{-i}) = k^*(\theta_i', \theta_{-i})$ otherwise ex-post efficiency is violated. By DSIC,

$$v_i(k^*(\theta_i^\epsilon, \theta_{-i}), \theta_i^\epsilon) + t_i(\theta_i^\epsilon, \theta_{-i}) \geq v_i(k^*(\theta_i, \theta_{-i}), \theta_i^\epsilon) + t_i(\theta_i, \theta_{-i})$$

$$\Rightarrow v_i(k^*(\theta_i^\epsilon, \theta_{-i}), \theta_i^\epsilon) + \sum_{j \neq i} v_j(k^*(\theta_i^\epsilon, \theta_{-i}), \theta_j) + h_i(\theta_i^\epsilon, \theta_{-i}) \geq v_i(k^*(\theta), \theta_i^\epsilon) + \sum_{j \neq i} v_j(k^*(\theta), \theta_j) + h$$

$\Rightarrow h_i(\theta_i^\epsilon, \theta_{-i}) + \epsilon \geq h_i(\theta_i, \theta_{-i})$ Then, since $k^*(\theta_i^\epsilon, \theta_{-i}) = k^*(\theta_i', \theta_{-i})$, by part [1], we have $h_i(\theta_i^\epsilon, \theta_{-i}) = h_i(\theta_i', \theta_{-i})$. Hence,

$$h_i(\theta_i', \theta_{-i}) + \epsilon \geq h_i(\theta_i, \theta_{-i})$$

However, when ϵ small enough, it will contradict with the hypothesis $h_i(\theta_i, \theta_{-i}) > h_i(\theta_i', \theta_{-i})$

□

Theorem 14. *Suppose, for each i, $\{v_i(\cdot, \theta_i) \mid \theta_i \in \Theta_i\} = \mathbb{R}^K$. Then, there does not exist a SCF f that is DSIC and fully ex-post efficient.*

Proof. (Easy and incomplete version: assume twice differentiability and $n = 2$.) Assume $K = \mathbb{R}, \theta_i \in [\underline{\theta}_i, \bar{\theta}_i] \subseteq \mathbb{R}, \partial^2 v_i / \partial k^2 < 0$, and $\partial^2 v_i / \partial k \partial \theta_i \neq 0$ FOC for reporting θ_i : $\frac{\partial v_i}{\partial k} \frac{\partial k^*}{\partial \theta_i} + \frac{\partial t_i}{\partial \theta_i} = 0 \Rightarrow -\frac{\partial^2 t_i}{\partial \theta_1 \partial \theta_2} = \frac{\partial^2 v_i}{\partial k^2} \frac{\partial k^*}{\partial \theta_1} \frac{\partial k^*}{\partial \theta_2} + \frac{\partial v_i}{\partial k} \frac{\partial^2 k^*}{\partial \theta_1 \partial \theta_2}$ Balanced budget $\Rightarrow t_1(\theta) + t_2(\theta) = 0, \forall \theta$, then,

$$\left(\frac{\partial^2 v_1}{\partial k^2} + \frac{\partial^2 v_2}{\partial k^2} \right) \frac{\partial k^*}{\partial \theta_1} \frac{\partial k^*}{\theta_2} + \left(\frac{\partial v_1}{\partial k} + \frac{\partial v_2}{\partial k} \right) \frac{\partial^2 k^*}{\partial \theta_1 \partial \theta_2} = 0$$

But, by ex-post efficiency, $\frac{\partial v_1}{\partial k} + \frac{\partial v_2}{\partial k} = 0$. And we can show that $\partial k^* / \partial \theta_i > 0$ by implicit function theorem so that $\left(\frac{\partial^2 v_1}{\partial k^2} + \frac{\partial^2 v_2}{\partial k^2} \right) \frac{\partial k^*}{\partial \theta_1} \frac{\partial k^*}{\theta_2} < 0$, contradiction. □

4.1 Rochet and Vorha Theorems

Definition 44 (Implementability)**.** *Suppose we have one agent k is implementable if $\exists t : \Theta \to R$*

$$v(\theta, k(\theta)) - t(\theta) \geq v(\theta, k(\theta')) - t(\theta') \quad \forall \theta, \theta'$$

Definition 45 (C-MON)**.** *k is C-MON if $(\theta_1, \dots \theta_n = \theta_1)$*

$$\sum_{i=1}^{m} v(\theta_{i+1}, k(\theta_i)) - v(\theta_i, k(\theta_i)) \leq 0$$

k is MON if it holds 2-cycles ($m = 3$)

Theorem 15 (Rochet 1989)**.** ***k is implementable $\iff$ k is C-MON***

Proof. $\Rightarrow$ Let $(\theta_1, \dots \theta_n, \theta_{n+1} = \theta_1)$ be a cycle. Then

$$v(\theta_1, k(\theta_1)) - t(\theta_1) \geq v(\theta_1, k(\theta_m)) - t(\theta_m)$$

$$v(\theta_2, k(\theta_2)) - t(\theta_2) \geq v(\theta_2, k(\theta_1)) - t(\theta_1)$$

$$\dots$$

$$v(\theta_m, k(\theta_m)) - t(\theta_m) \geq v(\theta_m, k(\theta_{m-1})) - t(\theta_{m-1})$$

sum it up to obtain:

$$\sum_{i=1}^{m} v(\theta_{i+1}, k(\theta_i) - v(\theta_i, k(\theta_i)) \leq 0$$

$\Rightarrow$

$$V(t,s) := \sup_{\gamma}\{\sum_{i=1}^{m} v(\theta_{i+1}, k(\theta_i) - v(\theta_i, k(\theta_i)) : \gamma = (\theta_1 = t, \dots \theta_m, \theta_{m+1} = s)\}$$

C-MON $\Rightarrow V(\hat{\theta}, \hat{\theta}) \leq 0$

$$V(\hat{\theta}, \hat{\theta}) \geq V(\hat{\theta}, \theta) + v(\hat{\theta}, k(\theta)) - v(\theta, k(\theta) \quad \forall \theta$$

therefore $V(\hat{\theta}, \theta) < \infty$. Again:

$$V(\hat{\theta}, \theta) \geq V(\hat{\theta}, s) + v(\theta, k(s)) - v(s, k(s)) \quad \forall \theta$$

let $t(\theta) = v(\theta, k(\theta)) - V(\hat{\theta}, \theta)$

$$v(\theta, k(\theta)) - t(\theta) \geq v(s, k(s)) - t(s) + v(\theta, k(s)) - v(s, k(s)) = v(\theta, k(s)) - t(s)$$

□

Revenue Equivalence Suppose k is implementable and $\exists t$. k exhibits revenue equivalence if $t, t^{'}$ that implement k modulo a constant $\exists c \in R$:

$$t(\theta) = t^{'}(\theta) + c$$

Definition 46 (Budget Balance). *A SCF is Fully EPE (FEPE) if it is EPE and* $f = (k, t)$

$$\sum_{i=1}^{m} t_i(\theta) = 0 \quad \forall \theta \quad \theta = (\theta_1, \ldots \theta_n)$$

Theorem 16. ***Let*** $V = \{v(\theta) \in \mathbb{R}^X : \theta \in \Theta\} = R^k$ $n > 1$ ***then*** $\nexists$ ***FEPE SCF.***

If asome agents' preferences are known then they can 'break the budget'.

Theorem 17. ***Suppose*** k ***is implementable.*** k ***exhibits Revenue Equivalence*** $\iff$

$$V(\theta, \theta^{'}) = -V(\theta^{'}, \theta)$$

Proof. TBD □

Theorem 18. ***Let*** $V = \{v(\theta) \in \mathbb{R}^X : \theta \in \Theta\}$***. If*** k ***is implementable and*** V ***is a convex set then*** k ***exhibits revenue equivalence***

Proof. TBD □

Corollary 2. *Grove's scheme uniquely implements an ex post efficient allocation if V is convex*

Expected Externality Mechanism

Suppose types are independent. Let for any h_i

$$t_i(\theta) = \mathbb{E}_{\theta_{-i}} \left[\sum_{j \neq i} v_j(z^*(\theta_i, \hat{\theta}_{-i}), \hat{\theta}_j) \right] + h_i(\theta_{-i})$$

Theorem 19. (z^*, t) ***is Bayesian IC***

Proof.

$$\mathbb{E}\left[v_i(z^*(\theta_i, \theta_{-i}), \theta_i) + t_i(\theta) | \theta_i\right] =$$

$$= \mathbb{E}\left[\sum_{j=1}^{n} v_j(z^*(\theta_i, \theta_{-i}), \theta_j) | \theta_i \right] + \mathbb{E}\left[h_i(\theta_{-i} | \theta_i\right] \geq$$

$$\geq \mathbb{E}\left[\sum_{j=1}^{n} v_j(z^*(\theta_i', \theta_{-i}), \theta_j)|\theta_i\right] + \mathbb{E}\left[h_i(\theta_{-i}|\theta_i\right]$$

-

$$= \mathbb{E}\left[v_i(z^*(\theta_i', \theta_{-i}), \theta_i) + t_i(\theta_i', \theta_{-i})|\theta_i\right]$$

□

$_i(\theta_i', \theta_{-i})$ does not depend on t_i. Now choose h_i to obtain Balanced Budget:

Let

$$t_i(\theta_i) = \mathbb{E}_{\theta_{-i}}[\sum_{j \neq i} v_j(z^*(\theta_i, \theta_{-i}), \theta_j)]$$

and

$$h_i(\theta_{-i}) = -\frac{1}{n-1}\sum_{j \neq i} t_j(\theta_j)$$

Let's check

$$\sum_{i=1}^{n} t_i(\theta) = \sum_{i=1}^{n} t_i(\theta_i) + \sum_{i=1}^{n} h_i(\theta_{-i}) = \sum_{i=1}^{n} t_i(\theta_i) - -\frac{1}{n-1}\sum_{i=1}^{n}\sum_{j \neq i} t_j(\theta_j) = 0$$

Every individual receive $t_i(\theta_i)$ and contributes an equal share $\frac{1}{n-1}$ of everyone else's payments so net transfer is

$$t_i(\theta_i) - \frac{1}{n-1}\sum_{j \neq i} t_j(\theta_j)$$

Transfers ⇒ a particular distribution of utility across types

4.2 Individual Rationality

- In general, without IR, there exists a Groves scheme that would be DSIC
- With IR, there would not exist such a Groves scheme

Example 7. *$K = \{0, 1\}, n = 2, \Theta_i = \{\underline{\theta}, \bar{\theta}\}, \bar{\theta} > 2\underline{\theta} > 0$. If $k = 1$, there is a cost $c \in (2\underline{\theta}, \bar{\theta})$. $v_i(0, \theta_i) = 0$ and $v_i(1, \theta_i) = \theta_i$*

- *The efficient allocation is*

$$k^* = \begin{cases} 1 & \text{if } \theta_1 = \bar{\theta} \text{ or } \theta_2 = \bar{\theta} \\ 0 & \text{if } \theta_1 = \theta_2 = \underline{\theta} \end{cases}$$

- *Ex-post IR:* $\underline{\theta} + t_1(\underline{\theta}, \bar{\theta}) \geq 0 \Rightarrow t_1(\underline{\theta}, \bar{\theta}) \geq -\underline{\theta}$.

- *DSIC:* $\bar{\theta} + t_1(\bar{\theta}, \bar{\theta}) \geq \bar{\theta} + t_1(\underline{\theta}, \bar{\theta}) \Rightarrow t_1(\bar{\theta}, \bar{\theta}) \geq t_1(\underline{\theta}, \bar{\theta}) \geq -\underline{\theta}$

- *By symmetry,* $t_2(\bar{\theta}, \bar{\theta}) \geq -\underline{\theta}$.

- *Therefore,* $t_1(\bar{\theta}, \bar{\theta}) + t_2(\bar{\theta}, \bar{\theta}) \geq -2\underline{\theta} > -c$;

- *But balanced budget requires* $t_1(\bar{\theta}, \bar{\theta}) + t_2(\bar{\theta}, \bar{\theta}) + c \leq 0$, *contradiction.*

Example 8. *Same as previous one*

- *But consider a prior* $\Pr(\theta_i) = \frac{1}{2}$.

- *Interim IR:* $\frac{1}{2}\left[\underline{\theta} + t_1(\underline{\theta}, \bar{\theta}) + t_1(\underline{\theta}, \underline{\theta})\right] \geq 0 \Rightarrow t_1(\underline{\theta}, \bar{\theta}) \geq -\underline{\theta}$

- *Interim IC:* $\frac{1}{2}\left[\bar{\theta} + t_1(\bar{\theta}, \underline{\theta}) + \bar{\theta} + t_1(\bar{\theta}, \bar{\theta})\right] \geq \frac{1}{2}\left[\bar{\theta} + t_1(\underline{\theta}, \bar{\theta}) + t_1(\underline{\theta}, \underline{\theta})\right] \Rightarrow t_1(\bar{\theta}, \bar{\theta}) + t_1(\bar{\theta}, \underline{\theta}) \geq -\bar{\theta}$;

- *By symmetircity:* $t_2(\bar{\theta}, \underline{\theta}) \geq -\underline{\theta}$ *and* $t_2(\bar{\theta}, \bar{\theta}) + t_2(\underline{\theta}, \bar{\theta}) \geq -\bar{\theta}$; *Therefore,* $\sum_i t_i(\theta) \geq -2\bar{\theta} - 2\underline{\theta}$

- *By the feasibility region:*

$$
\begin{aligned}
t_1(\bar{\theta}, \bar{\theta}) + t_2(\bar{\theta}, \bar{\theta}) &\leq -c \\
t_1(\bar{\theta}, \underline{\theta}) + t_2(\bar{\theta}, \underline{\theta}) &\leq -c \\
t_1(\underline{\theta}, \bar{\theta}) + t_2(\underline{\theta}, \bar{\theta}) &\leq -c \\
t_1(\underline{\theta}, \underline{\theta}) + t_2(\underline{\theta}, \underline{\theta}) &\leq 0
\end{aligned}
$$

- *Therefore,* $-3c \geq \sum_i t_i(\theta) \geq -2\bar{\theta} - 2\underline{\theta} \Rightarrow c \leq \frac{2}{3}(\bar{\theta} + \underline{\theta})$. *So, if* $\bar{\theta} < 3\underline{\theta}$, *then there is a contradiction.*

4.3 Bayesian Nash equilibrium implementation

Definition 47 (Bayesian Nash equilibrium). *of G is a strategy profile $\sigma^* = (\sigma_i^*)_{i\in I}$ (where $\sigma_i^* : \Theta_i \to M_i$) such that $\forall i \in I, \forall \theta_i \in \Theta_i, \forall s_i \in \Sigma_i$*

$$\mathbb{E}_{\theta_{-i}}\left[u_i\left(g\left(\sigma_i^*\left(\theta_i\right), \sigma_{-i}^*\left(\theta_{-i}\right)\right), \theta_i\right) \mid \theta_i\right] \geq \mathbb{E}_{\theta_{-i}}\left[u_i\left(g\left(m_i, \sigma_{-i}^*\left(\theta_{-i}\right)\right), \theta_i\right) \mid \theta_i\right]$$

We will use $BNE(G)$ to denote the set of Bayesian Nash equilibria of G.

Definition 48 (Implementation). *Γ implements social choice function $f : \Theta \to X$ in Bayesian Nash equilibrium if* $\text{BNE}(G) \neq \emptyset$ *and $\exists \sigma^* \in BNE(G)$ such that*

$$g\left(\sigma^*(\theta)\right) = f(\theta) \quad \forall \theta \in \text{support}(\phi)$$

Definition 49 (Bayes-Nash Incentive Compatible (BNIC)). *The SCF f truthfully implementable in Bayesian Nash equilibrium (or BNIC) if $s_i^*(\theta_i) = \theta_i$ and s^* is BN equilibrium of the direct revelation mechanism $\Gamma = (\Theta, f)$. That is:*

$$\mathbb{E}_{\theta_{-i}}\left[u_i\left(f\left(\theta_i, \theta_{-i}\right), \theta_i\right) \mid \theta_i\right] \geq \mathbb{E}_{\theta_{-i}}\left[u_i\left(f\left(\theta_i', \theta_{-i}\right), \theta_i\right) \mid \theta_i\right]$$

As in the d.s. equilibrium section, Γ strictly/fully implements f if the above the above condition holds for all $\sigma^* \in BNE(G)$ rather than just one of them.

Theorem 20 (Revelation principle for BN implementation). ***. If f is BN -implementable, then f is BN implementable by the direct mechanism $\Gamma_{direct} = (\Theta, f)$ with truth-telling as a BNE.***

Proof. Suppose Γ BN-implements f. Let σ^* be the BNE such that $g\left(\sigma^*(\theta)\right) = f(\theta), \forall \theta \in \text{support}(\phi)$. By definition of a BNE, $\forall i \in I, \forall \theta \in \Theta, \forall m_i \in M_i$

$$E_{\theta_{-i}}\left[u_i\left(g\left(\sigma_i^*\left(\theta_i\right), \sigma_{-i}^*\left(\theta_{-i}\right)\right), \theta_i\right) \mid \theta_i\right] \geq E_{\theta_{-i}}\left[u_i\left(g\left(m_i, \sigma_{-i}^*\left(\theta_{-i}\right)\right), \theta_i\right) \mid \theta_i\right]$$

Since $g\left(\sigma^*(\theta)\right) = f(\theta), \forall \theta \in \text{support}(\phi)$, we have $\forall i \in I, \forall \theta \in \Theta, \forall \theta_i' \in \Theta_i$

$$E_{\theta_{-i}}\left[u_i\left(f\left(\theta_i, \theta_{-i}\right), \theta_i\right) \mid \theta_i\right] \geq E_{\theta_{-i}}\left[u_i\left(f\left(\theta_i', \theta_{-i}\right), \theta_i\right) \mid \theta_i\right]$$

Thus truth-telling is a BNE in the direct game, i.e., f is strategy proof. □

Definition 50. *If f is strategy proof, we say that f is d.s. incentive compatible. If f is BN-implementable, we say that f is Baysesian incentive compatible. Note that d.s. IC implies BN IC, but not the other way around. In terms of set notation, BN IC $\supset$ d.s. IC.*

Because BN implementation is a weaker condition, it seems reasonable that a wider variety of social choice functions could be BN-implemented. Exactly which SCFs can be BN-implemented depends on the assumptions about the probability distribution ϕ. If ϕ is a product measure (i.e., θ_i are independently distributed), then the kinds of SCFs that can be BN-implemented are pretty similar to the ones that can be d.s. implemented (in this case there is a similar monotoncity condition that we will see shortly). However, if ϕ is not a product measure, then under weak regularity condition virtually all SCFs can be implemented. In this case, the transfer function that implements the SCF will often take a complex structure and be very sensitive to parameters.

Definition 51. *Let ϕ be a product measure and suppose that $t(\cdot)$ implements $k(\cdot) = (y_1(\cdot), \ldots, y_n(\cdot))$ For each $i \in I$, define $\bar{y}_i(\cdot)$ and $\bar{y}_i(\cdot)$ as*

$$\bar{t}_i(\theta_i) := E_{\theta_{-i}}\left[t_i(\theta_i, \theta_{-i})\right]$$

and

$$\bar{y}_i(\theta_i) := E_{\theta_{-i}}\left[y_i(\theta_i, \theta_{-i})\right]$$

These definitions allow us to write the BN incentive compatibility condition as

$$\bar{y}_i(\theta_i)\,\theta_i + \bar{t}_i(\theta_i) \geq \bar{y}_i\left(\theta_i'\right)\theta_i + \bar{t}_i\left(\theta_i'\right), \forall \theta_i' \in \Theta_i$$

Note that we don't have a separate IC condition for each θ_{-i} as we did in the DS implementation case. Instead, we just have one constraint given the averages $\bar{y}$ and $\bar{t}$ implied by ϕ.

Theorem 21. $k(\cdot)$ ***is*** *BN* ***-implementable only if it is "monotone" in the sense that*** $\forall i \in I, \bar{y}_i(\cdot)$ ***is weakly increasing.***

Proof. Suppose not. Then $\exists i \in I$ and $\theta_i, \theta_i' \in \Theta_i$ such that $\theta_i < \theta_i'$ but $\bar{y}_i(\theta) > \bar{y}_i\left(\theta_i'\right)$. Let $t(\cdot)$ be a transfer function that BN-implements $k(\cdot)$. If player i 's type is θ, then incentive compatibility implies that

$$\bar{y}_i(\theta_i)\,\theta_i + \bar{t}_i(\theta_i) \geq \bar{y}_i\left(\theta_i'\right)\theta_i + \bar{t}_i\left(\theta_i'\right)$$

If his type is θ_i', then IC implies that

$$\bar{y}_i\left(\theta_i'\right)\theta_i' + \bar{t}_i\left(\theta_i'\right) \geq \bar{y}_i(\theta_i)\,\theta_i' + \bar{t}_i(\theta_i)$$

Adding both inequalities together and cancelling out the transfers implies yields

$$\bar{y}_i\left(\theta_i'\right)\left(\theta_i'-\theta_i\right) \geq \bar{y}_i\left(\theta_i\right)\left(\theta_i'-\theta_i\right)$$

Then $\theta_i' - \theta_i$ cancels out as well, implying that

$$\bar{y}_i\left(\theta_i'\right) \geq \bar{y}_i\left(\theta_i\right)$$

This is a contradiction, so it must be that $k(\cdot)$ is "monotone." □

This leads us to the following question: Are all monotonic $k(\cdot)$ BN-implementable? Let's turn again to **Example 2**, the fact that $k(\cdot)$ was finite implied that y_i was also finite. Here, we are dealing with $\bar{y}_i$, which is a probability. Thus the benefit of dealing with a finite K is gone, i.e., we will now consider the most general case:

$$K=\left\{k=(y_1,\ldots,y_n)\in[0,1]^n : \sum_{i\in I} y_i \leq 1\right\}$$

Given type θ_i and any $m \in M_i$, player i gets expected utility of

$$p(m)\theta_i + t_i(m)$$

which is a linear function of θ_i. Thus player i 's optimal strategy is

$$s^* = \operatorname{argmax}\left\{p(m)\theta_i + t(s) : s \in \Sigma_i\right\}$$

Since $\Sigma_i = \Theta_i$ in the direct mechanism, this is equivalent to

$$\theta_i^* = \operatorname{argmax}\left\{p\left(\theta_i'\right)\theta_i + t\left(\theta_i'\right) : \theta_i' \in \Theta_i\right\}$$

In other words, player i will claim to have whichever type puts him on the upper envelope of the set of lines given above. Call this **upper envelope** $U_i\left(\theta_i\right)$.

Since $U_i\left(\theta_i\right)$ is the upper envelope of a set of linear functions, $U_i\left(\theta_i\right)$ is convex and continuous. This implies that is differentiable almost everywhere, and $U_i\left(\theta_i\right)$ is equal to the integral of its own derivatives, i.e.

$$U_i\left(\theta_i\right) = U_i(x) + \int_x^{\theta_i} U_i'(y)\mathrm{d}y$$

Note that $U_i'(y) = p(m)$ for some $m \in M_i$ If $k(\cdot)$ is incentive compatible, then $U_i'\left(\theta_i\right) = p\left(\theta_i\right) = \bar{y}_i\left(\theta_i\right)$, the probability of getting the object by playing type θ_i. Thus

$$U_i\left(\theta_i\right) = U_i(x) + \int_x^{\theta_i} \bar{y}_i(z)\mathrm{d}z$$

Assuming $\Theta_i = [\underline{\theta}_i, \bar{\theta}_i]$, we can also write this as

$$U_i(\theta_i) = \bar{y}_i(\theta_i)\,\theta_i + \bar{t}_i(\theta_i) = U_i(\underline{\theta}_i) + \int_{\underline{\theta}_i}^{\theta_i} \bar{y}_i(z)\mathrm{d}z$$

Rearranging, we have

$$\bar{t}_i(\theta_i) = \bar{t}_i(\underline{\theta}_i) - [\bar{y}_i(\theta_i)\,\theta_i - \bar{y}_i(\underline{\theta}_i)\,\underline{\theta}_i] + \int_{\underline{\theta}_i}^{\theta_i} \bar{y}_i(x)\mathrm{d}x.$$

Note that if $\underline{\theta}_i = 0$, this simplifies to

$$\bar{t}_i(\theta_i) = \bar{t}_i(\underline{\theta}_i) - \bar{y}_i(\theta_i)\,\theta_i + \int_0^{\theta_i} \bar{y}_i(x)\mathrm{d}x.$$

Thus if $t(\cdot)$ implements $k(\cdot), \bar{t}_i(\cdot)$ must take this form. Let's check to see that this $\bar{t}_i(\cdot)$ works. If $\theta_i < \theta_i'$, then we need to make sure that player i will not lie regardless of whether his true type is θ_i or θ_i'. Note that

$$U_i(\theta_i') - U\Big(i(\theta_i) = \int_{\theta_i}^{\theta_i'} \bar{y}_i(x)\mathrm{d}x.$$

This implies that

$$\bar{y}_i(\theta_i)\left(\theta_i' - \theta_i\right) \le U_i(\theta_i') - U\left(i(\theta_i) \le \bar{y}_i(\theta_i')\left(\theta_i' - \theta_i\right)\right.$$

Using the first inequality, we get

$$U_i(\theta_i') \ge U_i(\theta_i) + \bar{y}_i(\theta_i)\left(\theta_i' - \theta_i\right)$$

Using the defintion of $\bar{t}_i(\cdot)$ from above, we get

$$U_i(\theta_i') \ge \bar{t}_i(\theta_i) + \bar{y}_i(\theta_i)\,\theta_i'$$

Thus player i is better off by telling the truth when his true type is θ_i'. Using the second inequality, we get

$$U_i(\theta_i) \ge U_i(\theta_i') + \bar{y}_i(\theta_i')\left(\theta_i' - \theta_i\right)$$

Using the defintion of $\bar{t}_i(\cdot)$ from above, we get

$$U_i(\theta_i) \ge \bar{t}_i(\theta_i') + \bar{y}_i(\theta_i')\,\theta_i$$

Thus player i is also better off by telling the truth when his true type is θ_i. Therefore $\bar{t}_i(\cdot)$ defined above does indeed BN-implement $k(\cdot)$

To summarize, $\bar{y}_i(\cdot)$ is BN implementable $\iff$ it is "monotone" (in the BN sense). Moreover, if Θ_i is a connected subset of $\mathbb{R}$, then any $\bar{t}_i(\cdot)$ that BN-implements $\bar{y}_i(\cdot)$ must take the form

$$\bar{t}_i(\theta_i) = \bar{t}_i(\underline{\theta}_i) - \bar{y}_i(\theta_i)\,\theta_i + \int_0^{\theta_i} \bar{y}_i(x)\mathrm{d}x$$

4.4 Expected Externality Mechanism

Each agent has utility $u_i(k, \theta_i) = v_i(k, \theta_i) + t_i$

Definition 52. *SCF f is Bayesian Nash Incentive Compatible if, for all i, θ_i, θ_i', and θ_{-i}*

$$E\left[v_i(k(\theta), \theta_i) + t_i(\theta) \mid \theta_i\right] \geq E\left[v_i\left(k\left(\theta_i', \theta_{-i}\right), \theta_i\right) + t_i\left(\theta_i', \theta_{-i}\right) \mid \theta_i\right]$$

Theorem 22. ***Suppose agents' types are statistically independent. Let k^* satisfy ex-post efficiency and let $t_i(\theta) = E\left[\sum_{j \neq i} v_j\left(k^*(\theta), \theta_j\right) \mid \theta_i\right] + h_i(\theta_{-i})$. Then, $f = (k^*, t)$ is Bayesian incentive compatible and there exist some h_i for each i such that it satisfies balanced budget.***

Proof.

- BIC. For any θ_i, θ_i',

$$\begin{aligned}
&E\left[v_i(k^*(\theta), \theta_i) + t_i(\theta) \mid \theta_i\right] \\
=&E\left[v_i(k^*(\theta), \theta_i) \mid \theta_i\right] + E\left[\sum_{j \neq i} v_j\left(k^*(\theta), \theta_j\right) \mid \theta_i\right] + E\left[h_i(\theta_{-i})\right] \\
=&E\left[\sum_{i=1}^{n} v_i(k^*(\theta), \theta_i) \mid \theta_i\right] + E\left[h_i(\theta_{-i})\right] \\
\geq&E\left[\sum_{i=1}^{n} v_i\left(k^*\left(\theta_i', \theta_{-i}\right), \theta_i\right) \mid \theta_i\right] + E\left[h_i(\theta_{-i})\right] \\
=&E\left[v_i\left(k^*\left(\theta_i', \theta_{-i}\right), \theta_i\right) + t_i(\theta) \mid \theta_i\right]
\end{aligned}$$

- Balanced budget. Let

$$h_i(\theta_{-i}) = -\frac{1}{n-1} \sum_{j \neq i} \xi_j(\theta_j)$$

  where

$$\xi_i(\theta_i) = E\left[\sum_{j \neq i} v_j\left(k^*(\theta), \theta_j\right) \mid \theta_i\right].$$

  Then

$$\begin{aligned}
\sum_{i=1}^{n} t_i(\theta_i) &= \sum_{i=1}^{n} \xi_i(\theta_i) + \sum_{i=1}^{n} h_i(\theta_{-i}) \\
&= \sum_{i=1}^{n} \xi_i(\theta_i) - \frac{1}{n-1} \sum_{i=1}^{n} \sum_{j \neq i} \xi_j(\theta_j) = 0
\end{aligned}$$

□

Zero-sum transfers in BN-implementation (expected externality mechanism) In the DS implementation section, we found that it was impossible to allocate an object between two bidders using zero-sum transfers in a strategy-proof way. In the BN-implementation context, this is no longer the case. Start with a normal second-price auction and modify it so the winner pays a fixed amount to the other player equal to the average he would have paid in the normal second-price auction. In other words, define

$$\tau_i(\theta_i) = E_{\theta_{-i}}\left[-t_i(\theta_i, \theta_{-i})\right], \forall i = 1, 2$$

Then define

$$\hat{t}_i(\theta_i, \theta_{-i}) = -\tau_i(\theta_i) + \tau_{-i}(\theta_{-i}), \forall i = 1, 2$$

This shuts down the incentive to lie and implements the efficient $k^*(\cdot)$ in BN-equilibrium with a balanced budget (tranfers stay "within the system"). Note that in a more general context with more than two players, we can define

$$\hat{t}_i(\theta_i, \theta_{-i}) = \tau_i(\theta_i) + \sum_{j \neq i} \frac{1}{n-1} \tau_j(\theta_j)$$

This will implement the efficient allocation with zero-sum transfers.

5 Optimal Mechanisms

This section is based on seminal paper by Myerson (1981)

Set-up

- One seller has an object and n bidders, set of agents is $A = \{0, 1, \ldots, n\}$.
- The utility function for i is

$$v_i(a, \theta_i) = \begin{cases} \theta_i & \text{if } a = i \\ 0 & \text{otherwise} \end{cases}$$

- For each $i = 1, \ldots, n, \theta_i$ is independently drawn from a continuous PDF $\phi_i = (\theta_i)$ with support $\Theta_i \ [\underline{\theta}_i, \bar{\theta}_i]$
- The mechanism is (q, t) where: $-q_i(\theta) = \Pr(a = i \mid \theta)$ is the probability of agent i get the object; $-t_i(\theta) : \Theta \to \mathbb{R}$ is the transfer from buyers to the seller where $\Theta = \prod_{i=1}^{n} \Theta_i$
- Denote

$$\bar{q}_i(\theta_i) = E[q_i(\theta_i, \theta_{-i}) \mid \theta_{-i}] = \int\limits_{\Theta_{-i}} q_i(\theta)\phi_{-i}(\theta_{-i})d\theta_{-i}$$

- Denote

$$\bar{t}_i(\theta_i) = E[t_i(\theta_i, \theta_{-i}) \mid \theta_i]$$

- By revelation principal, restrict our focus on direct mechanism.
- $q_i(\theta)$ needs to be probabilities $(PR) : \sum_{i=1}^{n} q_i(\theta) \leq 1$ and $q_i(\theta) \geq 0 \forall i.$
- For $i = 1, \ldots, n$, given the mechanism (q, t), the expected utility is

$$\begin{aligned} U_i(\theta_i) &= \int_{\Theta_{-i}} (\theta_i q(\theta) - t_i(\theta)) \phi_{-i}(\theta_{-i}) d\theta_{-i} \\ &= \theta_i \bar{q}_i(\theta_i) - \bar{t}_i(\theta_i) \end{aligned}$$

- The seller's expected utility is

$$U_0(\theta_0) = \int_{\Theta} \left[\theta_0 \left(1 - \sum_{i=1}^{n} q_i(\theta) \right) + \sum_{i=1}^{n} t_i(\theta) \right] \phi(\theta) d\theta$$

- Individual rationality (IR) requires:

$$U_i(\theta_i) \geq 0 \forall i = 1, \ldots n, \forall \theta_i \in \Theta_i$$

- Incentive-compatibility (IC) requires: for all $\theta_i' \in \Theta_i$

$$U_i(\theta_i) \geq \int_{\Theta_{-i}} \left(\theta_i q\left(\theta_i', \theta_{-i}\right) - t_i\left(\theta_i', \theta_{-i}\right)\right) \phi_{-i}(\theta_{-i})\, d\theta_{-i}$$

Definition 53. *A mechanism (p, t) is feasible if it satisfies PR, IR, and IC.*

Lemma 5. *A mechanism (p, t) is feasible $\iff$*

1. *$\bar{q}_i$ is nondecreasing*
2. $U_i(\theta_i) = U_i(\underline{\theta}_i) + \int_{\underline{\theta}_i}^{\theta_i} \bar{q}_i(\tau_i)\, d\tau_i$
3. $U_i(\underline{\theta}_i) \geq 0 \forall i = 1, \ldots, n$
4. $\sum_{i=1}^{n} q_i(\theta) \leq 1$ *and* $q_i(\theta) \geq 0 \forall i$.

Proof. $\Rightarrow$ Take $\theta_i' > \theta_i$

$$U_i(\theta_i) \geq \theta_i \bar{q}_i\left(\theta_i'\right) - \bar{t}_i\left(\theta_i'\right) =$$

$$= U_i\left(\theta_i'\right) + \left(\theta_i - \theta_i'\right) \cdot \bar{q}_i\left(\theta_i'\right)$$

and

$$U_i\left(\theta_i'\right) \geq \theta_i' \bar{q}_i(\theta_i) - \bar{t}_i(\theta_i) =$$

$$= U_i(\theta_i) + \left(\theta_i' - \theta_i\right) \cdot \bar{q}_i(\theta_i)$$

so

$$\bar{q}_i\left(\theta_i'\right) \geq \frac{U_i\left(\theta_i'\right) - U_i(\theta_i)}{\theta_i' - \theta_i} \geq \bar{q}_i(\theta_i)$$

so $\bar{q}_i$ is non decreasing.

As $\theta_i' \to \theta_i$

$$U_i'(\theta_i) = \bar{q}_i(\theta_i)$$

so integrating it ensures

$$U_i(\theta_i) = U_i(\underline{\theta}_i) + \int_{\underline{\theta}_i}^{\theta_i} \bar{q}_i(\tau_i)\, d\tau_i$$

$\Leftarrow$ Take $\theta_i' < \theta_i$ (1) and (2) gives

$$U_i(\theta_i) - U_i(\theta_i') = \int_{\underline{\theta}_i}^{\theta_i} \bar{q}_i(\tau_i)\, d\tau_i \geq$$

$$\geq \int_{\underline{\theta}_i}^{\theta_i} \bar{q}_i(\tau_i')\, d\tau_i = (\theta_i - \theta_i') \cdot \bar{t}_i(\theta_i')$$

So we obtained IC. Similar argument holds for $\theta_i' > \theta_i$ □

Definition 54. *A mechanism (q, t) is an optimal auction if it maximizes $U_0(\theta_0)$ subject to feasibility.*

Lemma 6. *If $q : \Theta \to \mathbb{R}^n$ maximizes*

$$\int_\Theta \left[\sum_{i=1}^{n} \left(\theta_i - \frac{1 - \Phi(\theta_i)}{\phi(\theta_i)} - \theta_0 \right) q_i(\theta) \right] \phi(\theta) d\theta$$

subject to (1) and (4) from previous lemma, and

$$t_i(\theta) = \theta_i q_i(\theta) - \int_{\underline{\theta}_i}^{\theta_i} q_i(\theta_i, \theta_{-i})\, d\theta_i$$

Then, (q, t) is an optimal auction (maximizing objective of a seller).

Proof.

$$U_0(\theta_0) = \int_\Theta \left[\theta_0 \left(1 - \sum_{i=1}^{n} q_i(\theta) \right) + \sum_{i=1}^{n} t_i(\theta) \right] \phi(\theta) d\theta$$

$$= \theta_0 + \sum_{i=1}^{n} \int_\Theta q_i(\theta)(\theta_i - \theta_0)\phi(\theta) d\theta + \sum_{i=1}^{n} \int_\Theta [t_i(\theta) - \theta_i q_i(\theta)]\, \phi(\theta) d\theta$$

By feasibility and Lemma ,

$$\int_\Theta [t_i(\theta) - \theta_i q_i(\theta)]\, \phi(\theta) d\theta$$

$$= -\int_{\underline{\theta}_i}^{\bar{\theta}_i} U_i(\theta_i)\, \phi_i(\theta_i)\, d\theta_i$$

$$= -\int_{\underline{\theta}_i}^{\bar{\theta}_i} \left[U_i(\underline{\theta}_i) + \int_{\underline{\theta}_i}^{\theta_i} \bar{q}_i(\tau_i)\, d\tau_i \right] \phi_i(\theta_i)\, d\theta_i$$

$$= -U_i(\underline{\theta}_i) - \int_{\underline{\theta}_i}^{\bar{\theta}_i} \left[\int_{\underline{\theta}_i}^{\theta_i} \bar{q}_i(\tau_i)\, d\tau_i \right] \phi_i(\theta_i)\, d\theta_i$$

$$= -U_i(\underline{\theta}_i) - \int_{\underline{\theta}_i}^{\bar{\theta}_i} \int_{\tau_i}^{\bar{\theta}_i} \bar{q}_i(\tau_i)\, \phi_i(\theta_i)\, d\theta_i d\tau_i$$

$$= -U_i(\underline{\theta}_i) - \int_{\underline{\theta}_i}^{\bar{\theta}_i} \bar{q}_i(\tau_i)\, [1 - \Phi_i(\tau_i)]\, d\tau_i$$

$$= -U_i(\underline{\theta}_i) - \int_\Theta q_i(\theta)\, [1 - \Phi_i(\theta_i)]\, \phi_{-i}(\theta_{-i})\, d\theta.$$

Then,

$$U_0(\theta_0) = \theta_0 + \sum_{i=1}^{n} \int_\Theta q_i(\theta)(\theta_i - \theta_0)\phi(\theta)d\theta - \sum_{i=1}^{n} \int_\Theta q_i(\theta)[1-\Phi_i(\theta_i)]\phi_{-i}(\theta_{-i})d\theta - \sum_{i=1}^{n} U_i(\underline{\theta}_i)$$

And,

$$t_i(\theta) = \theta_i q_i(\theta) - \int_{\underline{\theta}_i}^{\theta_i} q_i(\theta_i, \theta_{-i})d\theta_i \Rightarrow \bar{t}_i(\underline{\theta}_i) = \underline{\theta}_i \bar{q}_i(\underline{\theta}_i)$$

Therefore,

$$U_i(\underline{\theta}_i) = \underline{\theta}_i \bar{q}_i(\underline{\theta}_i) - \bar{t}_i(\underline{\theta}_i) = 0 \forall i = 1, \ldots n$$

That is, t_i is chosen to maximize $-\sum_{i=1}^{n} U_i(\underline{\theta}_i) \leq 0$. Rearrange the objective can be simplified to

$$U_0(\theta_0) = \int_\Theta \left[\sum_{i=1}^{n}\left(\theta_i - \frac{1-\Phi_i(\theta_i)}{\phi_i(\theta_i)} - \theta_0\right) q_i(\theta)\right]\phi(\theta)d\theta + \theta_0.$$

where we can drop t_i from the problem. Since θ_0 is constant, if q_i is chosen to maximize

$$\int_\Theta \left[\sum_{i=1}^{n}\left(\theta_i - \frac{1-\Phi_i(\theta_i)}{\phi_i(\theta_i)} - \theta_0\right) q_i(\theta)\right]\phi(\theta)d\theta$$

subject to PR and nondecreasing, which are the only constraints regarding q_i. Then the solution is feasible and maximizes the objective. □

$$U_i(\theta_i, \theta_i') := \theta_i q_i(\theta_i', \theta_{-i}) - t_i(\theta_i', \theta_{-i})$$

Define

$$U_i(\theta_i) = \max_{\theta_i'} U_i(\theta_i, \theta_i')$$

Theorem 23 (Envelope).

$$\frac{dU_i(\theta_i)}{d\theta_i'} = q_i(\theta)$$

Corollary 3 (The Revenue-Equivalence Theorem). . *The seller's expected utility from a feasible auction mechanism is completely determined by the probability functions q_i and the numbers $U_i(\underline{\theta}_i)$ of each $i = 1, \ldots, n$*

Regular Case

Definition 55. *The problem is regular if for each i*

$$w_i(\theta_i) = \theta_i - \frac{1 - \Phi_i(\theta_i)}{\phi_i(\theta_i)}$$

is strictly monotone.

Relaxed problem Let drop (1) for a second $\Rightarrow$ objective goes to bidder with the highest w_i

- Consider the following auction mechanism:
- Seller keeps the object if $\theta_0 > \max_i (w_i(\theta_i))$
- Otherwise, give it to $i^* = \arg\min \{i \mid w_i(\theta_i) = \max_j (w_j(\theta_j))\}$ Set $t_i(\theta) = \theta_i q_i(\theta) - \int_{\underline{\theta}_i}^{\theta_i} q_i(\theta_i, \theta_{-i})\, d\theta_i$

Theorem 24. ***The auction mechanism*** (q, t) ***is optimal.***

Proof. By the construction,

$$q_i(\theta_i) > 0 \Rightarrow w_i(\theta_i) = \max_j (w_j(\theta_j))$$

Therefore, q maximizes $\sum_{i=1}^n (w_i(\theta_i) - \theta_0)\, q_i(\theta)$, subject to PR (4), and hence the objective. Moreover, to see that $\bar{q}_i$ is non-decreasing, we shall see that $q_i(\theta_i, \theta_{-i})$ is non-decreasing in θ_i for all θ_{-i}. $\theta_i \leq \theta_i' \Rightarrow w_i(\theta_i) \leq w_i(\theta_i')$ since w_i is nondecreasing, and suppose, for contradiction, $q_i(\theta_i, \theta_{-i}) > q_i(\theta_i', \theta_{-i})$. Then, it must the case that $q_i(\theta_i', \theta_{-i}) = 0$ and $q_i(\theta_i, \theta_{-i}) = 1$ $q_i(\theta_i, \theta_{-i}) = 1 \Rightarrow w_i(\theta_i) = \max_j (w_j(\theta_j))$ and $i = i^*$. However, $q_i(\theta_i', \theta_{-i}) = 0 \Rightarrow w_i(\theta_i') < \max_j (w_j(\theta_j))$ or $i > i^*$, a contradiction. □

- To see $t_i(\theta)$ intuitively,
- Define $z_i(\theta_{-i}) = \inf \{\theta_i \mid w_i(\theta_i) \geq \theta_0 \text{ and } w_i(\theta_i) \geq w_j(\theta_j)\ \forall j\}$, which is the minimum possible winning bid given θ_0 and θ_{-i} for i.
- Then, we can define

$$q_i(\theta) = \begin{cases} 1 & \text{if } \theta_i > z_i(\theta_{-i}) \\ 0 & \text{if } \theta_i < z_i(\theta_{-i}) \end{cases}$$

- Then,

$$\int_{\underline{\theta}_i}^{\theta_i} q_i(\theta_i, \theta_{-i}) \, d\theta_i = \begin{cases} \theta_i - z_i(\theta_{-i}) & \text{if } \theta_i \geq z_i(\theta_{-i}) \\ 0 & \text{if } \theta_i < z_i(\theta_{-i}) \end{cases}$$

- Finally,

$$t_i(\theta) = \begin{cases} z_i(\theta_{-i}) & \text{if } q_i(\theta) = 1 \\ 0 & \text{if } q_i(\theta) = 0 \end{cases}$$

- If the bidders are symmetric, i.e. $\Theta_i = [\underline{\theta}, \bar{\theta}]$ and $\phi_i(\theta_i) = \phi(\theta_i)$ for all $i = 1, \ldots, n$, then,

$$z_i(\theta_{-i}) = \max\left\{ w_i^{-1}(\theta_0), \max_{j \neq i} \theta_j \right\}$$

- that is, a modified Vickrey auction where the seller submits a bid or reserved price $w_i^{-1}(\theta_0)$.

Example 9. *$\underline{\theta}_i = 0, \bar{\theta}_i = 100$, and $\phi_i(\theta_i) = \frac{1}{100}$ for all i. Then,*

$$w_i(\theta_i) = \theta_i - \frac{1 - \frac{\theta_i}{100}}{100} = 2\theta_i - 100$$

then w_i is increasing . Suppose, $\theta_0 = 0$, then, the reserved price of the seller is $w_i^{-1}(0) = \frac{0+100}{2} = 50$. Seller aannouses reservation price of 50. And the risk of keeping the object is $\left(\frac{1}{2}\right)^n$. and is not sold. Return: highest price if sold.
Thus optimal auction is not ex post efficient.

Example 10. *For asymmetric bidders, the object may not go to the highest bidder. For example, $\phi_i(\theta_i) = \frac{1}{\bar{\theta}_i - \underline{\theta}_i}$ and $\theta_0 = 0$. Then, $w_i(\theta_i) = 2\theta_i - \bar{\theta}_i$. Therefore, even though $\theta_i > \theta_j$, if $\bar{\theta}_i \ll \bar{\theta}_j$, then, it is possible that $2\theta_i - \bar{\theta}_i < 2\theta_j - \bar{\theta}_j$ so that i could not win the bid.*

In general, the optimal auction may not be ex post efficient.

Example 11. *Interpretation in one buyer case If $n = 1$, then set*

$$q_1^*(\theta_1) = \begin{cases} 1 & \textit{if } w_1^*(\theta_1) \geq \theta_0 \\ 0 & \textit{otherwise} \end{cases}$$

and

$$t_1^*(\theta_1) = q_1^*(\theta_1) \cdot \min\{s_1 : w_1^*(s_1) \geq \theta_0\}$$

- *That is, the seller offer to sell the object at the price $w_1^{*-1}(\theta_0) = \min\{s_1 : w_1^*(s_1) \geq \theta_0\}$.*

- *If i is the only bidder who submit value θ_i, then the seller is willing to sell it if and only $w_i^*(\theta_i)$ is greater than θ_0*

General case Ironing technique

5.1 Nonlinear pricing

Problem Set-up

- One buyer with private information of preference $\theta \sim F$ on $[\underline{\theta}, \bar{\theta}]$ with pdf $f > 0$.
- The buyer's preference is $u(x, t, \theta) = v(x, \theta) - t$ where x is nr of goods, t payment.
- Assume v has the following properties:
  1. $v(0, \theta) = 0$ for all θ
  2. strictly increasing and strictly concave in x, and twice differentiable: $\frac{\partial v}{\partial x} > 0$ and $\frac{\partial^2 v}{\partial x^2} < 0$
  3. single-cross property (SCP) of v, i.e. $\frac{\partial^2 v}{\partial x \partial \theta} > 0$
- Monopolist: The marginal cost of producing one unit of good is constant $c > 0$

Lemma 7. *Single-cross property implies that* $\frac{\partial v}{\partial x} > 0$ *and* $\frac{\partial v}{\partial \theta} > 0$.

Proof.

$$v(x, \theta) = v(0, \underline{\theta}) + \int_{\underline{\theta}}^{\theta} \int_0^x \frac{\partial^2 v}{\partial x \partial \theta}(s, t) ds dt = \int_{\underline{\theta}}^{\theta} \int_0^x \frac{\partial^2 v}{\partial x \partial \theta}(s, t) ds dt \text{ . Therefore,}$$

$$\frac{\partial v}{\partial \theta}(x, \theta) = \int_0^x \frac{\partial^2 v}{\partial x \partial \theta}(s, \theta) ds > 0$$

□

Notation

- $q(\theta)$: amount of non-money allocation for θ;
- $t(\theta)$: amount of money;
- $U(\theta) = v(q(\theta), \theta) - t(\theta)$

The monopoly problem

$$\begin{array}{lll} \max_{q,t} & & E[\pi(\theta)] = \int_{\underline{\theta}}^{\bar{\theta}} [t(\theta) - cq(\theta)] f(\theta) d\theta \\ \text{s.t.} & \text{IR} & U(\theta) \geq 0 \\ & \text{IC} & U(\theta') \geq v(q(\theta), \theta') - t(\theta) \quad \forall \theta, \theta' \\ & & U(\theta) \geq v(q(\theta'), \theta) - t(\theta') \quad \forall \theta, \theta' \end{array}$$

Lemma 8. *The mechanism* (q,t) *is feasible (IR & IC)* $\Longleftrightarrow$

1. q *is monotone increasing,*
2. $U(\theta) = U(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} \frac{\partial v}{\partial \theta}(q(\tau),\tau)d\tau,$
3. $U(\underline{\theta}) \geq 0$

Proof. $\Rightarrow$ Note that

$$\begin{aligned} U(\theta) &\geq v\left(q\left(\theta'\right),\theta\right) - t\left(\theta'\right) \\ &= v\left(q\left(\theta'\right),\theta\right) - v\left(q\left(\theta'\right),\theta'\right) + v\left(q\left(\theta'\right),\theta'\right) - t\left(\theta'\right) \\ &= v\left(q\left(\theta'\right),\theta\right) - v\left(q\left(\theta'\right),\theta'\right) + U\left(\theta'\right) \\ \Leftrightarrow U(\theta) - U\left(\theta'\right) &\geq v\left(q\left(\theta'\right),\theta\right) - v\left(q\left(\theta'\right),\theta'\right) \end{aligned}$$

Similarly,

$$\begin{aligned} U\left(\theta'\right) &\geq v\left(q(\theta),\theta'\right) - t(\theta) \\ \Leftrightarrow U\left(\theta'\right) - U(\theta) &\geq v\left(q(\theta),\theta'\right) - v(q(\theta),\theta) \end{aligned}$$

Therefore,

$$\begin{aligned} v\left(q\left(\theta'\right),\theta'\right) - v\left(q\left(\theta'\right),\theta\right) \geq U\left(\theta'\right) - U(\theta) &\geq v\left(q(\theta),\theta'\right) - v(q(\theta),\theta) \\ \Leftrightarrow \int_{\theta}^{\theta'} \frac{\partial v}{\partial \theta}\left(q\left(\theta'\right),\tau\right)d\tau \geq U\left(\theta'\right) - U(\theta) &\geq \int_{\theta}^{\theta'} \frac{\partial v}{\partial \theta}(q(\theta),\tau)d\tau \\ \Rightarrow \int_{\theta}^{\theta'} \frac{\partial v}{\partial \theta}\left(q\left(\theta'\right),\tau\right)d\tau - &\int_{\theta}^{\theta'} \frac{\partial v}{\partial \theta}(q(\theta),\tau)d\tau \geq 0 \\ \Leftrightarrow &\int_{\theta}^{\theta'} \int_{q(\theta)}^{q(\theta')} \frac{\partial^2 v}{\partial x \partial \theta}(q,\tau)dqd\tau \geq 0 \end{aligned}$$

By SCP of v, the above inequality implies that $\theta \geq \theta' \Rightarrow q(\theta) \geq q\left(\theta'\right)$, i.e. q is monotone. Moreover, by Envelop Theorem,

$$\begin{aligned} U(\theta) &\geq v\left(q\left(\theta'\right),\theta\right) - t\left(\theta'\right) \forall \theta' \\ \Rightarrow U'(\theta) &= \frac{\partial v}{\partial \theta}(q(\theta),\theta) \text{ a.e.} \end{aligned}$$

Therefore, $U(\theta) = U(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} \frac{\partial v}{\partial \theta}(q(\tau),\tau)d\tau$. Finally, to show IC, $\forall \theta, U(\theta) = U(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} \frac{\partial v}{\partial \theta}(q(\tau),\tau)d\tau \geq 0$ by previous lemma and $U(\underline{\theta}) \geq 0$.

$\Leftarrow$ Clearly, IR is satisfied. Suppose, for contradiction, not IC. Then, wlog $\exists \theta > \theta'$:

$$U(\theta) < v\left(q\left(\theta'\right),\theta\right) + t\left(\theta'\right)$$

Then, LHS is:

$$U(\theta)=U(\underline{\theta})+\int_{\underline{\theta}}^{\theta}\frac{\partial v}{\partial\theta}(q(\tau),\tau)d\tau$$

and the RHS is:

$$\begin{aligned}
&v\left(q(\theta),\theta'\right)-t(\theta)\\
=&v\left(q\left(\theta'\right),\theta\right)-v\left(q\left(\theta'\right),\theta'\right)+U\left(\theta'\right)\\
=&v\left(q\left(\theta'\right),\theta\right)-v\left(q\left(\theta'\right),\theta'\right)+U(\underline{\theta})+\int_{\underline{\theta}}^{\theta'}\frac{\partial v}{\partial\theta}(q(\tau),\tau)d\tau.
\end{aligned}$$

Therefore, we have:

$$\begin{aligned}
&U(\underline{\theta})+\int_{\underline{\theta}}^{\theta}\frac{\partial v}{\partial\theta}(q(\tau),\tau)d\tau\\
&<v\left(q\left(\theta'\right),\theta\right)-v\left(q\left(\theta'\right),\theta'\right)+U(\underline{\theta})+\int_{\underline{\theta}}^{\theta'}\frac{\partial v}{\partial\theta}(q(\tau),\tau)d\tau\\
\Rightarrow&\int_{\theta'}^{\theta}\frac{\partial v}{\partial\theta}(q(\tau),\tau)d\tau<\int_{\theta'}^{\theta}\frac{\partial v}{\partial\theta}\left(q\left(\theta'\right),\tau\right)d\tau\\
\Rightarrow&\int_{\theta'}^{\theta}\int_{q(\theta')}^{q(\tau)}\frac{\partial^2 v}{\partial x\partial\theta}(q,\tau)dqd\tau<0
\end{aligned}$$

By SCP of v, the above inequality implies that $\exists\tau\in(\theta',\theta]:q(\tau)<q\left(\theta'\right)$, contradiction to monotonicity. □

By the above Lemma, we could rewrite the monopoly's problem, substituting $t(\theta)$ and $U(\theta)$, as follow:

$$\begin{aligned}
&\max_{q,t}E[\pi(\theta)]=\int_{\underline{\theta}}^{\bar{\theta}}\left[v(q(\theta),\theta)-U(\underline{\theta})-\int_{\underline{\theta}}^{\theta}\tfrac{\partial v}{\partial\theta}(q(\tau),\tau)d\tau-cq(\theta)\right]f(\theta)d\theta\\
&\text{s.t. } U(\underline{\theta})\geq 0\quad q'\geq 0.
\end{aligned}$$

Using the similar trick,

$$\begin{aligned}
&\int_{\underline{\theta}}^{\bar{\theta}}\left(\int_{\underline{\theta}}^{\theta}\frac{\partial v}{\partial\theta}(q(\tau),\tau)d\tau\right)f(\theta)d\theta\\
=&\int_{\underline{\theta}}^{\bar{\theta}}\int_{\tau}^{\bar{\theta}}\frac{\partial v}{\partial\theta}(q(\tau),\tau)f(\theta)d\theta d\tau\\
=&\int_{\underline{\theta}}^{\bar{\theta}}\frac{\partial v}{\partial\theta}(q(\tau),\tau)[1-F(\tau)]d\tau
\end{aligned}$$

Therefore,

$$E[\pi(\theta)]=\int_{\underline{\theta}}^{\bar{\theta}}\left[v(q(\theta),\theta)-U(\underline{\theta})-\frac{\partial v}{\partial\theta}(q(\theta),\theta)\frac{1-F(\theta)}{f(\theta)}-cq(\theta)\right]f(\theta)d\theta$$

Pointwise FOC:

$$\frac{\partial v}{\partial q}(q(\theta),\theta)-c-\frac{1-F(\theta)}{f(\theta)}\frac{\partial^2 v}{\partial x\partial\theta}(q(\theta),\theta)=0$$

If there is no distribution at top, $\frac{1-F(\bar{\theta})}{f(\bar{\theta})} = 0$, then no distortion for the top, i.e. $\frac{\partial v}{\partial q}(q(\bar{\theta}), \bar{\theta}) = c$. Efficiency means Marginal Benefit=Marginal Cost

6 Static Mirrlees taxation

Standard assumption in the Ramsey literature is that lump sum taxes are not allowed. Why aren't lump sum taxes used in practice?
One reason for this is they require truthful elicitation of agents characteristics, which might not be publicly observable. Moreover, agents might not have an incentive to reveal these characteristics truthfully.

We will consider a mechanism design problem in which agents true ability types are private and allow the designer to use arbitrary mechanisms and transfer schedules to achieve efficiency. Next, will consider implementations/decentralizations.

6.1 A Two Type Example

Consider an environment with a continuum of HHs characterized by a productivity level $\theta \in \Theta = \{\theta_H, \theta_L\}$ with $\theta_H > \theta_L > 0$.
A household of type θ who works l hours can produce $y = \theta l$ of output. Let $\pi(\theta)$ denote the probability that given household is of type θ. By the LLN, this is also the fraction of HHs with productivity θ.

Household preferences are given by $u(c) - v(l)$ but we will use $l = \frac{y}{\theta}$ to define the preferences as $U(c, y, \theta) = u(\mathrm{c}) - v\left(\frac{y}{\theta}\right)$. Assume $u' > 0 < u''$ and $v', v'' > 0$.

Suppose first that HH productivities are public information. Under full information a utilitarian planner (cares about all types equally) solves

$$\max_{c(\theta), y(\theta)} \pi(\theta_H)\left[u(c(\theta_H)) - v\left(\frac{y(\theta_H)}{\theta_H}\right)\right] + \pi(\theta_L)\left[u(c(\theta_L)) - v\left(\frac{y(\theta_L)}{\theta_L}\right)\right]$$

subject to

$$\pi(\theta_\mathrm{H})\,\mathrm{c}(\theta_\mathrm{H}) + \pi(\theta_\mathrm{L})\,c(\theta_\mathrm{L}) \leqslant \pi(\theta_\mathrm{H})\,\mathrm{y}(\theta_\mathrm{H}) + \pi(\theta_\mathrm{L})\,\mathrm{y}(\theta_\mathrm{L})$$

Let μ be the multiplier on the RC. Then the the FOCs are

$$u'(c(\theta_H)) = \mu = u'(c(\theta_L))$$
$$\frac{1}{\theta_H} v'(l(\theta_H)) = \mu = \frac{1}{\theta_L} v'(l(\theta_L))$$

Which implies

$$c(\theta_H) = c(\theta_L)$$
$$c(\theta) = \frac{1}{\theta} v'(l(\theta))$$
$$\frac{v'(l(\theta_H))}{v'(l(\theta_L))} = \frac{\theta_H}{\theta_L} > 1$$

where the last equation implies that $l(\theta_H) > l(\theta_L)$ since v is convex.

Now suppose that θ is private information. It is easy to see that the above allocation is not incentive compatible. A high type households strictly prefers to pretend to be a low type since the consumption levels are the same but hours worked is lower. From the revelation principle, that we can restrict ourselves to direct revelation mechanisms.

Definition 56 (Direct revelation mechanism)**.** *consists of action/message sets $M_i, i \in [0,1]$ such that for each i, $M_i = \Theta_i$ and outcome functions (c, y) where $c, y : \Theta \to \mathbb{R}_+$.*

Since there is no aggregate uncertainty (LLN), we will consider mechanisms that treat households anonymously, i.e. mechanisms that are independent of i.

Definition 57 (Revelation mechanism)**.** *is*

1. ***Incentive compatible (IC)*** *if and only if*

$$u(c_H) - v(l_H) \geq u(c_L) - v\left(\frac{\theta_L}{\theta_H} l_L\right) \tag{1}$$

$$u(c_L) - v(l_L) \geq u(c_H) - v\left(\frac{\theta_H}{\theta_L} l_H\right)$$

2. ***Resource feasible (FEAS)*** *if and only if*

$$\pi(\theta_H) c(\theta_H) + \pi(\theta_L) c(\theta_L) \leqslant \pi(\theta_H) y(\theta_H) + \pi(\theta_L) y(\theta_L) \tag{2}$$

Then, the Planner's/Mechanism designer's problem is

$$\max_{c(\theta), y(\theta)} \pi(\theta_H) \left[u(c(\theta_H)) - v\left(\frac{y(\theta_H)}{\theta_H}\right)\right] + \pi(\theta_L) \left[u(c(\theta_L)) - v\left(\frac{y(\theta_L)}{\theta_L}\right)\right]$$
$$\text{s.t. IC1, IC2, FEAS}$$

Notice that there are two incentive compatibility constraints, one for the high and one for the low type.

6.2 The relaxed problem

This problem is typically not concave/convex - concave objective function and a convex constraint set. This is because the $u(c)'s$ (and $v's$) appear on both the left and right hand side of the (IC) constraints. We need to reformulate our problem so we can use KKT.

Before that let's note some simple properties of contracts, i.e. combinations of (c_L, y_L) and (c_H, y_H) that must be true if (FEAS), (IC1), and (IC2) are all satisfied.

- Suppose $c_H > c_L$ but $y_H \leq y_L$. If this were true, then,

$$u(c_H) - v\left(\frac{y_H}{\theta_L}\right) > u(c_L) - v\left(\frac{y_H}{\theta_L}\right)$$

  since $c_H > c_L$ and $u(\cdot)$ is monotone. Moreover,

$$u(c_L) - v\left(\frac{y_H}{\theta_L}\right) \geq u(c_L) - v\left(\frac{y_L}{\theta_L}\right)$$

  since $y_L \geq y_H$ and $v(\cdot)$ is monotone. Thus,

$$u(c_H) - v\left(\frac{y_H}{\theta_L}\right) > u(c_L) - v\left(\frac{y_L}{\theta_L}\right)$$

  But this violates (IC2) and hence, these types of allocations are not feasible.

- A similar argument shows that combinations with $c_H \geq c_L$ and $y_H < y_L$ also are not feasible.

- Suppose $c_H < c_L$ but $y_L \leq y_H$. If this were true, as above, we would have

$$u(c_L) - v\left(\frac{y_L}{\theta_H}\right) > u(c_H) - v\left(\frac{y_H}{\theta_H}\right)$$

  i.e. (IC1) would be violated.

- A similar argument holds if $c_H \leq c_L$ but $y_L < y_H$. So, this cannot be the case in the solution.

We can summarize these in the following lemma.

Lemma 9. *If the contract (c_L, y_L) and (c_H, y_H) satisfies* (FEAS), (IC1), *and* (IC2) *in Problem (SP2), then, one of the following three configurations must hold:*

1. *$c_H > c_L$ and $y_H > y_L$*

2. *$c_L > c_H$ and $y_L > y_H$; or,*

3. *$c_L = c_H$ and $y_L = y_H$*

Here we show that only first allocation can be optimal and other two are impossible. To do so we will use **Variational methods**

Proof. • $c_L > c_H$ and $y_L > y_H$; is not possible.

Assume $(c_L, y_L) > (c_H, y_H)$. By (IC2),

$$u(c_L) - u(c_H) - \left[v\left(\frac{y_L}{\theta_L}\right) - v\left(\frac{y_H}{\theta_L}\right)\right] \geq 0$$

then by taking integral representation of last two we obtain:

$$u(c_L) - u(c_H) \geq \frac{1}{\theta_L}\int_{y_H}^{y_L} v'\left(\frac{y}{\theta_L}\right) dy$$

Notice that the first term, $u(c_L) - u(c_H)$ is positive, since c_L is assumed to be larger than c_H. Also, since $v'(\cdot) > 0$ and $y_L > y_H$, it follows that the second term is also positive. But since $\theta_H > \theta_L$ and $v(\cdot)$ is convex, it follows that $v'(y/\theta_H) < v'(y/\theta_L)$ for all y, and, as a result,

$$\int_{y_H}^{y_L} v'\left(\frac{y}{\theta_H}\right) dy < \int_{y_H}^{y_L} v\left(\frac{y}{\theta_L}\right) dy$$

Since $\theta_H > \theta_L$, we also have:

$$\frac{1}{\theta_H}\int_{y_H}^{y_L} v'\left(\frac{y}{\theta_H}\right) dy < \frac{1}{\theta_L}\int_{y_H}^{y_L} v'\left(\frac{y}{\theta_L}\right) dy$$

Thus,

$$\begin{aligned} u(c_L) - u(c_H) &\geq \frac{1}{\theta_L}\int_{y_H}^{y_L} v'\left(\frac{y}{\theta_L}\right) dy \\ &> \frac{1}{\theta_H}\int_{y_H}^{y_L} v'\left(\frac{y}{\theta_H}\right) dy \end{aligned}$$

implying that,

$$u(c_L) - u(c_H) - \frac{1}{\theta_H}\int_{y_H}^{y_L} v'\left(\frac{y}{\theta_H}\right) dy > 0$$

$$\therefore u(c_L) - u(c_H) - \left[v\left(\frac{y_L}{\theta_H}\right) - v\left(\frac{y_H}{\theta_H}\right)\right] > 0$$

That is high-type agents prefer (c_L, y_L) over (c_H, y_H), violating (IC1).

- $c_L = c_H$ and $y_L = y_H$

To show this formally, suppose (c, y) denotes the common consumption/production pair. We consider the following cases:

1. If
$$u'(c) < \frac{1}{\theta_H} v'\left(\frac{y}{\theta_H}\right)$$
since $\theta_H > \theta_L$, we have:
$$u'(c) < \frac{1}{\theta_L} v'\left(\frac{y}{\theta_L}\right)$$
Therefore, decreasing c and y at the same time by a small amount will keep the (IC)s holding and will be strictly better for both types.

2. If
$$u'(c) = \frac{1}{\theta_H} v'\left(\frac{y}{\theta_H}\right)$$
again, we have:
$$u'(c) < \frac{1}{\theta_L} v'\left(\frac{y}{\theta_L}\right)$$
Hence, decreasing consumption and production of the low types would make them better off, while high types have no incentives to deviate to the new allocation.

3. If
$$u'(c) > \frac{1}{\theta_H} v'\left(\frac{y}{\theta_H}\right) \text{ and } u'(c) > \frac{1}{\theta_L} v'\left(\frac{y}{\theta_L}\right)$$
by the same argument as in the first case, it is optimal to increase y and c at the same time. The case of
$$u'(c) > \frac{1}{\theta_H} v'\left(\frac{y}{\theta_H}\right) \text{ and } u'(c) = \frac{1}{\theta_L} v'\left(\frac{y}{\theta_L}\right)$$
is not optimal either, by the same logic as in the second case.

4. If
$$u'(c) > \frac{1}{\theta_H} v'\left(\frac{y}{\theta_H}\right) \text{ and } u'(c) < \frac{1}{\theta_L} v'\left(\frac{y}{\theta_L}\right)$$
an increase in the consumption of high types followed by an increase in their production, and a decrease in the consumption of low types followed by a decrease in their production would leave both types better off.

□

Lemma 10. *(IC1) holds with equality*

Proof. We will use **perturbation argument**.We will do this by supposing it is false and constructing a better contract. The dominating contract that we will construct will have better insurance over c without disrupting (IC1). Suppose that (IC1) is not satisfied at equality;

$$u\left(c_H\right)-v\left(\frac{y_H}{\theta_H}\right)>u\left(c_L\right)-v\left(\frac{y_L}{\theta_H}\right)$$

Notice that, if this holds, by continuity, it will still hold if we add a bit to c_L and subtract a bit from c_H

$$u\left(c_H-\varepsilon\right)-v\left(\frac{y_H}{\theta_H}\right)>u\left(c_L+\delta\right)-v\left(\frac{y_L}{\theta_H}\right)$$

as long as ε and δ are small enough. Consider the alternative contract given by $\left(c_H-\varepsilon, y_H\right)$ and $\left(c_L+\delta, y_L\right)$. Choose $\delta=\pi_H\varepsilon/\pi_L$. Then, if ε is small enough, (IC1) will still hold, and (FEAS) becomes:

$$\begin{aligned}\pi_H\left(c_H-\varepsilon\right)+\pi_L\left(c_L+\delta\right)&=\pi_H c_H+\pi_L c_L+\pi_H\varepsilon-\pi_L\frac{\pi_H}{\pi_L}\varepsilon\\&=\pi_H c_H+\pi_L c_L\end{aligned}$$

Thus, (FEAS) will hold because we didn't change y_L or y_H, and because of the way we constructed δ.

So, we only need to show that welfare goes up from this change, even when ε is small but positive. To see this note that the change in welfare is given by:

$$\Delta W=\pi_H\left[u\left(c_H-\varepsilon\right)-u\left(c_H\right)\right]+\pi_L\left[u\left(c_L+\frac{\pi_H}{\pi_L}\varepsilon\right)-u\left(c_L\right)\right]$$

The terms involving the y 's do not appear in this, since they are unchanged. If we take the derivative with respect to ε, at $\varepsilon=0$, we have:

$$\begin{aligned}\left.\frac{d\Delta W}{d\varepsilon}\right|_{\varepsilon=0}&=-\pi_H u'\left(c_H\right)+\pi_L\frac{\pi_H}{\pi_L}u'\left(c_L\right)\\&=-\pi_H u'\left(c_H\right)+\pi_H u'\left(c_L\right)\\&=\pi_H\left[u'\left(c_L\right)-u'\left(c_H\right)\right]\\&>0\end{aligned}$$

since $c_L<c_H$, and $u(\cdot)$ is assumed to be strictly concave. □

Lemma 11. *If (IC1) holds with equality then (IC2) is satisfied for $c_H > c_L$ and $y_H > y_L$*

Proof. Suppose not and that this constraint was violated. Then

$$
\begin{aligned}
& u\left(c\left(\theta_L\right)\right)-v\left(\frac{y\left(\theta_L\right)}{\theta_L}\right)<u\left(c\left(\theta_H\right)\right)-v\left(\frac{y\left(\theta_H\right)}{\theta_L}\right) \\
\Longrightarrow & v\left(\frac{y\left(\theta_H\right)}{\theta_L}\right)-v\left(\frac{y\left(\theta_L\right)}{\theta_L}\right)<u\left(c\left(\theta_H\right)\right)-u\left(c\left(\theta_L\right)\right) \\
\Longrightarrow & \frac{1}{\theta_L} \int_{y\left(\theta_L\right)}^{y\left(\theta_H\right)} v^{\prime}\left(\frac{y}{\theta_L}\right) d y<u\left(c\left(\theta_H\right)\right)-u\left(c\left(\theta_L\right)\right) \\
\Longrightarrow & \frac{1}{\theta_H} \int_{y\left(\theta_L\right)}^{y\left(\theta_H\right)} v^{\prime}\left(\frac{y}{\theta_H}\right) d y<\frac{1}{\theta_L} \int_{y\left(\theta_L\right)}^{y\left(\theta_H\right)} v^{\prime}\left(\frac{y}{\theta_L}\right) d y<u\left(c\left(\theta_H\right)\right)-u\left(c\left(\theta_L\right)\right) \\
\Longrightarrow & v\left(\frac{y\left(\theta_H\right)}{\theta_H}\right)-v\left(\frac{y\left(\theta_L\right)}{\theta_H}\right)<u\left(c\left(\theta_H\right)\right)-u\left(c\left(\theta_L\right)\right) \\
\Longrightarrow & u\left(c\left(\theta_L\right)\right)-v\left(\frac{y\left(\theta_L\right)}{\theta_H}\right)<u\left(c\left(\theta_H\right)\right)-v\left(\frac{y\left(\theta_H\right)}{\theta_H}\right)
\end{aligned}
$$

which contradicts the IC for the high type holding with equality. □

Finally we are ready to use KKT to solve **Relaxed problem**

$$
\max_{c(\theta), y(\theta)} \pi\left(\theta_H\right)\left[u\left(c\left(\theta_H\right)\right)-v\left(\frac{y\left(\theta_H\right)}{\theta_H}\right)\right]+\pi\left(\theta_L\right)\left[u\left(c\left(\theta_L\right)\right)-v\left(\frac{y\left(\theta_L\right)}{\theta_L}\right)\right]
$$

subject to

$$
\begin{gathered}
u\left(c\left(\theta_H\right)\right)-v\left(\tfrac{y\left(\theta_H\right)}{\theta_H}\right)=u\left(c\left(\theta_L\right)\right)-v\left(\tfrac{y\left(\theta_L\right)}{\theta_H}\right) \\
\pi\left(\theta_H\right) c\left(\theta_H\right)+\pi\left(\theta_L\right) c\left(\theta_L\right) \leqslant \pi\left(\theta_H\right) y\left(\theta_H\right)+\pi\left(\theta_L\right) c\left(\theta_L\right)
\end{gathered}
$$

Let λ be the multiplier on the first constraint and μ on the second.

The FOCs are

$$
\pi\left(\theta_{\mathrm{H}}\right) u^{\prime}\left(c\left(\theta_{\mathrm{H}}\right)\right)+\lambda u^{\prime}\left(c\left(\theta_{\mathrm{H}}\right)\right)-\pi\left(\theta_{\mathrm{H}}\right) \mu=0 \tag{3}
$$

$$
\pi\left(\theta_{\mathrm{L}}\right) u^{\prime}\left(c\left(\theta_{\mathrm{L}}\right)\right)-\lambda u^{\prime}\left(c\left(\theta_{\mathrm{L}}\right)\right)-\pi\left(\theta_{\mathrm{L}}\right) \mu=0 \tag{4}
$$

$$
-\frac{\pi\left(\theta_{\mathrm{H}}\right)}{\theta_{\mathrm{H}}} v^{\prime}\left(\frac{\mathrm{y}\left(\theta_{\mathrm{H}}\right)}{\theta_{\mathrm{H}}}\right)-\lambda \frac{1}{\theta_{\mathrm{H}}} v^{\prime}\left(\frac{\mathrm{y}\left(\theta_{\mathrm{H}}\right)}{\theta_{\mathrm{H}}}\right)+\pi\left(\theta_{\mathrm{H}}\right) \mu=0 \tag{5}
$$

$$
-\frac{\pi\left(\theta_{\mathrm{L}}\right)}{\theta_{\mathrm{L}}} v^{\prime}\left(\frac{\mathrm{y}\left(\theta_{\mathrm{L}}\right)}{\theta_{\mathrm{L}}}\right)+\lambda \frac{1}{\theta_{\mathrm{H}}} v^{\prime}\left(\frac{\mathrm{y}\left(\theta_{\mathrm{L}}\right)}{\theta_{\mathrm{H}}}\right)+\pi\left(\theta_{\mathrm{L}}\right) \mu=0 \tag{6}
$$

Combining (3) and (5) we obtain

$$u'\left(c\left(\theta_H\right)\right) = \frac{1}{\theta_H} v'\left(\frac{y\left(\theta_H\right)}{\theta_H}\right)$$

This says that the just as in the unconstrained problem, for the high type, the marginal utility of c equals the marginal disutility of working. In particular, the allocation for the high type households is ex-post efficient. This is sometimes referred to as "**no-distortion at the top**".

The mechanical reason for this is that, no type wants to pretend to be the high type and thus the planner does not need to distort. This will not be true for the low type. Next, combine (3) and (4) :

$$\frac{u'\left(c\left(\theta_H\right)\right)}{u'\left(c\left(\theta_L\right)\right)} = \frac{\pi\left(\theta_H\right)}{\pi\left(\theta_L\right)} \frac{\left[\pi\left(\theta_L\right) - \lambda\right]}{\left[\pi\left(\theta_H\right) + \lambda\right]} = \frac{\pi\left(\theta_H\right)\pi\left(\theta_L\right) - \lambda\pi\left(\theta_H\right)}{\pi\left(\theta_H\right)\pi\left(\theta_L\right) + \lambda\pi\left(\theta_L\right)} < 1$$

$$\Rightarrow u'\left(c\left(\theta_H\right)\right) < u'\left(c\left(\theta_L\right)\right) \quad \Rightarrow c\left(\theta_H\right) > c\left(\theta_L\right)$$

But then given the IC holds with equality, it must be that

$$v\left(\frac{\mathrm{y}\left(\theta_{\mathrm{H}}\right)}{\theta_{\mathrm{H}}}\right) > v\left(\frac{\mathrm{y}\left(\theta_{\mathrm{L}}\right)}{\theta_{\mathrm{H}}}\right)$$

$$\Rightarrow \mathrm{y}\left(\theta_{\mathrm{H}}\right) > \mathrm{y}\left(\theta_{\mathrm{L}}\right)$$

To see the allocation for the low type is distorted, combine (4) and (6) to obtain

$$\begin{aligned} u'\left(c\left(\theta_L\right)\right) &= \frac{1}{\theta_L} v'\left(\frac{y\left(\theta_L\right)}{\theta_L}\right) + \frac{\lambda}{\pi\left(\theta_L\right)}\left[u'\left(c\left(\theta_L\right)\right) - \frac{1}{\theta_H} v'\left(\frac{y\left(\theta_L\right)}{\theta_H}\right)\right] > \\ &> \frac{1}{\theta_L} v'\left(\frac{y\left(\theta_L\right)}{\theta_L}\right) + \frac{\lambda}{\pi\left(\theta_L\right)}\left[u'\left(c\left(\theta_H\right)\right) - \frac{1}{\theta_H} v'\left(\frac{y\left(\theta_H\right)}{\theta_H}\right)\right] \\ &= \frac{1}{\theta_L} v'\left(\frac{y\left(\theta_L\right)}{\theta_L}\right) \end{aligned}$$

To make sure that indeed we have solution which maximizes welfare need to check SOCs

$$\frac{d^2 L}{d(c_H)^2}: \quad \pi\left(\theta_{\mathrm{H}}\right) u''\left(c\left(\theta_{\mathrm{H}}\right)\right) + \lambda u''\left(c\left(\theta_{\mathrm{H}}\right)\right)$$

$$\frac{d^2 L}{d(c_L)^2}: \quad \pi\left(\theta_{\mathrm{L}}\right) u''\left(c\left(\theta_{\mathrm{L}}\right)\right) - \lambda u''\left(c\left(\theta_{\mathrm{L}}\right)\right)$$

$$\frac{d^2 L}{d(y_H)^2}: \quad -\frac{\pi\left(\theta_{\mathrm{H}}\right)}{\theta_{\mathrm{H}}^2} v''\left(\frac{\mathrm{y}\left(\theta_{\mathrm{H}}\right)}{\theta_{\mathrm{H}}}\right) - \lambda \frac{1}{\theta_{\mathrm{H}}^2} v''\left(\frac{\mathrm{y}\left(\theta_{\mathrm{H}}\right)}{\theta_{\mathrm{H}}}\right)$$

$$\frac{d^2L}{d(y_L)^2}: \quad -\frac{\pi\left(\theta_{\mathrm{L}}\right)}{\theta_{\mathrm{L}}^2}v''\left(\frac{\mathrm{y}\left(\theta_{\mathrm{L}}\right)}{\theta_{\mathrm{L}}}\right)+\lambda\frac{1}{\theta_{\mathrm{H}}}v''\left(\frac{\mathrm{y}\left(\theta_{\mathrm{L}}\right)}{\theta_{\mathrm{H}}^2}\right)$$

And all other cross derivatives are 0. Observe that from (1)

$$-\pi(\theta_H) < \lambda = \frac{\pi(\theta_L)u'(c_L) - \pi(\theta_H)u'(c_H)}{u'(c_H) + u'(c_L)} < \pi_(\theta_L)$$

Then all second order non zero derivatives are negative, so Hessian of our Langrasian is negative definite on whole $\mathbb{R}^4$ (in particular it is negative definite on kernel of linear epimorphism of Jacobian generated by constraints). So indeed we have solution to our problem whih maximizes welfare.

6.3 Don't distort at the top

Above tells us nothing about implementation, i.e. whether there exist tax systems, for example such that the equilibrium, given the tax system gives efficient allocation. We turn to this next. In particular, we will show that a non-linear income tax schedule can implement the efficient allocation.

Denote the optimal mechanism by (c^*, y^*). Define a tax function $T(y) = y - c$ if $y \in \{y^*(\theta_H), y^*(\theta_L)\}$ and $T(y) = y$ otherwise. Given this tax function, the household of type θ solves:

$$\max_{c,y} u(c) - v\left(\frac{y}{\theta}\right)$$
$$\text{st} \quad c \leq y - T(y)$$

The first order condition is

$$u'(c)\left(1 - T'(y)\right) = \frac{1}{\theta}v'\left(\frac{y}{\theta}\right)$$

Comparing this equation to the one in the planning problem implies that $T'(y_H^*) = 0$ and $T'(y_L^*) > 0$

6.4 Mirrlees with a continuum of types

We now consider a problem with a continuum of types. We characterize the efficient allocation and derive the Diamond-Mirrlees-Saez formula.

The main issue for continuum of types is how to simplify IC. In the two type case, we just dropped one but here we can not do things like that. But we have already seen how to deal with these constraints! We will use similar Myerson(1981)- like techniques and replace incentive compatibility with an local condition and a monotonicity condition.

As before, an allocation (c, y) is incentive compatible (GIC) if and only if

$$U(\theta) \equiv u(c(\theta)) - v\left(\frac{y(\theta)}{\theta}\right) \geq u(c(\hat{\theta})) - v\left(\frac{y(\hat{\theta})}{\theta}\right) \equiv u(\hat{\theta}, \theta) \quad \forall \theta, \theta \in \Theta$$

Lemma 12. *An allocation* (c, y) *satisfies global incentive compatibility* $\Longleftrightarrow$

1. $\mathrm{y}(\theta)$ *is increasing in* θ
2. $$u'(c(\theta))c'(\theta) = \frac{1}{\theta} y'(\theta) v'\left(\frac{y(\theta)}{\theta}\right)$$

As a result of lemma, we can now write down a relaxed planning problem

$$\max_{c,y} \int W\left(u(c(\theta)) - v\left(\frac{y(\theta)}{\theta}\right)\right) dF$$

subject to

$$\begin{gathered} \int_{\Theta} \mathrm{c}(\theta) dF(\theta) \leqslant \int_{\Theta} \mathrm{y}(\theta) dF(\theta) \\ \mathrm{u}'(\mathrm{c}(\theta))\mathrm{c}'(\theta) = \tfrac{1}{\theta} y'(\theta) v'\left(\tfrac{\mathrm{y}(\theta)}{\theta}\right) \\ \mathrm{y}'(\theta) \geqslant 0 \end{gathered}$$

Here F is the cdf of θ and W is a general weighting function instead of just assuming a utilitarian planner. This problem is still pretty intractable. We will use one more trick: replace the derivative condition with an **envelope condition**. In particular the derivative condition is equivalent to

$$\mathrm{U}(\theta) = \max_{\hat{\theta} \in \Theta} \mathrm{U}(\hat{\theta}, \theta)$$

The envelope condition for this maximization problem is

$$\begin{aligned} \mathrm{u}'(\theta) &= \frac{\partial}{\partial \theta} u(c(\hat{\theta})) - v\left(\frac{\mathrm{y}(\hat{\theta})}{\theta}\right)\Bigg|_{\hat{\theta}=\theta} \\ &= \frac{\mathrm{y}(\theta)}{\theta^2} v'\left(\frac{\mathrm{y}(\theta)}{\theta}\right) \end{aligned}$$

For a final time, the planning problem is

$$\max_{c,y} \int W(U(\theta))dF(\theta)$$

$$\text{st} \quad U(\theta) = u(c(\theta)) - v\left(\frac{y(\theta)}{\theta}\right)$$
$$\int_\Theta c(\theta)dF(\theta) \leqslant \int_\Theta y(\theta)dF(\theta)$$
$$U'(\theta) = \frac{y(\theta)}{\theta^2} v'\left(\frac{y(\theta)}{\theta}\right)$$
$$y'(\theta) \geqslant 0$$

A common mechanism design trick is to drop the monotonicity condition and check that the result allocation satisfies it ex-post. If it is violated then it usually means that there is "bunching". To deal with this situation we use an **ironing method**- see Myerson.

To solve the problem above we use techniques from the **calculus of variation**. We can write the lagrangian

$$\begin{aligned}\mathcal{L} = &\int W(U(\theta))dF(\theta) + \int \gamma(\theta)\left[u(c(\theta)) - v\left(\frac{y(\theta)}{\theta}\right) - U(\theta)\right] d\theta \\ &+ \lambda\left[\int_\Theta y(\theta)dF(\theta) - \int_\Theta c(\theta)dF(\theta)\right] \\ &+ \int \mu(\theta)\left[U'(\theta) - \frac{y(\theta)}{\theta^2}v'\left(\frac{y(\theta)}{\theta}\right)\right] d\theta\end{aligned}$$

Lets first deal with the $U'(\theta)$ term:

$$\begin{aligned}\int \mu(\theta)U'(\theta)d\theta &= \int \mu(\theta)dU(\theta) \\ &= \mu(\bar{\theta})U(\bar{\theta}) - \mu(\underline{\theta})U(\underline{\theta}) - \int \mu'(\theta)U(\theta)d\theta\end{aligned}$$

Now substitute this back in

$$\mathcal{L} = \int W(U(\theta))dF(\theta) + \int \gamma(\theta)\left[u(c(\theta)) - v\left(\frac{y(\theta)}{\theta}\right) - U(\theta)\right] d\theta$$

$$+\lambda\left[\int_\Theta y(\theta)dF(\theta) - \int_\Theta c(\theta)dF(\theta)\right]$$
$$-\int \mu(\theta)\frac{y(\theta)}{\theta^2}v'\left(\frac{y(\theta)}{\theta}\right)d\theta - \int \mu'(\theta)U(\theta)d\theta + \mu(\bar{\theta})u(\bar{\theta}) - \mu(\underline{\theta})U(\underline{\theta})$$

and take first order conditions:

$$U(\theta) : W'(U(\theta))f(\theta) - \gamma(\theta) - \mu'(\theta) = 0 \tag{7}$$

$$c(\theta) : \gamma(\theta)u'(c(\theta)) - \lambda f(\theta) = 0 \tag{8}$$

$$y(\theta) : -\gamma(\theta)\frac{1}{\theta}\nu'\left(\frac{y(\theta)}{\theta}\right) + \lambda f(\theta) - \mu(\theta)\left[\frac{1}{\theta^2}\nu'\left(\frac{y(\theta)}{\theta}\right) + \frac{y(\theta)}{\theta^3}v''\left(\frac{y(\theta)}{\theta}\right)\right] = 0 \tag{9}$$

We also have two boundary conditions: $\mu(\bar{\theta}) = \mu(\underline{\theta}) = 0$.

To see why these must hold notice that if $\mu(\bar{\theta}) > 0(\mu(\underline{\theta}) > 0)$ then the planner would like to set $\mathrm{U}(\bar{\theta}) = \infty(\mathrm{U}(\underline{\theta}) = -\infty)$ which would clearly violate incentive constraints. Using (7) and (8) we have

$$\begin{aligned}\mu(\theta) &= \int_{\theta}^{\bar{\theta}} \left[\gamma(z) - W'(\mathrm{U}(z))f(z)\right] \mathrm{d}z \\ &= \int_{\theta}^{\bar{\theta}} \left[\frac{\lambda f(z)}{u'(c(z))} - W'(\mathrm{U}(z))f(z)\right] \mathrm{d}z\end{aligned}$$

Then (9) becomes

$$\lambda f(\theta) - \frac{\lambda f(\theta)}{u'(c(\theta))}\frac{1}{\theta}v'\left(\frac{y(\theta)}{\theta}\right) =$$
$$= \left[\frac{1}{\theta^2}\nu'\left(\frac{y(\theta)}{\theta}\right) + \frac{y(\theta)}{\theta^3}v''\left(\frac{y(\theta)}{\theta}\right)\right] \int_{\theta}^{\bar{\theta}} \left[\frac{\lambda f(z)}{u'(c(z))} - W'(u(z))f(z)\right] dz$$

Divide both sides by $\frac{1}{\theta}v'\left(\frac{y(\theta)}{\theta}\right)$

$$\frac{1}{\frac{1}{\theta}v'\left(\frac{y(\theta)}{\theta}\right)} - \frac{1}{u'(c(\theta))} = \frac{1-F(\theta)}{\theta f(\theta)}\left[1 + \frac{y(\theta)}{\theta}\frac{v''\left(\frac{y(\theta)}{\theta}\right)}{v'\left(\frac{y(\theta)}{\theta}\right)}\right] \int_{\theta}^{\bar{\theta}} \left[\frac{1}{u'(c(z))} - \frac{W'(\mathrm{U}(z))}{\lambda}\right] \frac{dF(z)}{1-F(\theta)}$$

Now recall that we are interested in implementing the efficient allocation with a tax function $\mathrm{T(y)}$. In the decentralized problem $u'(\mathrm{c}(\theta))\left(1 - \mathrm{T}'(\mathrm{y}(\theta))\right) = \frac{1}{\theta}v'\left(\frac{\mathrm{y}(\theta)}{\theta}\right)$. Therefore the above equation can be written as

$$\frac{1}{u'(c(\theta))\left(1 - T'(y(\theta))\right)} - \frac{1}{u'(c(\theta))} =$$
$$\frac{1-F(\theta)}{\theta f(\theta)}\left[1 + \frac{y(\theta)}{\theta}\frac{v''\left(\frac{y(\theta)}{\theta}\right)}{v'\left(\frac{y(\theta)}{\theta}\right)}\right] \int_{\theta}^{\bar{\theta}} \left[\frac{1}{u'(c(z))} - \frac{W'(\mathrm{U}(z))}{\lambda}\right] \frac{dF(z)}{1-F(\theta)}$$

$$\frac{T'(\mathrm{y})}{1-T'(\mathrm{y})} = u'(c(\theta))\frac{1-\mathrm{F}(\theta)}{\theta f(\theta)}\left[1 + \frac{\mathrm{y}(\theta)}{\theta}\frac{v''\left(\frac{y(\theta)}{\theta}\right)}{v'\left(\frac{y(\theta)}{\theta}\right)}\right] \int_{\theta}^{\bar{\theta}} \left[\frac{1}{u'(c(z))} - \frac{W'(\mathrm{U}(z))}{\lambda}\right] \frac{dF(z)}{1-\mathrm{F}(\theta)} \tag{10}$$

Note that the LHS is increasing in τ. This is the famous **Diamond-Mirrlees-Saez formula**. This equation says that the optiaml marginal tax rates are determined by three things:

1. **The hazard rate** $\frac{1-F(\theta)}{f(\theta)}$ (f of the tail of the type distribution) In particular, for bounded distributions marginal taxes should be zero at the top. On the other hand if the distribution has fat tails, like the Pareto distribution then this term is positive.

2. **Labor supply elasticity** $\frac{v'\left(\frac{y(\theta)}{\theta}\right)}{\frac{y(\theta)}{\theta}v''\left(\frac{y(\theta)}{\theta}\right)}$: captures the effect of Frisch elasticity of labor supply (Captures the substitution effect of a marginal change in wage). The formula suggests that if labor is very elastic, then marginal tax rates should be low.

3. **Concern for redistribution**: If the planner loves redistribution then loosely the term $\int_{\theta}^{\bar{\theta}} \frac{W'(\mathrm{U}(z))}{\lambda}$ is small since planner cares more about the lower types. As a result marginal tax rates are higher.

7 Moral Hazard

For reference look at Laffont's book or Grossman Hart (1983) or Holmstrom (1979).

- There is a **Principal (P)** and **Agent (A)**
- Agent can exert effort $e \in \{e_H, e_L\}$
- Principal receives output $\pi \sim [\underline{\pi}, \bar{\pi}]$ with

$$F(\pi|e)\text{CDF} \quad f(\pi|e) > 0\text{PDF}$$

- Suppose $e_h >_{\text{FOSD}} e_L$: $F(\pi|e_H) \leq F(\pi|e_L)$ for all π with strict inequality on some subset of $[\underline{\pi}, \bar{\pi}]$ so

$$\mathbb{E}[\pi|e_H] > \mathbb{E}[\pi|e_L]$$

- Suppose A's preferences are given by

$$u(w, e) = v(w) - g(e)$$

  where w is a wage paid by P, $v' > 0, v'' \leq 0$. g is interpreted as effort cost: $g(e_H) > g(e_L)$

- A has reservation utility $\bar{U}$
- P has reservation utility 0 and is Risk Neutral

Timing

- P offers a take it or leave it contract to A.
- If A accepts then proceed with contract
- If A rejects then P and A get their outside options
- Suppose A accepting is worthwhile for P

7.1 Observed effort

P's problem

$$\max_{e,w}\{\mathbb{E}[\pi - w(\pi)|e] \quad \mathbb{E}[v(w(\pi))|e] \geq \bar{U}\}$$

P says to A : 'Exert effort e or I won't pay you'. WLOG solve this in 2 steps (bc we can solve separately for w and e):

- Step 2:

$$\max_{e} \mathbb{E}[\pi|e] - C(e)$$

- Step 1: Find $C(e)$:

$$C(e) = \min_{w}\{\mathbb{E}[w(\pi)|e] \quad \mathbb{E}[v(w(\pi))|e] - g(e) \geq \bar{U}\}$$

At the optimum IR constraint always binds , with assosiated multiplier γ FOCs

$$\frac{\partial}{\partial w(\pi)} : \quad -f(\pi|e) + \gamma v'(w(\pi))f(\pi|e) = 0$$

So

$$\frac{1}{v'(w(\pi))} = \gamma$$

If v' is strictly decreasing (i.e. if A is risk averse which is standard assumption) then w is constant! Fixed wage is optimal and equal:

$$\bar{w} = w(\pi) = (v')^{-1}(\frac{1}{\gamma})$$

Let's find it (we don't know γ). P chooses w_e^* s.t.:

$$v(w_e^*) - g(e) = \bar{U}$$

above comes from binding IR constraint in Step1 and as expected higher effort wage

$$w_{e_H}^* > w_{e_L}^*$$

If v' is constant then every w satisfies FOC:
'Everything' s.t. IR is optimal
Step 2:

$$\max_{e} \underbrace{\mathbb{E}[\pi|e]}_{\text{gross profit}} - \underbrace{(v')^{-1}(g(e) + \bar{U})}_{wagecost}$$

7.2 Unobservable effort

First suppose A is RN: $v(W) = w$

Let P 'sell' the firm to A:

$$w(\pi) = \pi - p$$

where p is price of the firm.

A chooses to maximize

$$\{\mathbb{E}[w(\pi)|e] - g(e) = \{\mathbb{E}[\pi|e] - g(e) - p$$

$v^{-1} = w$ this gives the same e^* as in observable effort problem.

Let p^* solve

$$\mathbb{E}[\pi|e] - g(e) - p^* = \bar{U}$$

Just like observable effort indeed

Now suppose A is risk averse. A faces following problem:

$$C(e) = \min_w \{ \quad \mathbb{E}[w(\pi)|e]$$

$$\text{st} \quad (\gamma) \quad \text{IR} \quad \mathbb{E}[v(w(\pi))|e] - g(e) \geq \bar{U}$$

$$(\mu) \quad \text{IC} \quad \mathbb{E}[v(w(\pi))|e] - g(e) \geq \mathbb{E}[v(w(\pi))|e'] - g(e') \quad \forall e' \quad \}$$

7.3 Limited Liability

Suppose $v = w$ but $w \geq 0$. Let's look at guy with high effort :

$$C(e_H) = \min_{w \geq 0} \{ \quad \mathbb{E}[w(\pi)|e_H]$$

$$\text{st} \quad (\gamma) \quad \text{IR} \quad \mathbb{E}[v(w(\pi))|e_H] - g(e_H) \geq \bar{U}$$

$$(\mu) \quad \text{IC} \quad \mathbb{E}[v(w(\pi))|e_H] - g(e_H) \geq \mathbb{E}[v(w(\pi))|e_L] - g(e_L) \quad \}$$

Individual Rationality (IR) is sometimes called Participation Constraint (PC).

FOCs:

$$\frac{\partial}{\partial w(\pi)} : \quad -f(\pi|e_H) + \gamma f(\pi|e_H) + \mu[f(\pi|e_H) - f(\pi|e_L)] \leq 0$$

$$1 \geq \gamma + \mu[1 - \frac{f(\pi|e_L)}{f(\pi|e_H)}] \tag{11}$$

Suppose

$$\mathbb{E}[\pi|e_H] - p^* = g(e_H) + \bar{U} \quad \Rightarrow \underline{\pi} - p^* < 0$$

$$c^*(e_H) = \max_{\gamma,\mu} \gamma(g(e_H) + \bar{U}) + \mu(g(e_H) - g(e_L))$$

$$\text{s.t.} \quad (11) \quad \text{holds}$$

Limited Liability (LL) does not only if $\mu = 0$

Observe that

$$\exists \pi : \quad \frac{f(\pi|e_L)}{f(\pi|e_H)} \lesseqgtr 1$$

but then $\gamma = 1 \Rightarrow$ from First Best:

$$w(\pi) = \pi - p^*$$

but this contradicts LL so $\mu > 0$

Lemma 13.

$$0 \geq \gamma \geq 1$$

Proof. Let's look at $1 \geq \gamma + \mu[1 - \frac{f(\pi|e_L)}{f(\pi|e_H)}]$ Let

$$LR^* = \max_{\pi} 1 - \frac{f(\pi|e_L)}{f(\pi|e_H)}$$

$$\gamma + \mu LR^* \leq 1$$

$g(e_H) - g(e_L) > 0$ inequality will bind and so

$$\mu = \frac{1-\gamma}{LR^*}$$

Objective:

$$\gamma(g(e_H) + \bar{U}) + \underbrace{(1-\gamma)(g(e_h) - g(e_L))}_{LR^*}$$

□

Back to A is RA. e_L is easy to implement

$$w^*_{e_L} = v^{-1}(\bar{U} + g(e_L))$$

is a constant wage.

To implement e_H we need IC:

$$C(e_H) = \min_w \{ \quad \mathbb{E}[w(\pi)|e_H]$$

$$\text{st} \quad (\gamma) \quad \text{IR} \quad \mathbb{E}[v(w(\pi))|e_H] - g(e_H) \geq \bar{U}$$

$$(\mu) \quad \text{IC} \quad \mathbb{E}[v(w(\pi))|e_H] - g(e_H) \geq \mathbb{E}[v(w(\pi))|e_L] - g(e_L) \quad \}$$

let's switch from wages to utils : $\varphi = v^{-1}$: $\varphi(\bar{v}(\pi)) = w(\pi)$ so we face following transformed problem:

$$C(e_H) = \min_{\bar{v}} \{ \quad \mathbb{E}[\varphi(\bar{v}(\pi))|e_H]$$

$$\text{st} \quad (\gamma) \quad \text{IR} \quad \mathbb{E}[\bar{v}(\pi)|e_H] - g(e_H) \geq \bar{U}$$

$$(\mu) \quad \text{IC} \quad \mathbb{E}[\bar{v}(\pi)|e_H] - g(e_H) \geq \mathbb{E}[\bar{v}(\pi)|e_L] - g(e_L) \quad \}$$

The purpose of it is to have well defined convex problem! Now we can take FOCs:

$$\frac{\partial}{\partial \bar{v}(\pi)} : \quad -\varphi'(\bar{v}(\pi))f(\pi|e_H) + \gamma f(\pi|e_H) + \mu[f(\pi|e_H) - f(\pi|e_L)] = 0$$

$$-\varphi'(\bar{v}(\pi)) = \gamma + \mu[1 - \frac{f(\pi|e_L)}{f(\pi|e_H)}]$$

Let's go back to wages :

$$\frac{1}{v'(w(\pi))} = \gamma + \mu[1 - \frac{f(\pi|e_L)}{f(\pi|e_H)}]$$

Lemma 14. *Both $\gamma > 0$ and $\mu > 0$*

Proof. Step1: $\gamma > 0$

Multiply FOC by $f(\pi|e_H)$

$$\frac{f(\pi|e_H)}{v'(w(\pi))} = \gamma f(\pi|e_H) + \mu f(\pi|e_H) - \mu f(\pi|e_L)$$

and integrate over *pi* to get:

$$\mathbb{E}\frac{1}{v'(w(\pi))} = \gamma + \mu - \mu = \gamma$$

$f(\pi|e) > 0, v' > 0$ so $\int f(\pi|e_H)\frac{1}{v'(w(\pi)} > 0$ which means $\gamma > 0$.

Step2: $\mu > 0$

Multiply FOC by $f(\pi|e_H)v(w(\pi))$

$$\frac{f(\pi|e_H)v(w(\pi))}{v'(w(\pi))} = \gamma v(w(\pi))f(\pi|e_H) + \mu f(\pi|e_H)v(w(\pi)) - \mu f(\pi|e_L)v(w(\pi))$$

and integrate over *pi* to get:

$$\mathbb{E}\frac{v(w(\pi))}{v'(w(\pi))} = \gamma\mathbb{E}v(w(\pi)) + \mu\int v(w(\pi)) \cdot (f(\pi|e_H) - f(\pi|e_L))$$

let's take last element on RHS . Keep in mind that IC is binding so $\lambda g(x) = 0$ so:

$$\mu \cdot \int (v(w(\pi))f(\pi|e_H) - C(e_H) - v(w(\pi))f(\pi|e_L) + C(e_L)) = 0$$

$$\mu \cdot \int (v(w(\pi))(f(\pi|e_H) - f(\pi|e_L)) = \mu(C(e_H) - C(e_L))$$

Plug it back to equation above

$$\mathbb{E}\frac{v(w(\pi))}{v'(w(\pi))} = \mathbb{E}\frac{1}{v'(w(\pi))}\mathbb{E}v(w(\pi)) + \mu(C(e_H) - C(e_L))$$

so

$$COV(v(w(\pi)), \frac{1}{v'(w(\pi))} = \mu(C(e_H) - C(e_L))$$

It is enough to show that if w then $cov(\cdot, \cdot) > 0$. To do so define:

$$f(w) = \frac{1}{v'(w)} \quad f'(w) = \frac{-v''(w)}{v'(w)^2} > 0$$

so if w then $u(w)$ and $f(w)$ so we have comovement and thus positive correlation. so indeed $\mu > 0$ □

Suppose $\hat{w}$ solves

$$\frac{1}{v'(\hat{w})} = \gamma$$

$$w(\pi) > \bar{w} \quad \frac{f(\pi|e_L)}{f(\pi|e_H)} < 1$$

$$w(\pi) < \bar{w} \quad \frac{f(\pi|e_L)}{f(\pi|e_H)} > 1$$

This rationalize using **Monotone likelihood ratio property (MLRP)**:

Definition 58 (MLRP). *f satisfies MLRP if $\frac{f(\pi|e_L)}{f(\pi|e_H)}$ is decreasing in π*

MLRP is not implied by FOSD

$\gamma > 0$ and $\mu > 0$ and IR binds so

$$\mathbb{E}[v(w(\pi))|e_H] = g(e_H) + \bar{U}$$

from Jensen inequality:

$$v(\mathbb{E}(w(\pi))|e_H] > g(e_H) + \bar{U}$$

7.4 More actions

With 2 actions $C(e)$ was well defined as lond as $e_H >_{\text{FOSD}} e_L$. With more actions: $\{e_L, e_M, e_H\} = \{(1,0), (\frac{1}{2}, \frac{1}{2}), (0,1)\}$ wwe say e_M is not implementable if

$$g(e_M) > \frac{1}{2}(g(e_L) + g)e_H))$$

Let's again look at transformed problem (we assume discrete space on x_i instaed of π):

$$C(e) = \min_{\bar{v}}\{ \quad \sum_i \varphi(\bar{v}(x_i)) f(x_i|e)$$

$$\text{st} \quad (\gamma) \quad \text{IR} \quad \sum_i f(x_i|e)\bar{v}(x_i) - g(e) \geq \bar{U}$$

$$(\mu) \quad \text{IC} \quad \sum_i f(x_i|e)\bar{v}(x_i) - g(e) \geq \sum_i f(x_i|e')[\bar{v}(\pi) - g(e') \qquad e'\}$$

FOCs:

$$\gamma + \sum_e \mu(e)[1 - \frac{f(x_i|e')}{f(x_i|e)}] = \varphi'(v_i)$$

FOA

$$\max_{w,a} \int_{\underline{x}}^{\bar{x}} (x - w(x)) f(x|a) dx$$

$$\text{st} \quad \int_{\underline{x}}^{\bar{x}} v(w(x)) f(x|a) dx - g(a) \geq \bar{U}$$

$$a \in \arg\max_{a'} \int_{\underline{x}}^{\bar{x}} v(w(x)) f(x|a') dx - g(a')$$

Assume. $g' > 0, g'' > 0$

FOCs:

$$\int_{\underline{x}}^{\bar{x}} v(w(x)) f_a(x|a) dx - g'(a) = 0$$

SOCs

$$\int_{\underline{x}}^{\bar{x}} v(w(x)) f_{a,a}(x|a) dx - g''(a) \leq 0$$

7.5 Relaxed problem

Same objective s.t. IR (γ) and FOC μ

Theorem 25. ***Suppose MLRP holds. Then FOA*** $\Rightarrow \mu > 0$

Proof.

$$\frac{\partial L}{\partial a}: \quad \int_{\underline{x}}^{\bar{x}} (x - w(x)) f_a(x|a) +$$

$$+\gamma[\int_{\underline{x}}^{\bar{x}} v(w(x)) f_a(x|a) dx - g'(a)] +$$

$$+\mu[\int_{\underline{x}}^{\bar{x}} v(w(x)) f_{a,a}(x|a) dx - g''(a)] = 0$$

Second element is 0 from FOC.

IF $\mu \leq 0$ then SOC $\Rightarrow$

$$\int_{\underline{x}}^{\bar{x}} (x - w(x)) f_a(x|a) \leq 0$$

Let w_{gamma} solve

$$\frac{1}{v'(w_\gamma))} = \gamma$$

$w(x)$ solves

$$\frac{1}{v'(w_(x))} = \gamma + \mu \frac{f_a(x|a)}{f_{aa}(x|a)} \quad x$$

$$\frac{\partial L}{\partial w(x)}: \quad -f(x|a) + \gamma f(x|a) v'(w(x)) + \mu f_a(x|a) v'(w(x)) = 0$$

$$\mu \leq 0 \quad \Rightarrow \quad [w(x) \leq w_\gamma \quad \iff \quad f_a(x|a) \geq 0]$$

Therefore

$$(x - w(x))f_a(x|a) \geq (x - w_\gamma)f_a(x|a)$$

$$\int_{\underline{x}}^{\bar{x}} (x - w(x))f_a(x|a)dx \geq \int_{\underline{x}}^{\bar{x}} (x - w_\gamma)f_a(x|a)dx$$

Integrating RHS by parts

$$(x - w(x))F_a(x|a)|_{\underline{x}}^{\bar{x}} - \int_{\underline{x}}^{\bar{x}} F_a(x|a)dx = 0 - \int_{\underline{x}}^{\bar{x}} F_a(x|a)dx > 0$$

it comes from fact that $F_a(\underline{x}|a) = F_a(\bar{x}|a) = 0$ (Suppose constant support) and FOSD
But this contradicts

$$\int_{\underline{x}}^{\bar{x}} (x - w(x))f_a(x|a) \leq 0$$

□

$v'' < 0$ so w is differentiable and if MLRP w is increasing by

$$\frac{1}{v'(W(x))} = \gamma + \mu \frac{f_a(x|a)}{f(x|a)}$$

Theorem 26. ***FOA is valid if F is convex in effort and MLRP***

Agent's payoff:

$$\int_{\underline{x}}^{\bar{x}} v(w(x))f(x|a)dx - g(a) = v(w(x))F(x|a)|_{\underline{x}}^{\bar{x}} - \int_{\underline{x}}^{\bar{x}} v'(w(x))w'(x)F(x|a)dx - g(a) =$$

$$= v(w(\bar{x})) - \int_{\underline{x}}^{\bar{x}} v'(w(x))w'(x)F(x|a)dx - g(a)$$

Assume that w is differentiable and differentiate twice wrt a:

$$-\int_{\underline{x}}^{\bar{x}} v'(w(x))w'(x) \underbrace{F_{aa}(x|a)}_{\geq 0} dx - \underbrace{g''(a)}_{>0} < 0 \quad \forall a$$

$\Rightarrow$ SOC is satisfied. However we cannot assume that the previous theorem to argue that $/mu > 0$ under MLRP, so that w is increasing and differentiable. What now?

7.6 Doubly Relaxed Problem

by Rogerson 1988. FOC≥ 0 instead of $= 0$. So:

$$\frac{d}{da}\mathbb{E}[U(w,a)] \geq 0$$

with $\delta \geq 0$ as new multiplier. MLRP $\Rightarrow$ increasing, differentiable solution to doubly relaxed problem.

WTS: The DRP is valid for the relaxed problem

Lemma 15. *at DRP*

$$\frac{d}{da}\mathbb{E}[U(w,a)] = 0$$

Proof. If $\delta > 0$ then we are done.

Suppose $\delta = 0$ so $\gamma > 0$

$$\frac{1}{v'(w_(x))} = \gamma + \delta\frac{f_a(x|a)}{f(x|a)} \quad x$$

Hence $w(x) = w_\gamma$ for all x. Integrate by parts

$$\mathbb{E}[x - w_\gamma|a] = (x - w_\gamma)F(x|a)|_{\underline{x}}^{\bar{x}} - \int_{\underline{x}}^{\bar{x}} {}_{\underline{x}}^{\bar{x}}F(x|a)dx =$$

$$= \bar{x} - w_\gamma - \int_{\underline{x}}^{\bar{x}} f(x|a)dx$$

differentiate wrt a:

$$\frac{\partial}{\partial a}: \quad -{}_{\underline{x}}^{\bar{x}}F_a(x|a)dx \geq 0$$

Since $F_a \leq 0$ by FOSD.

$\gamma > 0$ then FOC of the DRP yields:

$$\underbrace{\frac{d}{da}\mathbb{E}[x - w(x)|a]}_{\geq} + \underbrace{\gamma\frac{d}{da}\mathbb{E}[U(w,a)]}_{\Rightarrow\leq 0} = 0$$

But $\frac{d}{da}\mathbb{E}[U(w,a)] \geq 0$ therefore $\frac{d}{da}\mathbb{E}[U(w,a)] = 0$

□

Example 12 (Mirrlees 1978)**.** *Let a, z be scalars. Solve*

$$\max_a -(a-2)^2 - (z-1)^2$$

$$st \quad .z \in \arg\max ae^{-(z+1)^2} + e^{-(z-1)^2}$$

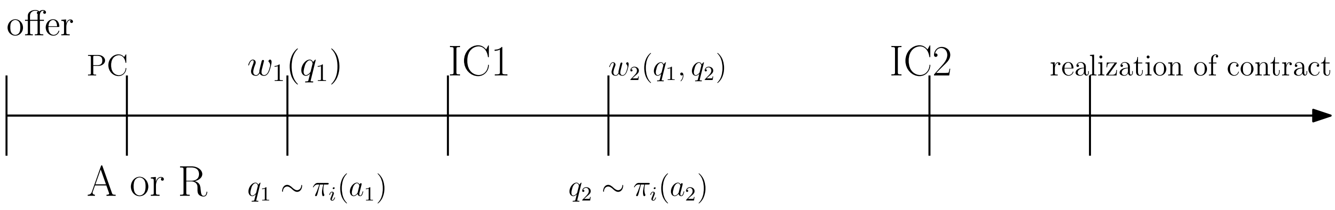

Figure 5: Timing of dynamic contract

7.7 Dynamic problem

Principal's preferences

$$S(q_1) - w_1 - \delta_p(S(q_2) - w_2)$$

Agent's preferences

$$u(w_1) - C(a_1) + \delta_a(u(w_2) - C(a_2))$$

For $q_1, q_2 \in \{q_L, q_H\}$ and $a_1, a_2 \in \{a_L, a_H\}$. Notation for $w_1(q_H)$, $w_1(q_L)$, $w_2(q_H, q_H)$, $w_2(q_H, q_L)$, $w_2(q_L, q_H)$, $w_2(q_L, q_L)$:
$\bar{u} := u(w_1(q_H))$, $\underline{u} := u(w_1(q_L))$, $u(\bar{q}_L) := u(w_2(q_L, q_H))$,
$u(\bar{q}_H) := u(w_2(q_H, q_H))$, $\underline{u(q_H)} := u(w_2(q_H, q_L))$,
$\underline{u(q_L)} := u(w_2(q_L, q_L))$.
$C(a_H) = \Psi$, $C(a_L) = 0$, $S_1 = S(q_H)$, $S_0 = S(q_L)$.
$\pi_1 = \pi_H(a_H)$, $\pi_0 = \pi_H(a_L)$, $1 - \pi_1 = \pi_L(a_H)$, $1 - \pi_0 = \pi_L(a_L)$.
Now let's solve two action, two levels productivity moral hazard:

$$\text{IC} \quad \sum_i \pi_i(a_H)u(w_i) - C(a_H) \geq \sum_i \pi_i(a_L)u(w_i) - C(a_L)$$

$$\pi_1\bar{u} + (1 - \pi_1)\underline{u} - \Psi \geq \pi_0\bar{u} + (1 - \pi_0)\underline{u} - 0$$

$$(\pi_1 - \pi_0)(\bar{u} - u) \geq \Psi \qquad \Delta\pi := \pi_1 - \pi_0 \geq 0$$

$$\Delta\pi(\bar{u} - u) \geq \Psi$$

$$\bar{u} - u \geq \frac{\Psi}{\Delta\pi}$$

in $t = 2$ $\forall q_1 \in \{q_H, q_L\}$

$$\text{(IC2)} \quad u(\bar{q}_1) - u(q_1) \geq \frac{\Psi}{\Delta\pi}$$

in $t = 1$ $\forall q_2 \in \{q_H, q_L\}$

$$\text{(IC1)} \quad \bar{u} + \delta_A \mathbb{E}_\pi \bar{u}(q_1) - u + \delta_A \mathbb{E}_\pi \underline{u}(q_1)$$

$$\text{(IC1)} \quad \bar{u} + \delta_A(\pi_1 \bar{u}(q_H) + (1-\pi_1)\bar{u}(q_L)) - \underline{u} + \delta_A(\pi_1 u(q_H) + (1-\pi_1)u(q_L))$$

PC $\quad \pi_1[\bar{u} + \delta_A(\pi_1 \bar{u}(q_H) + (1-\pi_1)\bar{u}(q_L)] + (1-\pi_1)[\underline{u} + \delta_A(\pi_1 \underline{u}(q_H) + (1-\pi_1)\underline{u}(q_L)] \geq \Psi + \delta_A \Psi$

Expected utility of principal $Ep := \mathbb{E}(S(q_1) - w_1 + \delta_P(S(q_2) - w_2)$

$$\pi_1(S_1 - w_H) + (1-\pi_1)(S_0 - w_L) + \delta_P(\pi_1 S_1 + (1-\pi_1)S_0 - (\pi_1 \pi_1 h(\bar{w_H}) + (1-\pi_1)\pi_1 h(\underline{w_H}) +$$

$$(1-\pi_1)\pi_1 h(\bar{w_L}) + (1-\pi_1)(1-\pi_1)h(\underline{w_L})$$

we want to solve max Ep s.t. (IC1), (IC2) and (PC) and we will do so by implementing recursive method and taking $u(q_1)$ as state with promised utility constraint

$$\pi_1 \bar{u}_(q_1) + (1-\pi_1)\underline{u}(q1) \geq u_2(q1)$$

for t=2 solve $V_2(u(q_1))$

$$V_2(u(q_1)) = \max\{\pi_1(S_1 - h(u(\bar{q}_1))) + (1-\pi_1)(S_0 - h(\underline{u}(q_1)))\}$$

s.t. (IC2) and (PC) holds. now we have 2 variables and 2 constraints, define C_2^{SB}

$$C_2^{SB} = \max\{\pi_1 h(\bar{u}_(q_1)) + (1-\pi_1)h(\underline{u}(q1))\}$$

$$V_2(u(q_1)) = \max\{\pi_1 S_1 + (1-\pi_1)S_0 c - C_2^{SB}(u(q_1))$$

let's take FOCs

$$\frac{\partial}{\partial u(q_1)} V_2(u(q_1)) = -(C_2^{SB})' = -\pi_1 h'(\bar{u}_(q_1)) - (1-\pi_1)h'(\underline{u}(q1))$$

To solve for $t = 1$ we solve max Ep with (IC1) and (PC).. Because I won't rewrite Langrange function (with λ and μ adjoint multipliers for IC1 and PC respectively) I just give FOCs for time 1 problem

$$\frac{\partial}{\partial \bar{u}} L = -\pi_1 h'(\bar{u}) + \lambda \pi_1 + \mu = 0$$

$$\frac{\partial}{\partial \underline{u}} L = -(1-\pi_1)h'(\underline{u}) + \lambda(1-\pi_1) - \mu = 0$$

$$\frac{\partial}{\partial u(q_H)} L = -\delta_P \pi_1 V_2'(u(q_H)) + \lambda \delta_A \pi_1 + \delta \mu = 0$$

$$\frac{\partial}{\partial u(q_L)} L = -\delta_P(1-\pi_1)V_2'(u(q_L)) + \lambda\delta_A(1-\pi_1) - \delta\mu = 0$$

Those 4 equations with 2 additional eqautions for period 2 can be solved (6 variables). Now let's add up first and third and we get

$$-\delta_P V_2'(u(q_H)) = \delta_A h'(\bar{u})$$

equation 2 and 4 when summed are equal to

$$\delta_P V_2'(u(q_L)) = \delta_A h'(\underline{u})$$

Recall that $u^{-1} = h$ and $h' = (u^{-1})'$ we obtain Euler equation (when we take weighted sum of result above)

$$\frac{\delta_A}{u'} = \mathbb{E}\frac{\delta_P}{u'(q_1)} \iff 1 = \mathbb{E}\frac{\delta_P u'}{\delta_A u'(q_1)}$$

h is convsvr and by analogue of analysis from macro (higher the $\beta = \frac{\delta_A}{\delta_B}$ the more agents prefer to get incentive in second period.

8 Informational Frictions in markets

8.1 Akerlof's market for lemons

- Buyer's valuation:
$$v = \begin{cases} 1 & \text{if peach} \\ 0 & \text{if lemon} \end{cases}$$

- Seller: π fraction are peach, and $(1-\pi)$ are lemon where $0 < \pi < 1$.

- The opportunity cost for seller:
$$c = \begin{cases} \frac{1}{2} & \text{if peach} \\ 0 & \text{if lemon} \end{cases}$$

- The market price is p

- So the buyer's expected payoff is $E(v \mid \text{sale}) = \begin{cases} \pi & \text{if } p \geq \frac{1}{2} \\ 0 & \text{if } p < \frac{1}{2} \end{cases}$.

- If $\pi \geq \frac{1}{2}$, then there will be some trade only because there exist so many peaches;

- If $\pi < \frac{1}{2}$, then there is no trade, and market breaks down.

8.2 Spence's signaling

A worker chooses education level $e \geq 0$ with private cost e/θ where θ is private type and the same as productivity.

- Competitive firm set wage at $w(e) = E(\theta \mid e)$.

- Two types of workers: θ', θ'' s.t. $0 < \theta' < \theta''$ with probability p' and $p'' = 1 - p'$.

- Let σ' and σ'' be some strategies for θ' and θ''

Lemma 16. *If* $\Pr(e' \mid \sigma') > 0$ *and* $\Pr(e'' \mid \sigma'') > 0$, *then* $e' \leq e''$

Proof.

$$\left.\begin{array}{r} w(e') - e'/\theta' \geq w(e'') - e''/\theta' \\ w(e'') - e''/\theta'' \geq w(e') - e'/\theta'' \end{array}\right\} \Rightarrow e''(1/\theta' - 1/\theta'') \geq e'(1/\theta' - 1/\theta'').$$

□

The separating equilibria: so wage is set at θ:

- Type θ' reveal his type and receive wage θ' and choose $e' = 0$;
- Type θ'' chooses e'' and receive wage θ'', for $(e' = 0, e'')$ being a separating equilibrium: $-\theta'' - e''/\theta'' \geq \theta' \Rightarrow e'' \leq \theta'' (\theta'' - \theta')$ and;

$$\theta' \geq \theta'' - e''/\theta' \Rightarrow e'' \geq \theta' \left(\theta'' - \theta'\right)$$

- Conversely, suppose $e'' \in [\theta' (\theta'' - \theta'), \theta'' (\theta'' - \theta')]$, consider the belief $\Pr(\theta' \mid e) = \begin{cases} 1 & \text{if } e \neq e'' \\ 0 & \text{if } e = e'' \end{cases}$.
- So we have a continuum of separating equilibria.
- The pooling equilibrium: let $\tilde{e} = e' = e''$ and wage is set at $w(\tilde{e}) = p'\theta' + p''\theta''$.
- The belief to support it is $\Pr(\theta' \mid e) = 1$ if $e \neq \tilde{e}$
- So, for $\tilde{e}$ to be a pooling equilibrium:

$$-\theta' \leq p'\theta' + p''\theta'' - \tilde{e}/\theta' \Rightarrow \tilde{e} \leq p''\theta' \left(\theta'' - \theta'\right)$$

- Note that $\theta' < \theta'' \Rightarrow \theta'' < p'\theta' + p''\theta'' < p'\theta' + p''\theta'' - \tilde{e}/\theta''$

8.3 Beer-quiche game

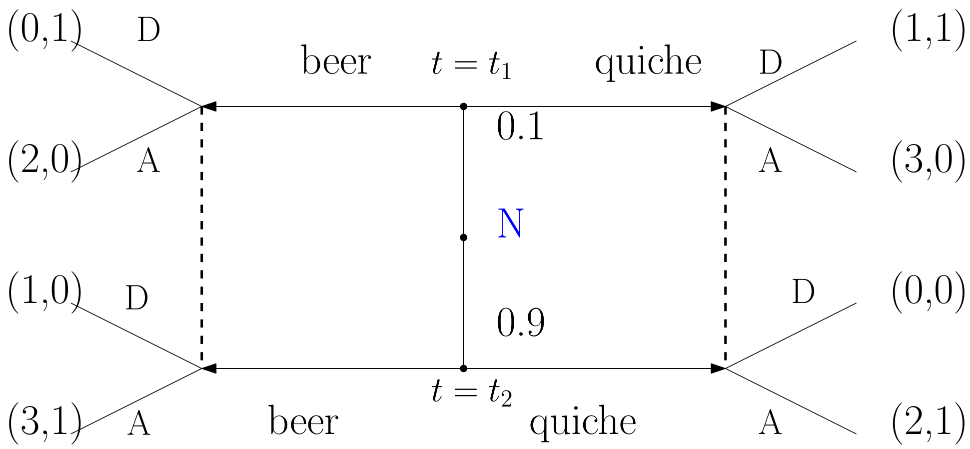

Figure 6: Beer Quiche Game- Cho, Kreps 1984

- There are two pooling equilibria: .9 type s and .1 type w
- Player 1 chooses beer, and Player 2 chooses "Avoid" if beer and chooses "Duel" if quiche, and player 2 believes $\Pr(t_w \mid \text{ quiche }) = 1$
- Same but beer and quiche are reversed.
- The separating equilibrium: type w chooses beer, and type s chooses quiche;
- Hybrid equilibrium for .1 type s and .9 type $-w$: $1/9$ type w chooses beer and all the type s chooses beer, $8/9$ type w choose quiche,
- player 2 has the correct belief and choose $\left(\frac{1}{2}, \frac{1}{2}\right)$ if beer and dual if quiche.
- Hybrid equilibrium
- P1 plays beer if A , beer with probability $\frac{1}{9}$ if W and quiche with probability $\frac{8}{9}$ if W
- P2 plays D if quiche, D with probability $\frac{1}{2}$ if beer and A with probability $\frac{1}{2}$ if beer
- P2 beer 0.5*1+0.5*0=0.5*0+0.5*1 and quiche D is clearly optimal.
- Beliefs $P(w|b) = \frac{0.9*\frac{1}{9}}{0.9*\frac{1}{9}+\frac{1}{10}*1} = \frac{1}{2}$ and $P(w|q) = 1$
- w: beer vs quiche $0.5*0+0.5*2=1$. A: $0.5*3+0.5*1>0$

8.4 Rothschild's and Stiglitz' insurance markets adverse selection

8.5 Grossman's Stiglitz' informational efficiency

8.6 Kyle's information aggregation

One insider i with valuatio $V \sim N(\mu, \sigma^2)$

signal

$$s_i = V + \epsilon_i \quad \epsilon_i \sim N(0, \sigma_i^2)$$

$x_i \in \mathbb{R}$ is a demand

Noise traders $x_0 \sim N(0, \sigma_0^2)$

$(V, x_0, \epsilon_i) \sim N$ independent

Market maker observes

$$X = x_0 + x_i$$

$$P(x) = \mathbf{E}[V|x] \qquad \text{'zero profit'}$$

$$\Pi_i(s_i) = \max_{x_i} \mathbb{E}[(V - P(x))x_i|s_i]$$

for convenience $\mathbb{E}[\cdot|s_i] = \mathbb{E}_i[\cdot]$ so we face following problem

$$\Pi_i(s_i) = \max_{x_i} \mathbb{E}_i[(V - P(x_0 + x_i))x_i]$$

Definition 59. *And equilibrium is a profile* (x_i, P) *s.t.*

$$x_i(s_i) \in \arg\max_{x_i} \mathbb{E}_i[(V - P(x_0 + x_i))x_i]$$

$$P(x) = \mathbf{E}[V|x]$$

Let's focus on linear equlibrium. Expectation and $P(x)$, and x_i are linear functions. Let

$$P(x) = a + bx$$

$$\Pi_i(s_i) = \mathbb{E}_i[(V - a - b(x_0 + x_i))x_i]$$

Focs

$$\mathbb{E}_i[(V - a - b(x_0 + x_i)] = \mathbb{E}_i[bx_i]$$

$$x_i = \frac{\mathbb{E}_i[V] - a}{2b}$$

Let's use projection theorem:

$$\mathbb{E}_i[V] = \mu + \frac{\sigma^2}{\sigma^2 + \sigma_i^2}(s_i - mu)$$

so

$$x_i(s_i) = \frac{\lambda_i(s_i - \mu) + \mu - a}{2b}$$

$$\mathbb{E}[V|X] = \mu + \frac{\lambda_i}{2b}\frac{\sigma^2}{\sigma^2 + \sigma_i^2}(s_i - mu)$$

Let

$$b = \frac{\lambda_i \lambda_0}{2b} \quad \lambda_0 = \frac{\sigma^2}{\sigma^2 + \sigma_i^2}$$

so

$$b = \sqrt{\frac{1}{2}\lambda_i \lambda_0}$$

$$a = \mu + \frac{\lambda_i \lambda_0}{2b}(-\frac{\mu + a}{2b}) \quad \Rightarrow \quad a = \mu$$

8.7 Leland's and Pyle's CAPM

- We are deciding about investing in project with initial capital K
- Future income is $\mu + \bar{x}$ where $\bar{x} \sim (0, \sigma^2)$
- Entrepreneur (E) retains α if equity
- Firm and E can both borrow at the riskless rate r
- Modigliani, Miller 1958, De Marzo, Duffie (ECTA)
- E knows mu it is her private info.
- Market value (price) of the project:

$$V(\alpha) = \frac{1}{1+r}[\mu(\alpha) - \lambda]$$

- where $\mu(\alpha)$ is market valuation schedule
- λ is market adjustment for the risk in $\bar{x}$

CAPM

- $\bar{m}$ income from market portfolio
- V_m value of market portfolio

$$\mathbb{E}[R_x] - R_f = \frac{\text{COV}(R_x, R_\mu)}{\text{VAR}(R_\mu)} \cdot (\mathbb{E}[R_\mu] - R_f)$$

- Huang and Litzenberger

$$\mathbb{E}[R_x] - R_f =$$

$$= \mathbb{E}[\bar{x} + \mu - \underbrace{(1+r)D - (V(\alpha) - D)}_{\text{firm's debt}} - [(1+r)(V(\alpha) - D) - (V(\alpha) - D)]] \cdot \frac{1}{V(\alpha) - D} =$$

$$= \frac{\lambda}{V(\alpha) - D}$$

- since $\mathbb{E}\bar{x} = 0$
- $R_m = \frac{\bar{m} - V_m}{\bar{m}_m}$ so

$$\mathbb{E}[R_m] - R_f = \frac{\mathbb{E}[\bar{m}] - V_m - [(1+r)V_m - V_m]}{V_m} = \frac{\mathbb{E}[\bar{m}] - (1+r)V_m}{V_m}$$

$$\frac{cov(R_x, R_m)}{Var(R_m)} = \frac{cov(\bar{x}, \bar{m})}{Var(\bar{m})} \cdot \frac{V_m^2}{(V(\alpha) - D)V_m}$$

$$\mathbb{E}[R_x] - R_f = \frac{\lambda}{V(\alpha) - D} = \frac{cov(R_x, R_m)}{Var(R_m)}(\mathbb{E}[R_\mu] - R_f) =$$

$$= \frac{cov(\bar{x}, \bar{m})}{Var(\bar{m})} \cdot \frac{V_m}{V(\alpha) - D} \frac{\mathbb{E}[\bar{m}] - (1+r)V_m}{V_m}$$

$$\lambda = \frac{cov(\bar{x}, \bar{m})}{Var(\bar{m})} \cdot (\mathbb{E}[\bar{m}] - (1+r)V_m)$$

Suppose $\mu(\alpha)$ is differentiable.

'Perfect competition': Project is small relative to market.

E maximizes EU by choosing

- financial structure of firm
- retained equity α
- individual holdings of market portfolio and riskless asset
- Budget constraint

$$W_0 + D + (1-\alpha)[V(\alpha) - D] - K = \beta V_M + Y$$

- W_0 individual wealth, y individual holdings of debt, *beta* fraction of market portfolio purchased by E
- returns to equity $\mu + \bar{x} - (1+r)D$
- End-of-period wealth

$$W_1 = \alpha[\mu + \bar{x} - (1+r)D] + \beta m + (1+r)Y = \alpha[\mu + \bar{x} - \mu(\alpha) + \lambda] +$$

$$+\beta[m + (1+r)V_m] + (W_0 - K)(1+r) + \mu(\alpha) - \lambda$$

- budget constraint determines $\alpha D - Y$
- $\mu(\alpha)$ is equilibrium schedule if

$$\mu(\alpha^*(\mu)) = \mu \quad \mu$$

- that induce E to undertake the project, given $\mu(\alpha)$ where α^* solves E's problem. to justify this condition take

$$\mu(\alpha^*(\mu)) > \mu$$

- outside investors make less than the return they receive for project's risk
- Suppose mean variance preferences (b CRRA coefficient)

$$\mathbb{E}[U(W_1)] = \mathbb{E}[W_1] - \frac{1}{2}bVAR(W_1)$$

- FOCs

$$\frac{d}{d\alpha}: \quad \mu - \mu(\alpha) + \lambda + (1-\alpha)\mu'(\alpha) - \alpha bVar(\bar{x}) - \beta bcov(\bar{x}, \bar{m}) = 0$$

$$\frac{d}{d\beta}: \quad E\bar{m} - (1+r)V_m - \beta bVar(\bar{m}) - \alpha bcov(\bar{x}, \bar{m}) = 0$$

$$\beta b = \frac{E\bar{m} - (1+r)V_m - \alpha bcov(\bar{x}, \bar{m})}{Var(\bar{m})} = \lambda^* - \frac{\alpha bcov(\bar{x}, \bar{m})}{Var(\bar{m})}$$

$$0 = \mu - \mu(\alpha) + \lambda + (1-\alpha)\mu'(\alpha) - \alpha bVar(\bar{x}) - \beta bcov(\bar{x}, \bar{m}) =$$

$$= [\lambda^* - \frac{\alpha bcov(\bar{x}, \bar{m})}{Var(\bar{m})}]cov(\bar{x}, \bar{m})$$

$$(1-\alpha)\mu'(\alpha) = \alpha bZ$$

$$Z = \frac{Var(\bar{x}Var(\bar{m}) - cov(\bar{x}, \bar{m})^2}{Var(\bar{m})} \geq 0$$

- Z is 'specific' risk

$$\mu'(\alpha) = \frac{\alpha}{1-\alpha}bZ$$

$$\mu(\alpha) = -bZ(\ln(1-\alpha) + \alpha) + (1+r)K + \lambda$$

- Suppose first the equilibrium μ is μ_Yst.

$$\mu_Y(0) > (1+r)K + \lambda$$

- E's arbitrary low mu would benefit from starting the project and returning zero equity. Investors would not get a return
- Suppose μ_L is equilibrium schedule μ_K beats μ_L in competition for E's. So there is no equilibrium

$$\mu(\alpha) = -bZ(\ln(1-\alpha) + \alpha) + (1+r)K + \lambda$$

$$V(\alpha) = \frac{1}{1+r}[-bZ(\ln(1-\alpha)+\alpha)] + K$$
$$V'(\alpha) = \frac{bZ}{1+r}\frac{\alpha}{1-\alpha} > 0$$

- $V(0) = K$, $V(\alpha) > K$

Lemma 17. *A project is undertaken* $\iff$ *its true market value given* μ *exceeds its cost*

9 Bargaining

9.1 Nash solution

Definition 60. *Two person bargaining* (F, v)

$$F \subset \mathbb{R} : F \cap \{(x_1, x_2) : x_1 \geq v_1, x_2 \geq v_2\} \neq \emptyset, \text{bounded, convex, closed}$$

- *F- set of feasible payoff allocations*
- *v- disagreement point*
- (F, v) *is essential* $\iff \exists y \in F : y_1 > v_1, y_2 > v_2$

$\varphi(F, v)$ solution of Nash Bargaining problem satisfy following axioms

Definition 61. *Strong Pareto Efficiency*

$$x \in F if \neg \exists y \in F \quad y \geq x \wedge y_i > x_i \text{for some } i$$

Definition 62. *Weakly Pareto Efficient*

$$z \in F \quad if \quad \neg \exists y \in F \quad y \geq z$$

Definition 63. *Strong Efficiency*

$$\varphi(F, v) \in F x \in F \quad x \geq \varphi(F, v) \quad \Rightarrow \quad x = \varphi(F, v)$$

Definition 64. *Individual Rationality*

$$\varphi(F, v) \geq v$$

Definition 65. *Scale Covariance*

$$\forall \lambda_1, \lambda_2 > 0 \mu_1, \mu_2 \in R \quad w = (\lambda_1 x_1 + v_1, \lambda_2 x_2 + v_2) \quad G = \{(\lambda_1 x_1 + \mu_1, \lambda_2 x_2 + \mu_2) x \in F\}$$

then

$$\varphi(G, w) = (\lambda_1 \varphi_1(F, v) + \mu_1, \lambda_2 \varphi_2(F, v) + \mu_2)$$

Definition 66. *Independent of Irrelevant Alternatives (IIA) For any closed convex* $G \subset F$

$$\text{if } G \subset F, \varphi(F, v) \in G \quad \Rightarrow \quad \varphi(F, v) = \varphi(G, v)$$

Definition 67. *Symmetry*

$$\text{if } v_1 = v_2 \quad F \text{ symmetric} \Rightarrow \varphi_1(F, v) = \varphi_2(F, v)$$

Theorem 27. ***Nash Let** (F, v) **be 2 person bargaining problem** $\varphi(F, v) \in F$ **it's unique solution satisfying SE, IR,SC, IIA and S** $\iff$*

$$\varphi(F, v) \in \arg \max_{y \in F, y \geq v} (y_1 - v_1) \cdot (y_2 - v_2)$$

Proof. TBD □

Example 13. $\Gamma = \{(1, 2), c_1, c_2, u_1, u_2\}$

$$F = \{(u_1(\mu), u_2(\mu))\mu\Delta(C)\} \quad u_i(\mu) = \sum_{c \in C} \mu(c)u_i(c)$$

if there is a moral hazard-so no regulation by contracts is possible then

$$F = \{(u_1(\mu), u_2(\mu)) \quad \mu \quad \text{is coreelated equilibrium of } \Gamma$$

How to pick v?

a) min max $v_1 = \min_{\sigma_2 \in \Delta(C_2)} \max_{\sigma_1 \in \Delta(C_1)} u_1(\sigma_1, \sigma_2)$

b) $(\sigma_1, \sigma_2$ focal equilibrium then $v_i = u_i(\sigma_1, \sigma_2$

c) rational threats

Solution of two bargaining problems

9.2 Interpersonal Comparision of Utilities

Consider two principles

- equal gain - egalitarian solution E^*
- greatest good - utilitarian solution U^*

E^* and U^* need not generally agree.

Let E^* select from (F, v) the unique point that is weakly efficient in F and

$$x_1 - v_1 = x_2 - v_2$$

U^* select from (F, v) x s.t.

$$x_1 + x_2 = \max_{y \in F} y_1 + y_2$$

z is λ-utilitarian solution if

$$\lambda_1(x_1 - v_1) = \lambda_2(x_2 - v_2)$$

Theorem 28. ***Suppose*** (F, v) ***is essential and let*** $x \in F$

$$x = \varphi(F, v) \iff \exists \lambda > 0 \quad \lambda_1(x_1 - v_1) = \lambda_2(x_2 - v_2) \quad \textbf{\textit{and}} \quad \lambda_1 x_1 + \lambda_2 x_2 = \max_{y \in F} \lambda_1 y_1 + \lambda_2 y_2$$

Example 14. $v = (0, 0)$ *Split \$30 and P1 is Risk Neutral and P2 is Risk Averse.*

$$F = \{(y_1, y_2) : 0 \leq y_1 \leq 30, 0 \leq y_2 \leq (30 - y_1)^{\frac{1}{2}}\}$$

Nash solution

$$\frac{d}{dy_1}[y_1(30 - y_1)^{\frac{1}{2}}] = (30 - y_1)^{\frac{1}{2}} - \frac{y_1}{2(30 - y_1)^{\frac{1}{2}}} = 0$$

$$30 - y_1 = \frac{y_1}{2} \quad \Rightarrow y_1 = 20$$

$$(20, 10^{\frac{1}{2}}) = (20, 3.162)$$

$$-\frac{x_2}{x_1} \frac{\sqrt{10}}{20} = \frac{\lambda_1}{\lambda_2}$$

$$20\lambda_1 = \sqrt{10}\lambda_2 \quad \lambda_1 = 1 \lambda_2 = \sqrt{40}$$

Consider the scaling factors λ_1, λ_2 *as above. P2's utility from a monetary gain of \$K is* $\sqrt{40} K^{\frac{1}{2}}$ *instead of* $K^{\frac{1}{2}}$*. P1 remains unchanged*

$$G = \{(y_1, y_2) : 0 \leq y_1 \leq 30, 0 \leq y_2 \leq 6.325(30 - y_1)^{\frac{1}{2}}\}$$

For $(G, (0, 0))$ *Nash is* $(20, 20)$ *which corresponds to* \$20, \$10 *both utilitarian and egalitarian*

9.3 Rubinstein (1982)

1. 2 players alternating makes decision, start from P1
2. P1 makes an offer (x_1, x_2) , P2 can choose to accept or reject:
   - if accept, game end;
   - if reject: with prob p_1 game end with disagreement and P2 gets v_2, P1 gets $w_1 \leq \max_{y \in F_v} y_1$; with prob $1 - p_1$, game continue.
3. If game continue, P2 makes an offer (y_1, y_2) , P1 A or R :
   - if A, game end;
   - if R, with p_2 game end with disagreement and P1 gets v_1, P2 gets $w_2 \leq \max_{y \in F_v} y_2$; with $1 - p_2$, game continue
4. Repeat 2 and 3

Acknowledgment

To put together this file I used notes of following people

- David Rahman
- Simeng Zeng
- Kim Sau Chung
- Yinxiao Chu
- Jonathan Levin
- Mathew O. Jackson