Advanced Microeconomic Theory 521B

Dirk Bergemann Jorge Balat

April 30, 2008

Abstract

The following notes contain the main material of the second part of 521b 2008.

1 Bayesian Games

Information / Knowledge / Uncertainty

There are two important contributions in the literature

1. Harsanyi (1967-68)

   Models without common knowledge are transformed into games of *incomplete* information, and then into *imperfect* information games.

   The idea is to introduce a notion of type which represents private information.

2. Mertens and Zamir (1985)

   Type information is represented by an "hierarchical construction."

1.1 "Applied" models

Let θ_i denote agent i's type, and let $u_i(x, \theta_i, \theta_{-i})$ be the payoff relevant types.

We say θ_i is payoff relevant if $\exists \theta_i', \theta_{-i}, x$ such that $u_i(x, \theta_i, \theta_{-i}) \neq u_i(x, \theta_i', \theta_{-i})$.

Agents have some common knowledge regarding the distribution of types, $p(\theta)$, i.e., common prior. Let $p_i(\theta_{-i}|\theta_i)$ be i's belief about other agents' types.

1.1.1 Type space

Θ: primitives of the model (payoff relevant states)

$t_i \in T_i$: agent i's type

Agent i's private info: $\pi_i : T_i \to \Delta(\Theta \times T_{-i})$. For the moment, we will assume Θ finite, and T_{-i} countable.

A type space is a collection $\mathcal{T} = ((T_i)_{i=1}^I, (\pi_i)_{i=1}^I)$

1.1.2 Example: Envelopes

Two agents: Alice and Bob. Both receive a closed envelope. It is common knowledge that the number of dollars in Alice's envelope is odd, and the number of dollars in Bob's envelope is even, and that both amounts are $1 apart, i.e., they are adjacent numbers.

Before opening the envelopes the information can be represented as in Table 1, where the stars denote the possible states of the world.

Table 1:

		Bob					
		0	2	4	6	8	…
	1	⋆	⋆				
	3		⋆	⋆			
Alice	5			⋆	⋆		
	7				⋆	⋆	
	⋮					…	

After each player opens its own envelope (hence private information) we can ask both of them "Is A richer than B?", and (except if B gets 0) neither know which one is richer.

Now let's make a public announcement: "No one has more than or equal to \$1000".

After the announcement, we ask repeatedly "Are you sure you are richer than your opponent?", the answer is public, and continue asking until a someone answers Yes. Can anything be learned? Say $A < 999$, then both answer No. Thus, B immediately learns that $A < 999$.

Having common knowledge on some event is important for the inference process even in environments where there is no hope of learning. In the previous example, say $A = 3$ and $B = 4$, there will be no learning without any further information, but after the announcement of the public event "$A, B < 1000$" they end up learning which one is richer.

Let E be the event "nobody has strictly more than \$5". This is represented in table 2.

Table 2:

		Bob					
		0	2	4	6	8	…
	1	⋆	⋆				
	3		⋆	⋆			
Alice	5			⋆	⋆		
	7				X	X	
	⋮					…	

How about the event "everybody knows that nobody has strictly more than \$5"? This is shown in table 3.

Let's now introduce probabilities. Let $p(1-p)^n$ be the probability that the smaller envelope contains n dollars. This information is represented in Table 4.

Now we ask "Is A richer than B?" Alice's belief at state n is

$$Pr_A(\text{A is richer than B}|n) = \frac{p(1-p)^{n-1}}{p(1-p)^{n-1} + p(1-p)^n} = \frac{1}{2-p} > \frac{1}{2} \tag{1}$$

And Bob's belief is

$$Pr_B(\text{A is richer than B}|n) = \frac{p(1-p)^{n}}{p(1-p)^{n-1} + p(1-p)^n} = \frac{1-p}{2-p} < \frac{1}{2} \tag{2}$$

Table 3:

		Bob					
		0	2	4	6	8	...
	1	$\star$	$\star$				
	3		$\star$	$\star$			
Alice	5			$\star$	$\xcancel{\star}$		
	7				$\xcancel{\star}$	$\xcancel{\star}$	
	$\vdots$					...	

Table 4:

		Bob				
		0	2	4	6	...
	1	p	$p(1-p)$			
	3		$p(1-p)^2$	$p(1-p)^3$		
Alice	5			$p(1-p)^4$	$p(1-p)^5$	
	7				$\ddots$	
	$\vdots$					

Thus, each one believes he or she is richer at each state.

Now fix an n. Ask the question, they will answer Yes and No respectively. Ask again, ... They will converge to the same posterior (this is Aumann's "Agreeing to disagree").

1.1.3 "Coordinated attack" problem

Two agents: A, B. A has the resources to attack, while B may or may not have the resources.

A and B want to attack if

1. agent i has the resources
2. agent j also attacks

hence, this is a coordination problem.

Communication protocol: B can send messages to A, but

- B sends a message if he has the resources
- the message may get lost

Let n be the maximum number of messages that could have been sent. The possible states are represented in table 5.

Coordinated attack problem: for the attack to occur we need (i) both players to have the resources, and (ii) both to attack. That is, we need to find n_A, n_B (number of messages) sufficiently large.

Table 5:

		B's resources					
		0	2	4	6	8	…
A's resources	1	$\star$	$\star$				
	3		$\star$	$\star$			
	5			$\star$	$\star$		
	7				$\star$	$\star$	
	⋮					…	

Claim 1 *Coordinated attack is impossible for all n.*

Proof. Suppose $\exists n_A^*, n_B^*$ such that $\forall n$ such that $n_A \geq n_A^*, n_B \geq n_B^*$ they attack.

The requirement that both have the resources imply that $n_B^* > 0$, and the requirement that both attack requires $n_A^* \geq n_B^* + 1$ and $n_B^* \geq n_A^* + 1$, which is a contradiction. ■

Therefore, there is no coordination when there is less than common knowledge.

2 Games with Incomplete Information (Jeffrey Ely)

Example: Envelopes. Two players are given envelopes with money inside, and are asked if they want to exchange the envelopes. There are two important questions: Will they trade?, and How is the decision to trade related to the amount of money in the envelope?

There are two approaches to model these games:

1. Type space: a player has a type that summarizes all that he knows and the beliefs about other players' types
2. States of the world

Going back to the example, list of things we need to specify in the model:

- To trade each player hast to ask himself what is the amount of money that both envelopes have. So the *belief* of the amount of money in the other envelope is going to matter. But this is not the only thing,
- The players also care about what the beliefs of the other players are. That is, the players also need to form a belief of what amount of money is on the other envelope conditional on that the other player is willing to trade. Hence, we will need to incorporate the belief of how much money is in the other player's envelope and the belief of what the other player's belief is
- …

Let Ω = set of realizations of underlying uncertainty. In our example, $\Omega = R \times R$. Also let $\beta_i^1 = \Delta(\Omega)$ be the first order belief, that is, the set of all probability distributions over Ω (the player will have *one* belief in that set). Now let the 2nd order belief be $\beta_i^2 = \Delta(\Omega \times \beta_{-i}^1)$, that is, the player has to form a belief about Ω and what the belief of player $-i$ is. But this is not enough, $\beta_i^3 = \Delta(\Omega \times \beta_{-i}^3)$, …

Therefore, $\{\beta_i^s\}_{s=1}^{\infty}$ are necessary elements of the model. This is the *language* that we need. Maybe it is not enough but it is necessary.

Now the problem is evident: we cannot even write down the model, less we can analyze it. The solution will be to model it in a different way such that still captures the relevant features.

Harsanyi Type Space. We will introduce an auxiliary object, set of types for i, T_i, and the belief mapping for i, $\mu_i : T_i \to \Delta(\Omega \times T_{-i})$ (note that $\mu_i \neq \beta_i^s$ for any s).

Note that this is a simpler model *and* captures everything from the starting model.

We require: μ_i to be measurable, and the conditional belief over Ω given t_i and t_{-i}, $\mu_i(t_i)(t_{-i}) \in \Delta(\Omega)$, to be jointly measurable.

Claim 2 *The type space and μ_i fully describes the hierarchies of beliefs.*

Proof. Fix a type t_i for i, and a mapping μ_i. Then

$$\beta_i^1(t_i) = \text{marg}_\Omega \mu_i(t_i)$$

Since the marginal is measurable, we have that the mapping $\beta_i^1 : T_i \to \Delta(\Omega)$ is measurable. Now we need to find $\beta_i^2(t_i) \in \Delta(\Omega \times \Delta\Omega)$. Let $E \subset \Omega \times \Delta\Omega$ be a measurable set, then define

$$\beta_i^2(t_i)(E) = \mu_i(t_i)(\{(\omega, t_{-i}) : (\omega, \beta_i^1(t_i)) \in F\})$$

Now continue in the same fashion for the higher order beliefs. ■

2.1 Relationship Between the Type Space and Hierarchies of Beliefs

Question: Is there any restriction that our simple model puts into the type of β_i^s we can model?

There are four valid answers:

1. I can model everything (Mertens-Zamir)
2. I can model even more than that (Rationalizability) (Type space is a *richer language*)
3. I can model essentially nothing
4. I can model only non-robust assumptions

Answer 1. Mertens-Zamir

There exists a (single) universal type space such that for all hierarchies of beliefs ...

Main idea: Suppose we take an hierarchy, how can we be assured that there exists a type space that describes it?

Let's pick an hierarchy

$$\beta_i = (\beta_i^1, \beta_i^2, \ldots) \in U_i(\Omega).$$

Reinterpret it as a probability measure over Ω (i.e., $\Delta(\Omega \times U_{-i}(\Omega))$, and denote it $\mu_i^*(\beta_i)$. Thus, taking marginals of this probability measure we can get all the β_i^s.

Note that we need to require Ω to be a complete, separable, metric space.

Coherence: any hierarchy captured by the probability measure $\Delta(\Omega \times U_{-i}(\Omega))$ has to be coherent (for example, the marginal of each β_i^s with respect to Ω has to be the same).

(Brandenburger and Dekel (1993))

Idea: use the Kolmogorov extension.

Now take $(U_i(\Omega), \mu_i^*)$ and this will be our type space !

Every hierarchy has a "representative" in the type space (universal type space).

When we write down a type space (i.e., a subset of the universal type space) we are implicitly making assumptions of what hierarchies we can represent, hence making restrictions in the hierarchies.

[Counterexample: Heifetz and Samet (1888)]

Answer 2. Rationalizability

Say we have a game with two players, and let A_i be the set of actions of player i. Define $A = A_1 \times A_2$. The type relevant payoffs (which summarize everything what matters to make a decision) are given by $u_i : A \times \Omega \to R$. Denote a behavioral strategy by $\sigma_i : T_i \to \Delta A_i$.

A *solution concept* is a mapping from the description of the game into strategies (or sets of).

A strategy σ_i for i is a best reply to σ_{-i} if for every t_i

$$\sigma_i(t_i) \in \arg\max_{a_i \in A_i} E_{\mu(\cdot|t_i)} u_i(a_i, \sigma_{-i}(t_{-i}), \omega)$$

Once we have a definition of best reply we then have a notion of rationalizability. For example, throw away everything that is not a best reply in the first round, ... If T_i and A_i are finite then we will eventually stop eliminating.

Note that the best reply definition given above is an ex ante definition. Another definition of best reply is the following: a_i is an interim best reply to a strategy σ_{-i} for a given type t_i if

$$a_i \in \arg\max_{a'_i \in A_i} E_{\mu(\cdot|t_i)} u_i(a'_i, \sigma_{-i}(t_{-i}), \omega)$$

Note that in this definition we are fixing t_i, and choosing σ_{-i}. Also note that σ_i is an interim best reply if for each t_i there is a σ_{-i} against which $\sigma_i(t_i)$ is an interim best reply.

We say that a profile of strategies $\epsilon = (\epsilon_1, \epsilon_2)$ has the best reply property if for each i, for each $\sigma_i \in \epsilon_i$ the strategy σ_i is an interim best reply given ϵ_{-i}, and it has the fixed-point property if there is no $\sigma_i \notin \epsilon_i$ with this feature.

Correlated rationalizability

Consider 2-players games. Replace the definition

$$\sigma_{-i} : T_{-i} \to \Delta A_{-i}$$

with

$$\sigma_{-i} : T_{-i} \times \Omega \to \Delta A_{-i},$$

that is, player i may think that the type of $-i$ could be correlated with the state of nature.

Example. 2 types, 2 players, 2 states of nature, with probabilities given by

	t_2	t'_2
t_1	1/4	0
t'_1	0	1/4

ω

	t_2	t'_2
t_1	0	1/4
t'_1	1/4	0

ω'

It is common knowledge that each player assigns equal probability to the states of nature given their type.

Now suppose the game they play is

	a_2	b_2	c_2
a_1	$1, 1$	$-10, -10$	$-10, 0$
b_1	$-10, -10$	$1, 1$	$-10, 0$
c_1	$0, -10$	$0, -10$	$0, 0$

ω

	a_2	b_2	c_2
a_1	$-10, -10$	$1, 1$	$-10, 0$
a_1	$1, 1$	$-10, -10$	$-10, 0$
a_1	$0, -10$	$0, -10$	$0, 0$

ω'

Note that all a, b, and c are rationalizable because they are all part of a Bayesian Nash Equilibrium.

But, there is a much simpler type space representation with only one type and probabilities given by

$$\begin{array}{cc} & t_2^* \\ t_1^* & \boxed{1/2} \\ & \omega \end{array} \qquad \begin{array}{cc} & t_2^* \\ t_1 & \boxed{1/2} \\ & \omega' \end{array}$$

With this type space representation they play (c, c). Hence the two representations give different results (i.e., they imply different behavior predictions in the game) even though they represent the *same* hierarchy of beliefs !

The point is that the same beliefs can be represented by different type spaces which in turn give different predictions about behavior. Thus we need a much richer language to distinguish the two.

$\Delta-$**hierarchies** (Ely and Peski (2006))

$\Delta-$hierarchies are necessary and sufficient (to predict behavior) for rationalizability in all 2-player games.

If two type spaces have the same $\Delta-$hierarchy then they will have the same rationalizable strategies in all games the players are presented with.

Answer 4.

There exist hierarchies that are similar but give us different predictions. We call them critical types.

Universal type space $U_i(\Delta\Omega)$ endowed with a product topology. It has a partial order, and this order is non-trivial (in Mertens-Zamir is trivial).

Common p-belief

$$F \subset U(\Delta\Omega)$$
$$C^p(F) \subset U(\Delta\Omega)$$

where $C^p(F)$ is the subset of common p-belief of the universal type space that attaches p-belief to the set F.

Theorem 1 *An hierarchy $u_i \in U_i(\Delta\Omega)$ is a critical type iff there is a closed, upper-contour, proper subset $W \subset U(\Delta\Omega)$ such that $u_i \in C^p(W)$ for some $p > 0$.*

The idea is that common p-belief is important for behavior.

Definition. u_i is a critical type if there is a game G, an action a_i such that a_i is rationalizable for u_i, but for some sequence $v_i^k \to u_i$, and for some $\epsilon > 0$, a_i is not ϵ-rationalizable for any v_i^k.

All type spaces modeled in the literature are "fragile".

Answer 3.

The idea is that the measure of hierarchies that we can model with critical types is really small.

3 Robust Mechanism Design (Bergemann and Morris (2005))

3.1 Introduction

Theoretical and practical success of Auction, Mechanism Design and Implementation literature… Mechanisms seem too complicated to use in practise… Successful applications of theory include ad hoc restrictions… Simplicity, non-parametric, detail free, ex post equilibrium, Wilson doctrine…

3.2 The Wilson Doctrine

"Game theory has a great advantage in explicitly analyzing the consequences of trading rules that presumably are really common knowledge; it is deficient to the extent that it assumes other features to be common knowledge, such as one agent's probability assessment about another's preferences or information.

I foresee the progress of game theory as depending on successive reductions in the base of common knowledge required to conduct useful analyses of practical problems. Only by repeated weakening of common knowledge assumptions will the theory approximate reality." Wilson (1987)

3.3 Weakening Common Knowledge

In game theory, Harsanyi (1967-68), Mertens and Zamir (1985) established that relaxing common knowledge assumptions is equivalent to adding types…

Environments with incomplete information can be modeled as a Bayesian game where wlog there is common knowledge among players of each player's type spaces and each type's beliefs over types of other players.

Economic analysis assumes smaller type spaces than universal type space *yet maintains common knowledge*

Are the implicit common knowledge assumptions that come from working with small type spaces problematic? Especially in mechanism design: **?** on surplus extraction. Especially in auctions: i) no strategic uncertainty among bidders, ii) designer and bidder i have identical information about all other bidders.

3.4 Bergemann and Morris' Approach

Introduce rich (higher order belief) types into mechanism design literature.

Fix payoff types and social choice correspondence (objective of designer is to implement social choice correspondence)

For fixed environment, we can construct many type spaces, where a player's type specifies: i) his payoff type, and ii) his belief type.

Crucially, there may be many types of a player with the same payoff type.

3.5 Comparative Statics

Study role of common knowledge assumptions by comparative statics on the type space, going from "naive" type space to "universal" type space.

The larger the type space is, the more incentive constraints there are, the harder it becomes to implement scc.

More "robust" the resulting mechanism will be.

Smallest type space: "naive type space" (possible types equal to payoff types). Standard construction in mechanism design.

Largest type space: "universal type space" (allow any (higher order) beliefs about other players' payoff relevant type).

3.6 Ex Post Equivalence

This paper focusses on "ex post equivalence". Question: When is (Bayesian) equilibrium (partial) implementation on the universal type space equivalent to ex post implementation? Answer: sometimes yes, sometimes no.

In private values case, ex post implementation is equivalent to dominant strategies implementation. Thus we answer the question: when is Bayesian implementation equivalent to dominant strategies implementation?

3.7 Payoff Environment

- Agent $i \in \mathcal{I} = \{1, 2, ..., I\}$
- i's "payoff type" $\theta_i \in \Theta_i$
- payoff type profile $\theta \in \Theta = \Theta_1 \times ... \times \Theta_I$
- social outcome $a \in A$

- utility function $u_i : A \times \Theta \to \mathbb{R}$
- social choice correspondence $F : \Theta \to 2^A / \{\varnothing\}$
- social choice function $f : \Theta \to A$.

Examples: efficient allocation, efficient allocation with budget balance.

3.8 Type Spaces

Richer type spaces than set of payoff types Θ

- i's *type* is $t_i \in T_i$
- t_i includes description of
  - payoff type: $\widehat{\theta}_i : T_i \to \Theta_i$. $\widehat{\theta}_i (t_i)$ is i's payoff type of t_i.
  - beliefs about types of other players: $\widehat{\pi}_i : T_i \to \Delta (T_{-i})$. $\widehat{\pi}_i (t_i)$ is i's belief type of t_i.
- type space is a collection

$$\mathcal{T} = \left\{ T_i, \widehat{\theta}_i, \widehat{\pi}_i \right\}_{i=1}^{I}$$

Example: Small Type Space

$T_i = \{l, h\}$, $\Theta_i = \{\theta_l, \theta_h\}$

Payoff type $\widehat{\theta}_i : T_i \to \Theta_i$, here identity mapping:

$$\begin{aligned} \widehat{\theta}_i &: l \longmapsto \theta_l \\ \widehat{\theta}_i &: h \longmapsto \theta_h \end{aligned}$$

Belief type given by common prior $\pi(t)$:

$$\begin{array}{ccc} & l & h \\ l & \frac{1}{3} & \frac{1}{6} \\ h & \frac{1}{6} & \frac{1}{3} \end{array}$$

Here $\widehat{\pi}_i (t_i) \in \Delta (t_{-i})$:

$$\begin{aligned} \widehat{\pi}_i &: l \longmapsto \left(\frac{2}{3}, \frac{1}{3} \right) \\ \widehat{\theta}_i &: h \longmapsto \left(\frac{1}{3}, \frac{2}{3} \right) \end{aligned}$$

Her belief type identifies payoff type.

Example: Large Type Space

$T_i = \{l, lh, h\}$, $\Theta_i = \{\theta_l, \theta_h\}$

Payoff type $\widehat{\theta}_i : T_i \to \Theta_i$,

$$\begin{aligned} \widehat{\theta}_i &: l \longmapsto \theta_l \\ \widehat{\theta}_i &: lh \longmapsto \theta_l \\ \widehat{\theta}_i &: h \longmapsto \theta_h \end{aligned}$$

Belief type given by common prior $\pi(t)$:

$$\begin{array}{cccc} & l & lh & h \\ l & 2 & a & 1 \\ lh & a & 2a^2 & 2a \\ h & 1 & 2a & 2 \end{array} \times \frac{1}{6+6a+2a^2}$$

For $a = 0$: large type space = small type space.
For $a > 0$:

$$\begin{aligned} \widehat{\pi}_i(l) &= \left(\frac{2}{a+3}, \frac{a}{a+3}, \frac{1}{a+3}\right) \\ &\neq \\ \widehat{\pi}_i(lh) &= \widehat{\pi}_i(h) = \left(\frac{1}{2a+3}, \frac{2a}{2a+3}, \frac{2}{2a+3}\right) \end{aligned}$$

Here belief type does not identify payoff type.

3.8.1 Properties of Type Spaces

- Type Space $\mathcal{T}$ is "naive" if $T_i = \Theta_i$ and each $\widehat{\theta}_i$ is the identity map.
- Type Space $\mathcal{T}$ is finite if each T_i is finite.
- Finite Type Space $\mathcal{T}$ has full support if $\widehat{\pi}_i(t_i)[t_{-i}] > 0$ for all i and t.
- Finite Type Space $\mathcal{T}$ satisfies the common prior assumption if there exists $\pi \in \Delta(T)$ such that

$$\sum_{t_{-i}} \pi(t_i, t_{-i}) > 0 \text{ for all } i \text{ and } t_i$$

and

$$\widehat{\pi}_i(t_i)[t_{-i}] = \frac{\pi(t_i, t_{-i})}{\sum\limits_{t'_{-i}} \pi(t_i, t'_{-i})} > 0.$$

3.8.2 Richer Type Spaces

- A Fixed Full Support Naive Common Prior Type Space
- All Full Support Naive Common Prior Type Spaces
- All Naive Common Prior Type Spaces
- All Common Prior Type Spaces
- All Type Spaces

modulo some technical details, union of all type spaces is identical to universal type space constructed a la Mertens-Zamir from payoff types...

3.9 Interim Implementation

Can we find a mechanism such that one equilibrium delivers outcomes consistent with F?

We appeal to the revelation principle and restrict attention to direct mechanisms where player i reports his type $t_i \in T_i \ldots$

Definition 1 *A social choice function $f \in A^T$ is interim incentive compatible on type space $\mathcal{T}$ if*

$$\int\limits_{t_{-i}} u_i \left(f\left(t_i, t_{-i}\right), \widehat{\theta}\left(t_i, t_{-i}\right) \right) d\widehat{\pi}_i\left(t_i\right) \geq \int\limits_{t_{-i}} u_i \left(f\left(t_i', t_{-i}\right), \widehat{\theta}\left(t_i, t_{-i}\right) \right) d\widehat{\pi}_i\left(t_i\right)$$

for all i, $t \in T$ and $t_i' \in T_i$.

Definition 2 *Social choice function $f \in A^T$ on $\mathcal{T}$ is a selection from F if*

$$f\left(t\right) \in F\left(\widehat{\theta}\left(t\right)\right)$$

for all $t \in T$.

Set of permissible allocations is conditioned on $\widehat{\theta}(t)$
Selection of f can be conditioned on t

Definition 3 *A social choice correspondence F is interim implementable on $\mathcal{T}$ if there exists $f \in A^T$ such that f is interim incentive compatible on $\mathcal{T}$ and f is a selection from F.*

3.10 Ex Post Implementation

Definition 4 *A social choice function $f \in A^\Theta$ is ex post incentive compatible if, for all i and $\theta \in \Theta$,*

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

for all $\theta_i' \in \Theta_i$.

Compare: A scf is dominant strategy incentive compatible if for all i and all θ, θ' :

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

Definition 5 *A social choice correspondence F is ex post implementable if there exists $f \in A^\Theta$ such that f is ex post incentive compatible and*

$$f\left(\theta\right) \in F\left(\theta\right)$$

for all $\theta \in \Theta$.

3.11 Ex Post Equivalence

"Ex post equivalence" holds for a given environment if:

> SCC F is ex post implementable if and only if it is interim implementable on the union of all type spaces.

Ex post equivalence holds if

- the SCC is a function;

- if SCC represents (efficient) outcomes in quasi-linear environment without budget balance

Ex post equivalence fails in general.

If SCC represents efficient outcomes in quasi-linear environment with budget balance, then ex post equivalence holds if

- $\#\Theta_i = 1$ for some i;
- $\#\Theta_i \leq 2$ for all i;
- $I \leq 2$

Ex post equivalence fails in an example with $I = 3$, $\#\Theta_1 = 3$ and $\#\Theta_2 = \#\Theta_3 = 2$.

3.12 Example 1

Easy to interim implement on any type space. Impossible to ex post implement.

- $\Theta_1 = \{\theta_1, \theta_1'\}$, $\Theta_2 = \{\theta_2, \theta_2'\}$, $A = \{a, b, c\}$
- Payoffs

a	θ_2	θ_2'
θ_1	$1, 0$	$-1, 2$
θ_1'	$0, 0$	$0, 0$

b	θ_2	θ_2'
θ_1	$-1, 2$	$1, 0$
θ_1'	$0, 0$	$0, 0$

c	θ_2	θ_2'
θ_1	$0, 0$	$0, 0$
θ_1'	$1, 1$	$1, 1$

- Social Choice Correspondence

F	θ_2	θ_2'
θ_1	$\{a, b\}$	$\{a, b\}$
θ_1'	$\{c\}$	$\{c\}$

3.13 Example 2: Allowing Lotteries

Easy to interim implement on any type space. Impossible to ex post implement.

- $\Theta_1 = \{\theta_1, \theta_1', \theta_1''\}$
- $\Theta_2 = \{\theta_2, \theta_2'\}$
- $A = \Delta(\{a, b, c, d\})$.
- Payoffs

a	θ_2	θ_2'
θ_1	$1, 0$	$-1, 2$
θ_1'	$0, 0$	$0, 0$
θ_1''	$0, 0$	$0, 0$

b	θ_2	θ_2'
θ_1	$-1, 2$	$1, 0$
θ_1'	$0, 0$	$0, 0$
θ_1''	$0, 0$	$0, 0$

c	θ_2	θ_2'
θ_1	$-1, 0$	$\frac{1}{3}, 0$
θ_1'	$1, 1$	$1, 1$
θ_1''	$0, 0$	$0, 0$

d	θ_2	θ_2'
θ_1	$\frac{1}{3}, 0$	$-1, 0$
θ_1''	$0, 0$	$0, 0$
θ_1'	$1, 1$	$1, 1$

- Social Choice Correspondence:

	θ_2	θ_2'
θ_1	$\{a, b\}$	$\{a, b\}$
θ_1'	$\{c\}$	$\{c\}$
θ_1''	$\{d\}$	$\{d\}$

Consider the following randomized mechanism:

	θ_2	θ_2'
θ_1	$(p, 1-p)$ on $\{a, b\}$	$(p', 1-p')$ on $\{a, b\}$
θ_1'	$\{c\}$	$\{c\}$
θ_1''	$\{d\}$	$\{d\}$

For θ_1 to tell the truth if he is sure that θ_2:

$$p - (1-p) \geq \frac{1}{3} \Leftrightarrow p \geq \frac{2}{3}$$

For θ_1 to tell the truth if he is sure that θ_2':

$$-p' + (1-p') \geq \frac{1}{3} \quad \Leftrightarrow \quad p' \leq \frac{1}{3}$$

But for θ_2 to tell the truth if he is sure that θ_1:

$$2(1-p) \geq 2(1-p') \quad \Leftrightarrow \quad p' \geq p.$$

Thus

$$p' \geq p.$$

- this example has interdependent values and F independent of θ_2; we can dispense with both these assumptions easily...

3.14 Example 3

scc is not ex post implementable, interim implementable on ANY naive type space, not interim implementable on some larger type spaces.

- $I = 2$, $\Theta_+ = \{\theta_+^1, \theta_+^2, \theta_+^3\}$, $\Theta_- = \{\theta_-^1, \theta_-^3\}$.
- $A = \{-1, -\frac{7}{8}, 0, \frac{7}{8}, 1\} \times \{+, -\}$
- Preferences over location:

$$u_i(a, \theta_i, \theta_j) = u_j(a, \theta_i, \theta_j) = (\theta_i - \theta_j)\, a.$$

- Utility $\frac{1}{8}$ from having own name.
- Planner maximizes sum of utility minus a^2.

3.15 General Environments

Proposition 1 *If F is ex post implementable, then F is interim implementable on any type space*

Proposition 2 *If F is single valued, then F is ex post implementable if and only if F is interim implementable on every $\mathcal{T}$*

Proposition 3 *F is ex post implementable if and only if F is interim implementable on every $\mathcal{T}$ using (WITH loss of generality) mechanisms where players ONLY report their payoff type*

c.f., Ledyard (1978) and Dasgupta, Hammond, and Maskin (1979)...

3.16 Augmented Ex Post Incentive Compatibility

- strengthen equilibrium conditions to obtain general equivalence
- each agent i reports θ_i and $m_i \in M_i(\theta_i)$.

Definition 6 *A social choice rule* $f \in A^M$ *is augmented ex post incentive compatible given* $\mathcal{M}$ *if, for all* i, $\theta_i \in \Theta_i$ *and* $\lambda_i \in \Delta(M_{-i})$, *there exists* $m_i \in M_i(\theta_i)$ *such that* (θ_i, m_i) *is a maximizer of*

$$\int \lambda_i(\theta_{-i}, m_{-i}) u_i\left(f\left((\theta_i', \theta_{-i}), (m_i', m_{-i})\right), (\theta_i, \theta_{-i})\right).$$

Definition 7 *A social choice correspondence* F *is augmented ex post implementable if there exists* $\mathcal{M}$ *and* $f \in A^M$ *such that* f *is augmented ex post incentive compatible and*

$$f(\theta, m) \in F(\theta)$$

for all $(\theta, m) \in M$.

Proposition 4 *Social Choice Correspondence* F *is augmented ex post implementable if and only if* F *is interim implementable on every* $\mathcal{T}$.

3.17 Quasi-Linear Environment

- $A = Z \times \mathbb{R}^I$
- $u_i((z,y), \theta) = v_i(z, \theta) + y_i$
- $F_\xi(\theta) = \{(z,y) \in A : z = \xi(\theta)\}$
- $v_i : Z \times \Theta \to \mathbb{R}$ for each i and $\xi : \Theta \to Z$.

Lemma 1 *If* F_ξ *is interim implementable on every full support common prior naive type space* $\mathcal{T}$, *then* F_ξ *is ex post implementable.*

Proposition 5 *The following are equivalent:*
(1) F_ξ *is interim implementable on all type spaces;*
(2) F_ξ *is interim implementable on all common prior type spaces;*
(3) F_ξ *is interim implementable on all common prior naive type spaces;*
(4) F_ξ *is interim implementable on all common prior full support naive type spaces;*
(5) F_ξ *is ex post equilibrium implementable in the naive type space.*

3.17.1 Interim Implementability for Quasi-Linear Environments

Key properties of type spaces:

Definition 8 $\mathcal{T}$ *satisfies one-to-one property if*

$$\widehat{\pi}_i(t_i') = \widehat{\pi}_i(t_i) \Rightarrow \widehat{\theta}_i(t_i') = \widehat{\theta}_i(t_i).$$

Definition 9 $\mathcal{T}$ *satisfies product property if for all* θ_i *and* π_i *in the range of* $\widehat{\pi}_i$, $\exists t_i \in T_i$ *s.t.* $\widehat{\pi}_i(t_i) = \pi_i$ *and* $\widehat{\theta}_i(t_i') = \theta_i$.

3.17.2 Large Payoff Beliefs

- i's beliefs about other agents' payoff types:
$$\psi_i \in \Delta\left(\Theta_{-i}\right)$$
- type depedent payoff beliefs, let
$$\widehat{\psi}_i : \Pi_i \rightarrow \Delta\left(\Theta_{-i}\right)$$
be defined by:
$$\widehat{\psi}_i\left(\pi_i\right)\left[\theta_{-i}\right] \triangleq \sum_{\left\{t_{-i}:\widehat{\theta}_{-i}\left(t_{-i}\right)=\theta_{-i}\right\}} \widehat{\pi}_i\left(t_i\right)\left[t_{-i}\right]$$

3.17.3 Interim Payoffs

- player i's expected utility of outcome $\xi\left(\theta_i', \theta_{-i}\right)$:
$$\widehat{v}_i\left(\theta_i, \theta_i', \psi_i\right) \triangleq \sum_{\theta_{-i}\in\Theta_{-i}} \psi_i\left(\theta_{-i}\right) v_i\left(\xi\left(\theta_i', \theta_{-i}\right), \left(\theta_i, \theta_{-i}\right)\right)$$
- i's beliefs ψ_i over opponent's payoff type
- i has payoff type θ_i
- i reports payoff type θ_i'

3.17.4 Large Interim Implementability

Theorem 2 *ξ is interim implementable iff $\forall i$, $\exists\ y_i^* : Z_i \rightarrow \mathbb{R}$ s.th. $\forall\left(\theta_i, \pi_i\right), \left(\theta_i', \pi_i\right) \in Z_i$:*

$$v_i\left(\theta_i, \theta_i, \widehat{\psi}_i\left(\pi_i\right)\right) + y_i^*\left(\theta_i, \pi_i\right) \geq v_i\left(\theta_i, \theta_i', \widehat{\psi}_i\left(\pi_i\right)\right) + y_i^*\left(\theta_i', \pi_i\right)$$

- $(*)$ holds if type space satisfies one-to-one property
- if one-to-one property fails, $(*)$ reduces to conditions with independent types in naive type space
- if enough beliefs over payoff types of opponent are consistent with a given belief type
$\Rightarrow$ reduces to ex post implementability conditions

3.18 Quasi-Linear Environment with Budget Balance

- $A = Z \times \mathbb{R}^I$
- $u_i\left(\left(z, y\right), \theta\right) = v_i\left(z, \theta\right) + y_i$
- $F_\xi\left(\theta\right) = \left\{\left(z, y\right) \in A : z = \xi\left(\theta\right) \text{ and } \sum\limits_{i=1}^{I} y_i = 0\right\}$
- $v_i : Z \times \Theta \rightarrow \mathbb{R}$ for each i and $\xi : \Theta \rightarrow Z$

3.18.1 Ex Post Implementability

- gains from misrepresentation:

$$\zeta_i\left(\theta_i,\theta_i',\theta_{-i}\right)\equiv v_i\left(\xi\left(\theta_i',\theta_{-i}\right),\theta\right)-v_i\left(\xi\left(\theta_i,\theta_{-i}\right),\theta\right).$$

- transfer functions $y=(y_1,...,y_I)$, each $y_i:\Theta\to\mathbb{R}$
- ex post incentive compatibility:

$$y_i\left(\theta_i,\theta_{-i}\right)-y_i\left(\theta_i',\theta_{-i}\right)\geq\zeta_i\left(\theta_i,\theta_i',\theta_{-i}\right)$$

  for all i, θ_i, θ_i' and θ_{-i}.

- budget balance:

$$\sum_{i=1}^{I}y_i\left(\theta_i,\theta_{-i}\right)=0$$

  for all $\theta\in\Theta$.

3.18.2 Linear Programming

Ex post implementation as linear programming problem:

$$\max_{\{y_i\}}\left\{\sum_{i=1}^{I}\sum_{\theta\in\Theta}\alpha_i\left(\theta\right)y_i\left(\theta\right)\right\}$$

subject to

$$y_i\left(\theta_i,\theta_{-i}\right)-y_i\left(\theta_i',\theta_{-i}\right)\geq\zeta_i\left(\theta_i,\theta_i',\theta_{-i}\right),\qquad\lambda_i\left(\theta_i,\theta_i',\theta_{-i}\right)$$

subject to

$$\sum_{i=1}^{I}y_i\left(\theta_i,\theta_{-i}\right)=0,\qquad\nu\left(\theta\right)$$

with associated multipliers

$$\lambda_i:\Theta_i^2\times\Theta_{-i}\to\mathbb{R}_+;$$

and

$$\nu:\Theta\to\mathbb{R}$$

3.18.3 Dual Approach

- Idea of argument
  - proof by contrapositive
  - not ex post implementable implies not interim implementable
- Theorem of alternative (farkas lemma) for dual characterization
- Recover ex post constraints in interim constraints through rich type space

Dual Characterization of Ex Post Implementability

- multipliers ν and λ satisfy the flow condition (F) if:

$$\nu(\theta) = \sum_{\theta_i'} \lambda_i(\theta_i, \theta_i', \theta_{-i}) - \sum_{\theta_i'} \lambda_i(\theta_i', \theta_i, \theta_{-i})$$

- mulitpliers λ satisfy the weighting condition (W) if

$$\sum_{i=1}^{I} \sum_{\theta \in \Theta} \sum_{\theta_i'} \lambda_i(\theta_i, \theta_i', \theta_{-i}) \zeta_i(\theta_i, \theta_i', \theta_{-i}) > 0;$$

Lemma 2 *SCC F_ξ is ex post implementable if and only if there exists (ν, λ) satisfying F and W*

Dual Characterization of Interim Implementability

- transfer functions $y = (y_1, ..., y_N)$, each $y_i : T \to \mathbb{R}$
- interim incentive compatibility:

$$\sum_{t_{-i}} \widehat{\pi}_i[t_i](t_{-i}) [y_i(t_i, t_{-i}) - y_i(t_i', t_{-i})] \geq \sum_{t_{-i}} \widehat{\pi}_i[t_i](t_{-i}) \zeta_i\left(\widehat{\theta}_i(t_i), \widehat{\theta}_i(t_i'), \widehat{\theta}_{-i}(t_i)\right), \quad \forall i, t_i, t_i'.$$

- budget balance:

$$\sum_{i=1}^{I} y_i(t_i, t_{-i}) = 0, \ \forall t$$

- multipliers $\widetilde{\nu} : T \to \mathbb{R}$ and $\widetilde{\lambda} = \left(\widetilde{\lambda}_1, ..., \widetilde{\lambda}_I\right)$, each

$$\widetilde{\lambda}_i : T_i^2 \to \mathbb{R}_+;$$

- satisfy the flow condition (F) if

$$\widetilde{\nu}(\theta) = \sum_{t_i'} \widehat{\pi}_i[t_i](t_{-i}) \widetilde{\lambda}_i(t_i, t_i') - \sum_{t_i'} \widehat{\pi}_i[t_i](t_{-i}) \widetilde{\lambda}_i(t_i', t_i)$$

- satisfies the weighting condition (W) if

$$\sum_{i=1}^{I} \sum_{t \in T} \sum_{t_i'} \widehat{\pi}_i[t_i](t_{-i}) \lambda_i(t_i, t_i') \zeta_i\left(\widehat{\theta}_i(t_i), \widehat{\theta}_i(t_i'), \widehat{\theta}_{-i}(t_{-i})\right) > 0;$$

Lemma 3 *SCC F_ξ is interim implementable if and only if there exists $\left(\widetilde{\nu}, \widetilde{\lambda}\right)$ satisfying F and W*

3.19 Results

Proposition 6 *If $N = 2$, F_ξ^* is ex post implementable if and only if F_ξ^* is interim implementable on all common prior type spaces*

Example 3 showed that F_ξ^* may fail to be interim implementable on all common prior type spaces, even when it is interim implementable on all common prior naive type spaces.

Proposition 7 *If $\#\Theta_i \leq 2$, the following are equivalent:*
(1) F_ξ is interim implementable on all type spaces;
(2) F_ξ is interim implementable on all common prior type spaces;
(3) F_ξ is interim implementable on all common prior naive type spaces;
(4) F_ξ is interim implementable on all common prior full support naive type spaces;
(5) F_ξ is ex post equilibrium implementable in the naive type space.

An example (with $N = 3$ and $\#\Theta_i \geq 3$ for some i) shows that F_ξ^* may be interim implementable on all type spaces, even when it is not ex post implementable.

3.20 Future Questions

- Full Implementation
- Revenue Maximization
- Single Crossing Conditions in Rich Type Spaces

4 Dynamic Allocation Problems

Dynamic allocation problems are also known as *multi-armed bandit problems.* This terminology comes from Las Vegas. *One-armed bandits* are machines where you pull an arm causing a dial with possible outcomes to spin, and where after the dial stops spinning you only see one outcome (but do not see the other possible outcomes in the dial). Hence, there is uncertainty about the dial (i.e., the distribution of returns).

Now suppose we face many such dials (hence, the term multi-armed). With many bandits the problem is to identify the highest expected return bandit. You can sample the bandits. To make it an economically interesting problem, you can sample one bandit at a time and you discount the future. Basically, this is a sequence of stopping problems (but not terminal stopping problems because you can go back to a previous bandit).

This literature started in the '50s-'60s for testing drugs and later to determine for how long to try a new technology.

4.1 Deterministic Problem

Say we have N alternatives denoted by $i = 1, \ldots, N$, with returns $x_i(t_i) \in R_+$, where t_i is the count of times (or periods) of usage of alternative i.

Note that the return depends on how many times we have used the alternative, but it is independent of the usage of alternative j, $\forall j \neq i$. It could be the case that $x_i(t_i + 1) < x_i(t_i)$, that is, the alternative experiences depreciation; or it could be that $x_i(t_i + 1) > x_i(t_i)$, for example from learning-by-doing.

Also note that in this problem the mapping $x_i : N \to R_+$ is known and deterministic.

Let $t = 0, 1, \ldots$ denote time, then $\sum_{i=1}^N t_i = t$. Also let $0 < \delta < 1$ be the discount factor.

Our problem is to allocate the alternatives across time. In fact, this is a classical scheduling problem (we can use one alternative at a time and we need to find the optimal order of tasks).

Now let's denote a policy by $t^* = (t_1^*, \ldots, t_N^*)$, where $t_i^* : N \to N$ is the counter of alternative i, with $t_I^*(t+1) \in \{t_i^*(t), t_i^*(t) + 1\}$, that is, the counter of i can stay the same or increase by 1.

We also require the policy to satisfy $\sum_{i=1}^N t_i^*(t) = t \; \forall t$.

We can now state the problem as

$$\max_{\{t_i(t)\}_{i=1}^N} \sum_{t=0}^{\infty} \left(\delta^t \sum_{i=1}^{N} x_i(t_i(t))[t_i(t+1) - t_i(t)] \right) \tag{3}$$

subject to

$$t_i^*(t+1) \in \{t_i^*(t), t_i^*(t)+1\} \tag{4}$$

$$\sum_{i=1}^{N} t_i^*(t) = t, \quad \forall t \tag{5}$$

Note that a sufficient condition for the problem to be well defined is $\forall i \sum_{t=0}^{\infty} \delta^t x_i(t) < \infty$.

We are going to solve this problem using dynamic programming. Rewrite the problem as

$$V(t_1, \ldots, t_N) = \max_{\{t_i(t+1)\}_{i=1}^N} \left\{ \sum_i x_i(t_i)[t_i(t+1) - t_i(t)] + \delta V(t_1(t+1), \ldots, t_N(t+1)) \right\} \tag{6}$$

subject to (4) and,

$$\sum_{i=1}^{N} t_i(t+1) = t+1 \tag{7}$$

Suppose for a moment that $x_i(t_1+1) < x_i(t_i) \quad \forall i, t_i$. The optimal solution will be to pick i^* such that $i^* = \text{argmax}_i x_i(t_i)$ at $(t_1, \ldots, t_N)$. Hence, the solution is myopic.

Now suppose that $x_i(t_i) = \alpha_i + \epsilon_i t_i$, with $\alpha_i, \epsilon_i \in R_+$ all i. We are going to suggest a solution criterion based on the individual stopping time.

Define for alternative i the average payoff if previous usage of i is s_i

$$v_i(s_i) = \max_{\tau_i} \left\{ \frac{\sum_{t=0}^{\tau_i} \delta^t x_i(t+s_i)}{\sum_{t+0}^{\tau_i} \delta^t} \right\} \tag{8}$$

where τ_i is the stopping time. We also call v_i the index of alternative i.

The optimal index policy is to choose at each state $s = (s_1, \ldots, s_N)$ alternative i with the highest index $v_i(s_i)$.

Note that we now have N 1-dimensional optimization problems !

A simple proof is the following. Say the use of alternatives prescribed under the optimal policy is

$$\begin{array}{l} \text{time: } \quad \overbrace{1 \quad 2 \quad 3 \quad 4}^{i} \quad \overbrace{5 \quad 6 \quad 7}^{j} \quad \overbrace{8 \quad 9 \quad 10}^{i} \quad \overbrace{11 \quad 12}^{k} \quad \overbrace{13 \quad 14 \quad 15 \quad 16}^{i} \quad \ldots \\ \text{payoffs: } \quad \underbrace{x_i(1) \quad x_i(2) \quad x_i(3) \quad x_i(4)}_{\Leftrightarrow \quad v_i(0) \text{ with } \tau_i^*=4} \qquad \underbrace{x_i(5) \quad x_i(6) \quad x_i(7)}_{\Leftrightarrow \quad v_i(5) \text{ with } \tau_i^*=3} \qquad \underbrace{x_i(8) \quad x_i(9) \quad x_i(10) \quad x_i(11)}_{\Leftrightarrow \quad v_i(8) \text{ with } \tau_i^*=4} \end{array}$$

Note that for each chunk we can replace the payoffs with the average (or index).

It is easy to see that the average for each successive chuck of alternative i should be decreasing over time (otherwise we can stick them together in a bigger chunk, i.e., stop later).

Now since the averages are decreasing we are back to the myopic rule but replacing payoffs with averages.

Going back to dynamic programming. Say we now have to choose between alternative i and a constant retirement option m. The problem is now

$$V_i(t_i, m) = \max \left\{ x_i(t_i) + \delta V_i(t_i+1, m), \frac{m}{1+\delta} \right\} \tag{9}$$

The question is now how to choose the largest m such that

$$x_i(t_i) + \delta V_i(t_i + 1, m) \geq \frac{m}{1+\delta}$$

5 Dynamic Marginal Contribution Mechanism (Bergemann and Välimäki (2007))

5.1 Introduction

Intertemporal Efficiency with Private Information. The idea is to extend VCG to dynamic environments (in which the auctions and/or information change over time).

Motivation:

- random arrival of buyers, sellers and/or objects
  - selling seats for an airplane with random arrival of buyers
  - bidding on ebay
  - bidding for construction projects with uncertain arrival of new projects
- bidding for links in sponsored search (Google, Yahoo, etc.)
  - uncertainty about click-through probability
  - uncertainty about conversion probability
- leasing resource over time
  - auction of renewable license, right, capacity over time
  - web serving, computational resource (bandwidth, CPU)

5.1.1 Benchmark: Static Efficiency with Private Information

In private value environment:

- Vickrey (1961): single or multiple unit discriminatory auctions implement socially efficient allocation
  - in private value environments
  - in (weakly) dominant strategies
- Clarke (1971) and Groves (1973) extend to general allocation problems in private value environments
  - agent i internalizes the social objective and is led to report her type truthfully

Pivot Mechanism

- Green and Laffont (1977) analyze specific VCG mechanism
- i internalizes social objective if i pays her externality cost
- externality cost of i

utility of $I\backslash i$ when i is absent - utility of $I\backslash i$ when i is present

- marginal contribution of i = utility of i - externality cost of i
- in Pivot mechanism:

1. payoff of i is her marginal contribution to social value
2. participation constraint holds ex post and no budget deficit

5.1.2 Dynamic Marginal Contribution Mechanism

- payoff in Pivot mechanism = marginal contribution
- develop marginal contribution mechanism in intertemporal environments with new arrival of information regarding:
  - preferences
  - agents
  - allocations
- design sequence of payments so that each agent receives **flow** marginal contribution in every period
- solve intertemporal problem as a completely contigent plan
- embed intertemporal problem in a static problem (as in an Arrow Debreu economy) ...
- ... and then appeal to the classic VCG results.
- BUT the contingent view fails to account for strategic possibilities of the agents in the sequential model

5.1.3 Sequential Incentive and Participation Constraints

- information arrives over time
- report of agent i in period t responds to private information of agent i, but may also respond to past reports of other agents (possibly inferred from allocative decisions)
- truthtelling (generally) fails to be a weakly dominant strategy
- with forward looking agents, participation constraint is required to be satisfied at every point in time (and not only in the initial period)

5.1.4 Results

- marginal contribution mechanism is dynamically efficient
- periodic ex post: with respect to information available at period t
- satisfies (periodic) ex post incentive constraints
- satisfies (periodic) ex post participation constraints
- adding efficient exit condition (weak “online” condition): if agent i does not impact future decisions, then agent i does not receive future payments, uniquely identifies marginal contribution mechanism

5.2 Literature

- Dolan (1978): priority queuing
- Parkes et al. (2003): delayed VCG without participation or budget balance constraints
- Bergemann and Välimäki (2006): complete information, repeated allocation of single object over time, first price bidding
- Athey and Segal (2007): balanced budget rather than participation constraints

5.3 Scheduling

- scheduling tasks
- discrete time, infinite horizon: $t = 0, 1, ...$
- common discount factor δ
- finite number of agents: $i \in \{0, 1, ..., I\}$
- each agent i has a single task
- value of task for i is:

$$v_i > 0$$

- quasilinear utility: $v_i - p_i$
- values are given wlog in descending order:

$$v_0 > v_1 > \cdots > v_I > 0$$

- marginal contribution of task i : difference in welfare with i and without i
- efficient task assignment policy:

$$\begin{array}{rcccccccccc}
\text{policy without } i: & 0 & 1 & \cdots & i-1 & \text{i+1} & \text{i+2} & \cdots & & \text{I} & \\
 & & & & & \searrow & \searrow & & & \searrow & \\
\text{policy with } i: & 0 & 1 & \cdots & i-1 & \text{i} & \text{i+1} & \text{i+2} & \cdots & & \text{I} \\
 & & & & & \uparrow & & & & &
\end{array}$$

Marginal Contribution

- insert valuable task i:
- raise the value of all future tasks: $t > i$
- marginal contribution M_i:

$$M_i = \sum_{t=0}^{I} \delta^t v_t - \left(\sum_{t=0}^{i-1} \delta^t v_t + \sum_{t=i+1}^{I} \delta^{t-1} v_t \right)$$

or

$$M_i = \sum_{t=i}^{I} \delta^t \left(v_t - v_{t+1} \right) \geq 0$$

Externality

- from marginal contribution to externality pricing:

$$M_i = v_i - p_i$$

- externality cost of task i is:

$$p_i = v_{i+1} - \sum_{t=i+1}^{I} \delta^{t-i} \overbrace{(v_t - v_{t+1})}^{>0}$$

- task i directly replaces task $i+1$, but also:
- task i raises the value of all future tasks

Incomplete Information

- v_i is private information to agent i at $t = 0$
- incentive compatibility and efficient sorting: when would i like to win against j versus $j+1$:

$$(\mathrm{v_i} - v_j) - \sum_{t=j}^{I} \delta^{t-(j-1)} (v_t - v_{t+1}) \geq \delta (\mathrm{v_i} - v_{j+1}) - \sum_{t=j+1}^{I} \delta^{t-j} (v_t - v_{t+1})$$

- reduces to cost of delay:

$$(1-\delta)\,\mathrm{v_i} \geq (1-\delta)\,v_j.$$

- report thruthful if others report truthful: ex post equilibrium

Bidding vs Direct Revelation Mechanism

- an ascending (English) auction in every period
- winning bidder i pays bid of second highest bidder
- bid by agent i in period t:

$$b_i^t$$

- bid should reflect value of task but ...
- value of task today versus value of task tomorrow
- value = utility - option value
- bidding strategy b_i^t determined recursively in i and t
- option value is value of realizing task tomorrow

$$\delta\left(v_i - p_i^{t+1}\right)$$

and the price tomorrow is

$$p_i^{t+1} \triangleq \max_{j \neq i} \left\{..., b_j^{t+1}, ...\right\}$$

- net value of realizing task today is

$$v_i - \delta \left(v_i - p_i^{t+1} \right)$$

Dynamic Bidding

- bidding strategy of agent i is given

$$b_i^t = v_i - \delta \left(v_i - p_{i+1}^t \right) = (1 - \delta) v_i + \delta b_{i+1}^{t+1}$$

- ascending auction gives efficient assignment in all periods
- Bergemann and Välimäki (2006): dynamic price competition, complete information, first price bidding
- Edelman, Ostrovsky, and Schwartz (2007): static price competition, incomplete information, second price bidding

5.4 Information Arrival: Licensing

- sequential allocation of a single indivisible object with initially uncertain value to the bidders
- bidder i receives additional information only in periods in which i is assigned the object
- license to use facility or to explore resource for a limited time

Examples

- renting store space in mall
  - current winner (lessee) gets traffic data, purchase behavior
  - current looser does not get traffic data, purchase behavior
- bidding for keywords
  - current winner gets information about click-through rate, sales conversion rate
  - current looser doesn't get information about click-through rate, sales conversion rate

Single Unit Auction

- single unit auction repeated over time
- discrete time, infinite horizon: $t = 0, 1, ...$
- finite number of bidders: $i \in \{1, ..., I\}$
- realized value of object for winning bidder in period t is

$$v_{i,t} = \omega_i + \varepsilon_{i,t}$$

- $\varepsilon_{i,t}$ is i.i.d. over time with $\mathbb{E}[\varepsilon_{i,t}] = 0$
- ω_i is true value of object
- $\varepsilon_{i,t}$ is random noise

Information Flow

- at $t = 0$: common prior distribution $F_i(\omega_i)$ for each agent i
- at $t \geq 0$: **winning bidder** receives informative signal $s_{i,t+1}$:

$$s_{i,t+1} = v_{i,t} = \omega_i + \varepsilon_{i,t}$$

- realized value in period t constitutes private information for period $t + 1$
- at $t \geq 0$: **loosing bidders** don't receive additional information:

$$s_{i,t+1} = s_{i,t}$$

Histories

- private history of bidder i:

$$s_i^t = (s_{i,0}, ..., s_{i,t})$$

- superscript ^{t} for vector of histories, subscript $_t$ for periodic event
- **expected value** for bidder i in perid t :

$$v_{i,t}\left(s_i^t\right) \triangleq \mathbb{E}\left[\omega_i \left| s_i^t \right.\right]$$

5.4.1 Dynamic Direct Mechanism

- bidder i is asked to report her signal in every period t
- initial reports:

$$r_0 = (r_{1,0}, ..., r_{I,0})$$

- inductively, a history of reports:

$$r^t = \left(r^{t-1},\ r_{1,t}, ..., r_{I,t}\right) \in R^t$$

- allocation rule:

$$x_t : R^{t-1} \times R_t \to [0,1]^I$$

- transfer (or pricing) rule is given by:

$$p_t : R^{t-1} \times R_t \to \mathbb{R}^I$$

Strategies

- reporting strategy for agent i:

$$r_{i,t} : R^{t-1} \times S_i \to S_i.$$

- expected payoff for bidder i :

$$\mathbb{E} \sum_{t=0}^{\infty} \delta^t \left[x_{i,t}\left(r^t\right) v_i\left(s_i^t\right) - p_{i,t}\left(r^t\right)\right].$$

- reporting strategy of i solves sequential optimization problem $V_i(s_i^t, r^{t-1})$:

$$\max_{r_{i,t}\in S_i} \mathbb{E}\left\{x_{i,t}\left(r^t\right) v_{i,t}\left(s_i^t\right) - p_{i,t}\left(r^t\right) + \delta V_i\left(s_i^{t+1}, r^t\right)\right\}$$

- taking expectation $\mathbb{E}$ wrt $(s_{-i,t}, r_{-i,t})$

Equilibrium

- denote by $s_{-(i,t)}^t \triangleq s^t \setminus s_{i,t}$
- *Bayesian incentive compatible* if $r_{i,t} = s_{i,t}$ solves

$$\max_{r_{i,t}\in S_i} \mathbb{E}\left\{x_{i,t}\left(r_{i,t}, s_{-(i,t)}^t\right) v_i\left(s_i^t\right) - p_{i,t}\left(r_{i,t}, s_{-(i,t)}^t\right) + \delta V_i\left(r_{i,t}, s_{-(i,t)}^t\right)\right\}$$

- periodic ex post: with respect to all the information available at period t
- (periodic) *ex post incentive compatible* if $r_{i,t} = s_{i,t}$ solves

$$\max_{r_{i,t}\in S_i} \left\{x_{i,t}\left(r_{i,t}, s_{-(i,t)}^t\right) v_i\left(s_i^t\right) - p_{i,t}\left(s_{i,t}, s_{-(i,t)}^t\right) + \delta V_i\left(r_{i,t}, s_{-(i,t)}^t\right)\right\}$$

  for all $s_{-i,t} \in S_{-i}$

Social Efficiency

- socially efficient assignment policy

$$W\left(s^u\right) = \max_{\{x_t(s^t)\}_{t=u}^{\infty}} \mathbb{E}\sum_{t=u}^{\infty}\sum_{i=1}^{N} \delta^{t-u} x_{i,t}\left(s^t\right) v_i\left(s_i^t\right)$$

- optimal assignment is a multi–armed bandit problem
- optimal policy is an index policy:

$$\gamma_i\left(s_i^u\right) = \max_{\tau} \mathbb{E}\left\{\frac{\sum_{t=0}^{\tau} \delta^t v_i\left(s_i^{u+t}\right)}{\sum_{t=0}^{\tau} \delta^t}\right\}$$

- socially efficient allocation policy $\mathbf{x}^* = \{x_t^*\}_{t=0}^{\infty}$:

$$x_{i,t}^* > 0 \text{ if } \gamma_i\left(s_i^t\right) \geq \gamma_j\left(s_j^t\right) \text{ for all } j.$$

Marginal Contribution

- value of social program after removing bidder i

$$W_{-i}\left(s^u\right) = \max_{\{x_{-i,t}(s^t)\}_{t=u}^{\infty}} \mathbb{E}\sum_{t=u}^{\infty}\sum_{j\neq i} \delta^{t-u} x_j^t\left(s^t\right) v_j\left(s_j^t\right)$$

- *marginal contribution* $M_i\left(s^t\right)$ of bidder i at history s^t is:

$$M_i\left(s^t\right) = W\left(s^t\right) - W_{-i}\left(s^t\right)$$

- value M conditional on history s^u and allocation x_u:

$$M\left(s^u, x_u\right)$$

- *flow marginal contribution* $m_i\left(s^t\right)$:

$$M_i\left(s^t\right) = m_i\left(s^t\right) + \delta M_i\left(s^t, x_t^*\right)$$

Flow Marginal Contribution

- flow marginal contribution:

$$m_i\left(s^t\right) = M_i\left(s^t\right) - \delta M_i\left(s^t, x_t^*\right)$$

- expanding flow term with respect to time

$$m_i\left(s^t\right) = \overbrace{\left(W\left(s^t\right) - W_{-i}\left(s^t\right)\right)}^{\text{M}_\text{i}\text{ starting at t}} - \overbrace{\delta\left(W\left(s^t, x_t^*\right) - W_{-i}\left(s^t, x_t^*\right)\right)}^{\text{M}_\text{i}\text{ starting at t+1 and x}_\text{t}^*}$$

- expanding flow term with respect to identity (rearranging)

$$m_i\left(s^t\right) = \overbrace{\left(W\left(s^t\right) - \delta W\left(s^t, x_t^*\right)\right)}^{\text{current value with i}} - \overbrace{\left(W_{-i}\left(s^t\right) - \delta W_{-i}\left(s^t, x_t^*\right)\right)}^{\text{current value without i but x}_\text{t}^*}$$

Efficient Assignment

$$m_i\left(s^t\right) = \left(W\left(s^t\right) - \delta W\left(s^t, x_t^*\right)\right) - \left(W_{-i}\left(s^t\right) - \delta W_{-i}\left(s^t, x_t^*\right)\right)$$

- consider efficient assignment $x_t^* = i$:
- information about $x_t^* = i$ is worthless without i:

$$W_{-i}\left(s^t, i\right) = W_{-i}\left(s^t\right)$$

leads to

$$m_i\left(s^t\right) = v_i\left(s_i^t\right) - (1-\delta)\, W_{-i}\left(s^t\right)$$

- consider inefficient bidder: $x_t^* \neq j$:

$$x_{-j,t}^* = x_t^*$$

leads to

$$m_j\left(s^t\right) = 0$$

Dynamic Second Price Auction

- match net payoff to flow marginal contribution
- for winner i:

$$m_i\left(s^t\right) = v_i\left(s^t\right) - p_i\left(s^t\right)$$

- for losers, $j \neq i$:

$$m_j\left(s^t\right) = -p_j\left(s^t\right)$$

Result

Theorem 3 (Dynamic Second Price Auction) *The socially efficient allocation rule* $\mathbf{x}^*$ *satisfies ex post incentive and participation constraints with payment* $\mathbf{p}^*$*:*

$$p_j^*\left(s^t\right) = \begin{cases} (1-\delta)\, W_{-j}\left(s^t\right) & \text{if } x_{j,t}^* = 1, \\ 0 & \text{if } x_{j,t}^{t*} = 0. \end{cases}$$

- price equals intertemporal opportunity cost
- delay $(1-\delta)$ of the optimal program for all but j

5.5 Dominant versus Ex Post Incentive Compatibility

- with private values, static mechanism satisfies incentive compatibility in weakly dominant strategies
- in dynamic mechanism, dominant incentive compatibility fails to hold in private value environment
- truthtelling after all histories fails to be a weakly dominant strategy as it removes the ability to respond to past announcements
- yet ex post incentive compatibility can be satisfied in dynamic mechanism

5.6 General Allocation Problems

- description of a dynamic Vickrey-Clarke-Groves mechanism
- general specification of utility of each agent and arrival of private information over time
- dynamic VCG mechanism is time consistent
  - social choice function can be implemented by a sequential mechanism without ex ante commitment by the designer
  - thruthtelling strategy in the dynamic setting forms an ex-post equilibrium rather than an equilibrium in weakly dominant strategies
- extend single unit auction to general allocation model
- net expected flow utility of agent i in period t :

$$v_i\left(x_t, s_i^t\right) - p_{i,t}$$

- private signal of agent i in period $t+1$ is generated by conditional distribution function:

$$s_{i,t+1} \sim G_i\left(\cdot \left| x_t, s_i^t\right.\right).$$

- generalize information flow by allowing signal $s_{i,t+1}$ of agent i in period $t+1$ to depend on current decision x_t and entire past history of private signals of i

Dynamic VCG Mechanism

- efficiency
- marginal contribution pricing

Theorem 4 (Dynamic VCG Mechanism) *The socially efficient allocation rule* $\{x^*\}$ *satisfies ex post incentive and ex post participation constraint with payment* p^**:*

$$p_{i,t}^*\left(x^*\left(s^t\right), s_{-i}^t\right) = v_i\left(x^*\left(s^t\right), s_i^t\right) - m_i\left(s^t\right).$$

- characterization of transfer prices via marginal contribution
- in specific environments additional insights from observing how social policies are affected by removal of agent i

Efficient Exit

- many transfer rules support ex post incentive and ex post participation constraints in dynamic setting
- temporal separation between allocative influence and monetary payments may be undesirable for may reasons:
  - agent i could be tempted to leave the mechanisms and break her commitment *after* she ceases to have a pivotal role but *before* her payments come due
  - if arrival and departure of agents is random, then an agent could falsely claim to depart to avoid future payments
- in intertemporal environment if agent i ceases to influence current or future allocative decisions in period t, then she also ceases to have monetary obligations
- presence of i is not pivotal for the future at τ and s^τ if

  $$x_t^*\left(s^t\right) = x_{-i,t}^*\left(s^t\right),$$

  for all $t \geq \tau$ and for all $s^t = (s^\tau, s_{\tau+1}, ..., s_t)$

Definition 10 (Efficient Exit) *A mechanism satisfies the efficient exit condition if for all* i*,* τ*, and* s^τ :

$$p_{i,t}\left(s^t\right) = 0, \quad \textit{for all } t \geq \tau.$$

Uniqueness

- weak online requirement: decisions regarding agent i have to be made in the presence of agent i

Theorem 5 (Uniqueness) *If a dynamic direct mechanism is efficient, satisfies ex post incentive and participation constraints and the efficient exit condition, then it is the dynamic marginal contribution mechanism.*

5.7 Conclusion

- direct dynamic mechanism in private value environments with transferable utility
- design of monetary transfers relies on notions of marginal contribution and flow marginal contribution
- transfer the insights of VCG mechanism from static to dynamic settings
- many interesting questions are left open
  - current contribution is silent on issue of revenue maximizing mechanisms
  - characterization of implementable allocations in dynamic setting will first be necessary
  - restriction to private value environments

6 Dynamic Mechanism Design

- $t = 0, 1, \ldots, T, \ldots$ i.e., finite or infinite time horizon
- $0 < \delta \leq 1$ discount factor
- $\theta_{i,t}$ willingness to pay of player i at time t, $i = 1, \ldots, I$

  Note that the type of i is evolving over time
- evolution of type
  $$p(\theta_{i,t+1}|\theta_{i,t}) \neq p(\theta_{i,t+1})$$
  i.e., we have some persistence.

  Special case: fully persistent type $\theta_{i,t+1} = \theta_{i,t}$
- dynamics come from
  - evolution of types
  - evolution of beliefs posteriors

6.1 Coase Conjecture (mechanisms without commitment)

Coase formulated in 1937 his conjecture about durable goods. We will follow Gul, Sonnenschein, and Wilson (1986).

We have one ($i = 1$) informed buyer, with $\theta \sim F(\theta)$, and one uninformed seller. Infinite horizon $t = 0, 1, \ldots$. Let p_t be the price of the single object at time t.

In the static problem
$$\max_p \{p(1 - F(p))\}$$
which is the monopolist problem.

Now if the monopolist sets p_0 and the buyer does not buy, in the next period the seller will want to lower the price
$$\max_p \{p(1 - F_1(p))\}$$
where F_1 is the truncated prior.

Note that the monopolist faces competition, i.e., himself tomorrow. Without commitment, the selling policy tomorrow competes with the selling policy today.

In this context, Coase conjecture is that as $\delta \to 1 \Rightarrow \pi = \underline{\theta}$.

6.1.1 Extensions

Hart and Tirole (1988) and Schmidt (1990) look at the Coase conjecture with rentals rather than sales, but the problem does not go away.

Feinberg and Skrypacz (2005): Coase conjecture relies heavily on the informed party to fully predict the future behavior of the seller. What if there is some uncertainty in higher order beliefs?

Skreta and .. (): generalization to multiagent allocation problems.

6.2 Mechanisms with Commitment

What if now there is commitment from the principal's point of view?

6.2.1 Baron and Besanko (1984)

Now the seller can commit to an entire price path. Result: cannot do better than the static price at all t.

6.2.2 Freixas, Guesnerie and Tirole (1985), Bullow and Stockey (1984)

These authors introduced the term "rachet effect" (a rachet is a dial that allows movement in only one direction).

Regulation model: a firm produces q with cost $c(q, \theta)$, where θ is the firm's private information. The firm is a natural monopolist and is supervised by a regulator. We have two periods $t = 0, 1$.

Note that any information revealed at $t = 0$ will be used against the firm at $t = 1$, hence this is an incentive to overstate costs in the first period. Thus, there will be some information rent.

6.3 Sequential Screening (Courty and Li (2000))

Two periods $t = 0, 1$. Two types of consumers which need to make a single transaction. Say both have to make a trip in the future, and hence need to buy an airline ticket. The idea is that as time passes and they approach the date of the trip they have more information about the value of the trip. This is stated as,

$$p(\theta_{i,t+1}|\theta_{i,t}) \neq p(\theta_{i,t+1})$$

Let the marginal cost of a seat in a plane be $c = 1$. There are two types of travelers: leisure and business travelers. At time $t = 0$, leisure travelers have a valuation distributed $uniform[1, 2]$, and business travelers have a valuation distributed $uniform[0, 1] \cup uniform[2, 3]$, i.e, the meeting may be of low or high value. The airlines knows that $\frac{1}{3}$ are leisure travelers and $\frac{2}{3}$ business travelers. At time $t = 1$ all information is realized.

Now, what is the optimal selling policy of the airline at $t = 1$?

Valuation in the market will be distributed $uniform[0, 3]$, hence the problem of the firm is

$$\max_p (p - c)(1 - F(p))$$

The optimal price will be $p^* = 2$, with profits $\pi^* = 1 \times \frac{1}{3} = \frac{1}{3}$.

Note that at $t = 1$ we have a lot of types, in fact the continuum $[0, 3]$. But at $t = 0$ we only have 2 types (leisure and business).

At $t = 0$ the firm will offer for leisure travelers an unconditional ticket (i.e., with no refund) at price $p_l = 1.5 = \mathrm{E}[\theta_{i,l}]$. And for business travelers it will offer a ticket at a higher price but with the option to a refund if the the realized value at $t = 1$ is low.

Note that all business travelers have to pay p_b, but $\frac{1}{2}$ of them will use the refund r_b. Hence the following must hold for business travelers

$$\frac{1}{2}\left(\mathrm{E}[\theta_{i,b}|\text{high value meeting}] - p_b\right) + \frac{1}{2}\left(r_b - p_b\right) = \frac{1}{2}2.5 - p_b + \frac{1}{2}r_b \geq 0$$

Under efficient exit, $r_b = 1$. Thus,

$$\begin{aligned} p_b &= 1.75 \\ r_b &= 1 \end{aligned}$$

And the profits will be

$$\begin{aligned}\pi_0 &= \frac{1}{3}(\underbrace{1.5}_{p_l} - \underbrace{1}_{\text{cost}}) + \frac{2}{3}\left[\frac{1}{2}(\underbrace{1.75}_{p_b} - \underbrace{1}_{\text{cost}}) + \frac{1}{2}(\underbrace{1.75}_{p_b} - \underbrace{1}_{r_b})\right] \\ &= \frac{2}{3} \\ &> \frac{1}{3}\end{aligned}$$

Note that,

1. Over time, incentive constraints multiply, thus more private info $\Leftrightarrow$ more info rent $\Rightarrow$ firm wants to sell the object earlier to avoid the info rent
2. Participation constraint only has to hold ex-ante, thus less private info means less public info $\Rightarrow$ less efficient decision

This is the trade-off in dynamic mechanism design.

Also note that we need commitment, otherwise if the firm offers

$$\begin{cases} p_l^* = 1.5 \\ p_b^* = 1.75, r_b^* = 1 \end{cases}$$

at both times, everyone will wait until the valuation is realized, and the only ticket sold will be the leisure ticket, thus half of the leisure travelers and half of the business travelers buy a ticket.

Now assume that there is only one buyer ($i = 1$) with valuation $v_i \sim uniform[0, 1]$, and that the marginal cost is $c = 0$.

Design problem: choose the information structure for the agent and the choice variable to maximize the profits of the firm. Hence, the designer can choose:

1. info structure: $I_i^* : v_i \to \Delta(s)$
2. p^*

The info structure will determine $\Pr(v_i|s_i, I_i)$, where s_i is the privately observed signal. Given this the buyer will have a realization of the valuation and decide to buy or not.

6.4 Sequential Screening (Eso and Szentes (2007))

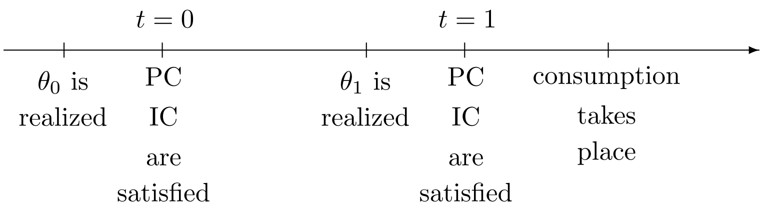
Recall that in Courty and Li (2000) θ_t was the valuation or willingness to pay for a single object, and this valuation evolves over time following a markov process $p(\theta_{t+1}|\theta_t)$ (i.e., more information arrives over time). The choice variable from the seller's point of view is the optimal time t^* to sell the object. And hence, the participation and incentive constraints have to hold at time t^* (remember the seller can commit not to sell before or after t^*, and the consumers have to want to participate at t^*, no later, nor before).

Eso and Szentes (2007) allow for decision/commitment by the agents in all time periods, i.e., not only at t^*.

Let the value of the object to a buyer be $v = \alpha\theta_0 + (1 - \alpha\theta_1)$, $\alpha \in [0, 1]$, where θ_0 is realized at $t = 0$, and θ_1 at $t = 1$.

We will look at a mechanism for agents to reveal their information at $t = 0$ and $t = 1$. Now the participation and incentive constraints should be satisfied in period 0 *and* 1.

Suppose, for example, the problem of selling a single good to a single agent. We want to design a revenue maximizing selling contract for the seller. The timing is as follows.



A mechanism comprises a transfer at $t = 0$, $t_0(\theta_0)$, a transfer at $t = 1$, $t_1(\theta_0, \theta_1)$, and a quantity (or probability of getting the object), $q(\theta_0, \theta_1)$.

Ex post utility (i.e., after the realization of the θ's) is given by

$$v(\theta_0, \theta_1)q(\theta_0, \theta_1) - t_0(\theta_0) - t_1(\theta_0, \theta_1)$$

Assume the seller has no valuation for the good. Let $\theta_0 \sim F_0(\theta_0)$, $\theta_1 \sim F_1(\theta_1)$ be the common prior, and let θ_0, θ_1 be private information to the buyer.

Consider the following special cases:

1. $\alpha = 1$

   That is, late information has no value for the buyer. The seller chooses the optimal monopoly pricing rule in the static case

   $$p_0^* \in \arg\max_p \{p(1 - F_0(p)\}$$

   And the interim IC and IR hold.

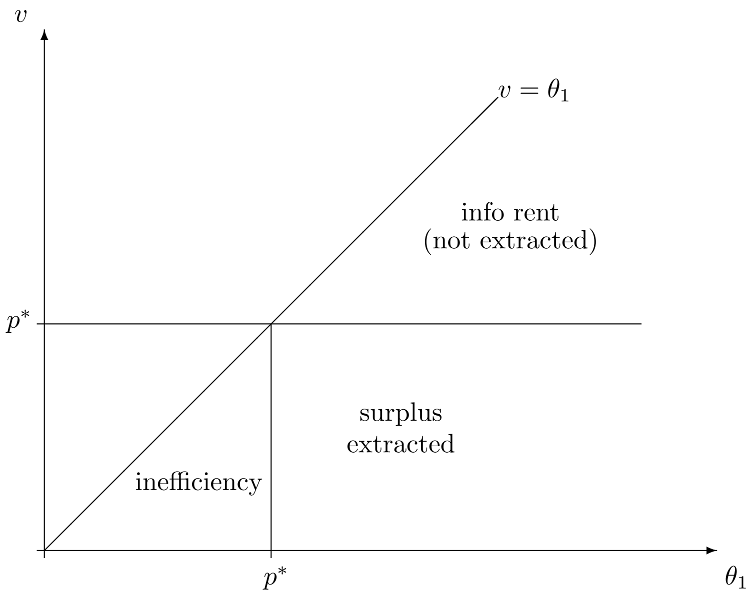
2. $\alpha = 0$

   That is, the initial information has no value. Then, the optimal contract at $t = 0$ is

   $$\begin{aligned} p_0 &= \mathrm{E}[\theta_1] \\ p_1 &= 0 \end{aligned}$$

   By backwards induction, at $t = 1$, if the buyer entered the contract at $t = 0$, the PC holds because $p_1 = 0$. Also, at $t = 0$, since θ_1 is still not realized, the PC is satisfied.

   Note 1: Why isn't $p_0 = 0$, $p_1 = p_1^* \in \arg\max_p\{p(1 - f_1(p)\}$ the optimal contract? There is a trade-off between efficiency and rent extraction by the seller. Say we set a price p^*



Note 2: Generalization to auctions (many bidders). Let $v_i(\theta_i) = \alpha\theta_{0i} + (1 - \alpha_{1i} = \theta_{1i}, \forall i$. Also let $\theta_i \sim F(\theta_i)$ *iid.* From an ex-post point of view, the seller wants to wait till the θ_1's are realized and pick $\theta_1^* = \max\{\theta_{11}, \dots, \theta_{1I}\}$. Thus, use a second price auction without reserve price. This will give us the efficient assignment, but we don't extract all the rent.

Other suggestion, charge $p_0 = \mathrm{E}[\theta_i]$, or can we combine the two contracts? We can charge the net gain to participate in the auction (i.e., the info rent) at the beginning. That is,

$$\begin{aligned} p_0 &= \text{entry fee} \\ &= \text{expected net value of the SPA} \\ &= \frac{1}{I}\mathrm{E}_0\left[\max\{v_i\} - \max\left\{\{v_i\}\backslash\max\{v_i\}\right\}\right] \\ p_1 &= \text{second price} \\ &= \max\left\{\{v_i\}\backslash\{\max\{v_i\}\right\} \end{aligned}$$

If the seller can charge a price before θ is realized, he can extract all rent.

3. $\alpha \in (0, 1)$

   Intuitively, we should combine is some way the two special cases.

   There will be a trade-off between efficiency and informational rents.

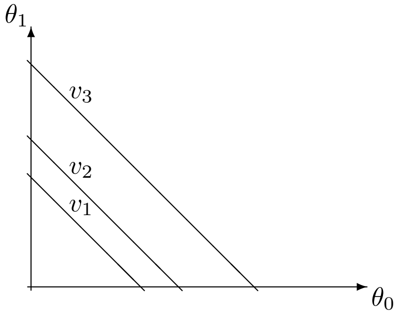
   We want to offer 2 prices: an unconditional payment $t_0(\theta_0)$ (unconditional to the final transaction, but can depend on θ_0, i.e., this is a payment for θ_0-types, unconditional on wether the θ_0-buyer buys the object at $t = 1$ or not. This is like an entrance fee), and a conditional payment $t_1(\theta_0, \theta_1)$ (i.e., a payment conditional on being type θ_0, θ_1 *and* buying the object).

   First guess for t_1, it should be monotonic in $\theta_1|\theta_0$. Second guess for t_1, $t_1(\theta_0)$, i.e., does not depend on θ_1 (this is the logic in the SPA).

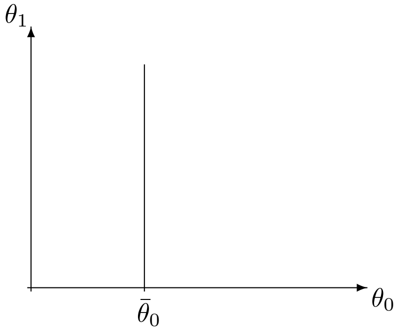
   Hence, we have 2 payments that depend only on θ_0, one with probability 1, the other with positive probability.

   Now, remember $v(\theta_0, \theta_1) = \alpha\theta_0 + (1 - \alpha)\theta_1$.

Say $\alpha = 1/2$, then the isoutilities (i.e., the combinations of θ_0 and θ_1 that give the same level of utility) will take the following form:



Now consider the utility levels that the buyer can achieve given a value of $\theta_0 = \bar{\theta}_0$:



Then it must be the case that

$$\text{sign}\frac{\partial t_0}{\partial \theta_0} \neq \text{sign}\frac{\partial t_1}{\partial \theta_0}$$

otherwise the IC won't hold.

In fact, in the optimal contract

$$\frac{\partial t_0^*}{\partial \theta_0} > 0, \qquad \frac{\partial t_1^*}{\partial \theta_0} < 0$$

6.5 Pavan, Segal and Toikka (2008)

As before, $v(\theta_0, \theta_1) = \alpha\theta_0 + (1-\alpha)\theta_1$.

Previously, we had i) θ_0 and θ_1 independent, and ii) one final decision.

The first assumption is not important, we only have to take the "new" information $\theta_1|\theta_0$ to be independent of θ_0

$$\theta_1|\theta_0 \perp \theta_0$$

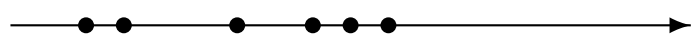
and then we can implement the same mechanism as before.

The big insight of PST is the relaxation of the second assumption. They consider a stream of decisions and analyze dynamic allocations.

7 Internet Auctions

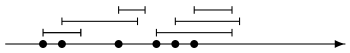
What do we mean by internet or ebay auctions? There are various characteristics

- Large number of buyers and sellers, of course, a large number of objects (similar or not)
- Similar objects are auctioned at different times:



  Objects are offered at random times (from the point of view of the buyers)

- The seller can decide the length of the action. Thus, there will be overlapping auctions, and nonsynchronized (buyers and sellers do not show at the same time, that's why we need to have a window):



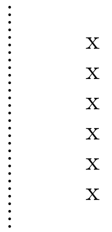
  Note: Traditional art auctions performed in auction houses are synchronized (hence no need for window).

Now the sellers have to choose the entry time, the window, and the reserve price. And the buyers have to choose the entry time in a given auction, and across auctions.

Simultaneous auctions introduce strategic effects. To see this, suppose we have two parallel auctions. The bidder has an incentive to enter by the end of the auction, because if he enters earlier he reveals how much he values the item and, hence, all other bidders jump to the other auction hoping there is no high bidder in the other auction.

7.1 Peters and Severinow (1997)

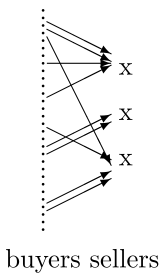
Suppose we have lots of buyers and sellers for an identical good. All auctions are run at the same time. Also suppose that the number of buyers is greater than the number of sellers.



buyers>sellers

At $t = 0$ the sellers choose the price and quantity rules (p_i, q_i) (for example, a reserve price, 2nd price auction).

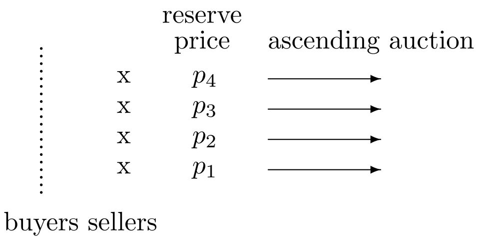
At $t = 1$ buyers decide what auction to enter, and can only enter in only one auction (this is a physical restriction, suppose you have to be present in the auction).



7.2 Peters and Severinow (2006)

Now allow buyers to participate in more than one auction at a time. The format of the auctions are as English (ascending price) auctions.

Suppose we have N buyers and M sellers. From the point of view of the sellers, the buyers enter randomly. The sellers have to decide a reserve price (p_i).



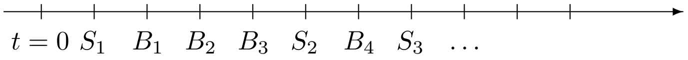
In this setting, we have M simultaneous ascending price auctions.

Main result: There exists a Perfect Bayesian Equilibrium which leads to an ex post efficient allocation (subject to a reserve price). And, as $N \to \infty$, the reserve price goes to 0.

7.3 Nekipelov (2007)

This paper provides an empirical and theoretical contribution to eBay auctions.

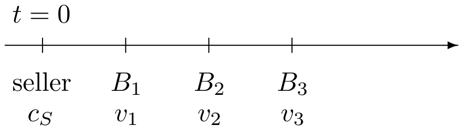
One way to model asynchronous auctions is the following. Suppose we have continuous time, and buyers and sellers entering at random times. Buyers arriving at a rate λ_B, with valuation v (identical to all buyers), and sellers arriving at a rate λ_S, with cost of selling c (identical to all sellers). Clearly, a transaction occurs if the price p is such that $v - p \geq 0$ and $p - c \geq 0$. For example, buyers and sellers can appear as follows:



Let the discounted payoffs be $e^{-r(t-s}(v - p_t)$ for the buyer and $e^{-r(t-s}(p_t - c)$ for the seller. Denote by a_{it} the seller i's asking price, and by b_{jt} the buyer j's offering price. If a transaction occurs, the price is set at $p = \frac{1}{2}(b_{it} + a_{jt})$. Let λ_B, λ_S, v, and c be common knowledge.

In the previous figure note that Buyer 1 cannot set b_1 too low because the seller has the option to wait for the next buyer. In equilibrium, the price will be determined by the number of sellers, and the number of buyers, in fact, what matters for the price determination is just the difference between the number of sellers and the number of buyers. Also, note that in equilibrium sellers and buyers match instantaneously.

What is missing in this model from eBay auctions? For example, buyers arrive with different valuations. Think of a single seller (with cost c_S) and random arrival of buyers (with valuations v_i):



We want efficient trade. If $v_i = v\ \forall i$, then the optimal policy is to sell to the first buyer. If the v_i's differ, then the seller can wait for a better offer (but waiting is costly), then the problem is how long does the seller want wait for a new sample?

7.4 Satterthwaite and Shneyerov (2007)

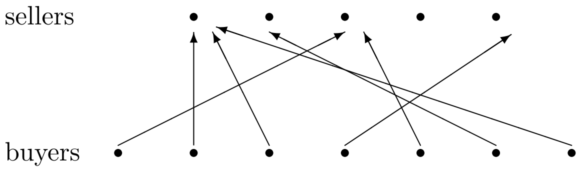
This paper takes a different approach. It is an incomplete information model of decentralized trade. There is a large number of buyers and sellers, actually, a continuum of them. Let buyers have valuation v distributed as $G_B(v)$, with $g_B(v) > \bar{\epsilon}$, and let sellers have valuation c distributed as $G_S(c)$, with $g_S(c) > \bar{\epsilon}$. The price formation rule is a 1st-price auction. Denote $b(v)$ the bids, and $r(c)$ the reserve price.

We see a transaction if and only if $\max\{b(v)\} - r(c) \geq 0$ (and the transaction occurs at the 1st-price rule). Net utilities realized are $v - b(v)$ for the winning buyer, and $b(v) - c$ for the seller.

Suppose we have a measure 1 of sellers and a measure a of buyers. Let δ be the period length.

The benchmark for this model is the static centralized market (or Walrasian market) where all buyers and sellers are present at the same time and time in the auction floor. We want to find the price p_W that clears demand and supply: $G_S(p_W) = a(1 - G_B(p_W))$ (note that the supply $G_S(p_W)$ is increasing in p_W, and demand $a(1 - G_B(p_W))$ decreasing, hence a unique crossing).

Now consider the more realistic model of dynamic decentralized markets. The idea is to allow for an infinite time horizon and analyze the steady state. We will have many simultaneous auctions created by a matching process (every buyer is matched with a single seller). Suppose for the moment a discrete number of buyers ($b = 1, \ldots, B$) and sellers ($s = 1, \ldots, S$). In every period t, buyers are matched uniformly but randomly across sellers. That is, each seller gets a buyer with probability $1/B$. For example, in period t the following match is realized:



If there is a transaction, the buyer and seller get out, they are replaced by a new buyer and seller, there is a new matching, … (In the continuous case, after each period, δ new sellers enter, and $a\delta$ new buyers enter).

At every period, buyers and sellers decide $r(c)$ and $b(v)$ before the market volume is realized (that is, before the matching).

The idea is that each seller, at every period, faces a small sample of buyers, then by setting a high enough $r(.)$ can reject a low realization of buyers and defer the auction to the next period (i.e., the seller can choose to postpone the auction and get a new sample of buyers, from the distribution of buyers).

Let K be the cost of participation per unit of time, hence if a time period is δ, the cost of participation at each period is δK, and the total cost of participation is δK times the number of periods it takes the seller (buyer) to sell (acquire) the item. Also, let β be the discount factor.

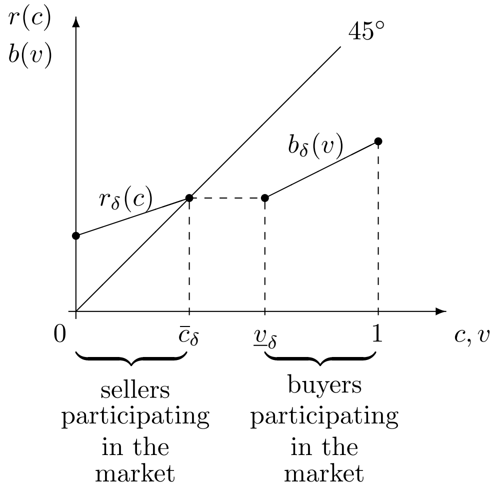
Fix K, we want to analyze the equilibrium as $\delta \to 0$ (i.e, new transactions come at a fast rate) and $\beta \to 0$.

Theorem 6 *Fix $K > 0$, and fix $\beta \geq 0$. Suppose there exits a $\delta > 0$ such that an equilibrium with positive trade exists. Then*

$$\lim_{\delta\to 0} \bar{c}_\delta = \lim_{\delta\to 0} \underline{c}_\delta = \lim_{\delta\to 0} \underline{v}_\delta = \lim_{\delta\to 0} \underline{p}_\delta = \lim_{\delta\to 0} \bar{p}_\delta = p_W,$$

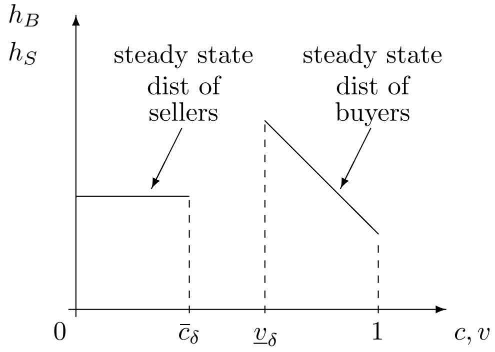
where $\bar{c}_\delta$ is the maximum cost type present in equilibrium, $\underline{c}_\delta$ is the minimum acceptable bid, $\bar{p}_\delta$ is the maximum bid, $\underline{c}_\delta$ is the minimum bid, and $\underline{v}_\delta$ is the lowest valuation in equilibrium.

To see what an equilibrium looks like consider the following. Fix $\delta > 0$



In the Walrasian market, $\bar{c}_\delta - \underline{v}_\delta = 0$, but now, $\bar{c}_\delta - \underline{v}_\delta > 0$, and what determines this difference is the participation cost.

Now consider the steady state distributions of buyers and sellers, h_B and h_S, when $a > 1$, and the (incoming flow) distributions G_B and G_S are Uniform [0,1] and bidding and reserve price strategies are as in the above graph



All sellers that have at least one bidder matched are going to make a transaction, then the steady state distribution is the same as the incoming flow distribution. But for the buyers, $\underline{v}_\delta$ type will only make a transaction if she is lucky to be the only bidder matched with a buyer, hence they will stay more often than the higher types.

Now think about types and willingness to pay in the auction. A buyer can get the item today or wait and get it tomorrow, then her valuation is going to be

$$v - e^{-\delta\beta} W_B(v)$$

where $W_B(v)$ is the continuation value. Hence bids in steady state should be based on this intertemporal value, i.e., $b(v - e^{-\delta\beta}W_S(v))$. Similarly, for the buyers we have,

$$c + e^{-\delta\beta}W_S(c)$$

and the steady state reserve price will be $r(c + e^{-\delta\beta}W_S(c))$.

As $\delta, \beta \to 0$, $e^{-\delta\beta} \to 1$, thus,

$$\begin{aligned} W_B(v) &= \max[v - p_W, 0] \\ W_S(v) &= \max[p_W - c, 0] \end{aligned}$$

and

$$\begin{aligned} b(v - (v - p_W)) &\to b(p_W) \\ r(c + (p_W - c)) &\to r(p_W) \end{aligned}$$

What happens to the private information? It's gone, it becomes irrelevant. The continuation value imposes the heterogeneous players to behave as a homogeneous agent.

8 Information Constraints

8.1 Introduction

In the real world, economic agents face constraints when solving for their optimization problems. We can think of these constraints in two ways: under perfect information, the problem that the agent faces is difficult to solve, or if the agent can solve the problem, he faces imperfect information:

1. Perfect information
   - Complexity of economic decision making. First introduced by H. A. Simon ("satisfycing")
   - Early 70's, Gilboa et. al showed that the complexity of computing a Nash Equilibrium is as difficult as an NP complete problem. Then agents may relax the optimality conditions to allow for easier to solve problems. In experiments, player biases may be explained by simple decision rules, such as a rule of thumb, that they adopt.
   - Complexity of the optimal decision
2. Imperfect information
   - Maybe the decision maker can solve the optimization problem but the constraint she faces is from information processing
   - There is a limit on the information that the decision maker can process
   - With limited information (i.e., with less variance in the information) we will have less variance in the decision

Let ω be the true state of the world. An agent chooses an action y to

$$\min(y - \omega)^2$$

In a world with perfect information the decision rule would be $y^*(\omega) = \omega$. What happens if we have imperfect information? Say that instead of observing ω we observe a noisy signal

$$x = \omega + \epsilon_x$$

with $\epsilon_x \sim N(0, \sigma_x^2)$, $\omega \sim N(0, \sigma_\omega^2)$. Then the optimal policy cannot depend on ω, but will depend on x, $y^*(x)$.

Let the agent know the conditional probability distribution $p(x|\omega)$. The agent wants to infer ω after observing x, that is, $p(\omega|x)$. We can use Bayes' rule:

$$p(\omega|x) = \frac{Pr(x \cap \omega)}{Pr(x)}$$

An estimate of ω given that we observe a signal x is

$$\begin{aligned}
\mathrm{E}[\omega|x] &= \frac{\frac{1}{\sigma_\omega^2}0 + \frac{1}{\sigma_x^2}x}{\frac{1}{\sigma_\omega^2} + \frac{1}{\sigma_x^2}} \\
&= \frac{\frac{1}{\sigma_x^2}x}{\frac{1}{\sigma_\omega^2} + \frac{1}{\sigma_x^2}} \\
&= x\frac{1}{1 + \frac{\sigma_x^2}{\sigma_\omega^2}}
\end{aligned}$$

Then the problem of the agent is to minimize the expected loss

$$\min \mathrm{E}_x (y - \omega)^2$$

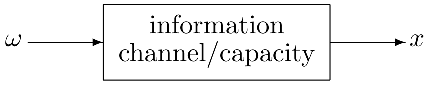
and the optimal decision rule will be $y^* = x\frac{1}{1+\frac{\sigma_x^2}{\sigma_\omega^2}}$.

One way of saying that we have "less information" is to say that σ_x^2 is large (limited information means $\sigma_x^2 > 0$). Thus, we are going to have less variance in our posterior ($\mathrm{E}(\omega|x)$) and less variance in our decision $y^*(x)$.

The idea is that information constraints lead to slow or sticky decision making. This idea is formalized in Sims (1998), Sims (2003), Sims (2006), and Woodford (2007).

8.2 Introduction to Information Theory

Say we have an input ω and an output x, *information* encodes the input into the output:



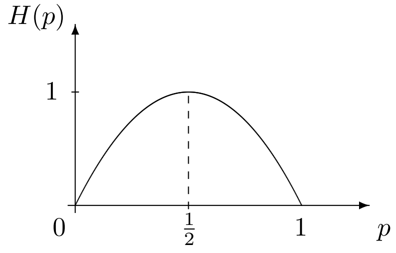
Thus, information constraints are constraints on the capacity. We are going to use some notions of Information Theory from engineering (Shannon (1948)).

We have a random variable $\omega \sim p(\omega)$. For example, consider a discrete random variable $\omega \in \{\omega_1, \ldots, \omega_n\}$.

Definition 11 *The entropy of a random variable $p(\omega)$ is defined by*

$$\begin{aligned}
H(p) &= -\sum_\omega p(\omega) \log_2 p(\omega) \\
&= \sum_\omega p(\omega) \log_2 \frac{1}{p(\omega)} \\
&= E_p \log_2 \frac{1}{p(\omega)}
\end{aligned}$$

Example. Coin toss: outcomes $\{H, T\}$ with probabilities $p, (1-p)$. Then, $H(p) = -p \log p - (1-p)\log(1-p)$



Remark. Why do we use base 2 logarithms in the definitions? This comes from encoding data into bits. Consider the following example. Suppose there is a horse race, there are 32 horses $n = 1, \ldots, 32$, there is one winner and all horses are equally likely to win, i.e., $Pr(n \, \text{wins}) = 1/32 \, \forall n$. The problem is how to communicate the identity of the winner to someone else. In fact, we need 5 bits to carry the information of the winner's identity, since $2^5 = 32$. Now, let's compute the entropy

$$\begin{aligned} H(p) &= -32\frac{1}{32}\log_2\frac{1}{32} \\ &= 5 \end{aligned}$$

which is exactly the number of bits needed.

Now suppose we have 8 horses, again one winner, but the probabilities of winning are $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{64}, \frac{1}{64}, \frac{1}{64}, \frac{1}{64}$. How many bits do we need to communicate the identity of the winner? An upper-bound, i.e., one code for each horse, is 3 bits since $2^3 = 8$, but $H(p) = 2$ and $H(p)$ still captures the amount of information that we can transmit. How can we reconciliate the two?

Note that some horses are more likely to win than others, then we can reencode the message to minimize the number of bits used. Say for example we use a 0 to inform that horse 1 won, and a 1 to say some other horse won and a second bit to say that horse 2 won or not, etc. The entropy measures the expected number of bits needed.

Entropy is the average minimal number of bits to communicate the realization of a random variable. We can think of entropy as the "cost" of communicating the realization of a random variable.

In many environments, we may not observe the random variable x, but a signal y.

Let x and y be two discrete random variables with joint distribution $p(x, y)$.

Definition 12 *The joint entropy is defined by*

$$H(X, Y) = -\sum_x \sum_y p(x, y)\log_2 p(x, y)$$

Definition 13 *The conditional entropy is defined by*

$$H(X|y) = -\sum_x p(x|y)\log_2 p(x|y)$$

Definition 14 *The expected conditional entropy is defined by*

$$H(X|Y) = -\sum_y \sum_x p(y)p(x|y)\log_2 p(x|y)$$

Definition 15 *Let $p(x)$ and $q(x)$ be two distribution functions, the relative entropy is given by*

$$D(p||q) = \sum_x p(x) \log_2 \frac{p(x)}{q(x)}$$

The idea is that relative entropy measures the loss in information or increase in cost of communicating the outcome of p through q.

Note that $D(p||q) \geq 0$ and equals 0 if and only if $p(x) = q(x)\ \forall x$. Also note that $D(.||.)$ is not symmetric and does not satisfy the triangular inequality.

We now want to define a concept to measure how much information does a random variable carries about another random variable.

Definition 16 *Given two random variables x, y. Their mutual information is given by*

$$I(X,Y) = D(p(x,y)||p(x)p(y))$$

Note that

$$\begin{aligned} I(X,Y) &= H(X) + H(Y) - H(X,Y) \\ &= H(X) - H(X|Y) \\ &= H(Y) - H(Y|X) \\ &= I(Y,X) \end{aligned}$$

since $H(X,Y) = H(X) + H(Y|X) = H(Y) + H(X|Y)$.

Example. Noiseless binary channel. Say we have an input x with two possible realizations 0,1, and an outcome y, with two realizations 0,1, and

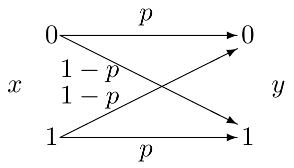
$$\begin{array}{ccc} x & & y \\ 0 & \longrightarrow & 0 \\ 1 & \longrightarrow & 1 \end{array}$$

The information is noiseless because observing the output we can infer the input without error. The channel is giving us the information $p(y|x)$. The capacity of the information channel is given by

$$C = \max_{p(x)} I(x,y)$$

that is, C is the number of bits we need to transfer information without error. In this example, $C = 1$.

Example. Noisy system, binary channel.



In this example $C > 1$.

8.3 Slow Decision Making

The first models of slow decision making were those of Simon, Calvo and Mankiw. Those models introduced a physical cost of change. In particular, the sticky prices models of Calvo and Mankiw assume that there is a ficed cost to change prices. Hence, we will observe infrequent price changes, and when they change they change by a big amount. The fixed cost leads to the well known (s, S) policy.

Instead of physical costs, Sims (2003), Woodford (2007), Reis (2006a), and Reis (2006b) consider information costs. In Woodford (2007) agents cannot process all the information and hence, they face a marginal price for information. Reis considers a fixed cost in the sense that if the information changes from the last period, it does not matter if it changes just a bit or a lot, to process it again the agents have to pay a fixed cost of processing.

8.3.1 Sims (2003)

Consider an agent choosing consumption over two periods. With complete information, the problem is simply

$$\max_{0\le c\le w} u(c) + u(w-c)$$

where $u(c) = \log c$.

We now add information constraints. Let $G(w)$ be the true distribution of wealth, and $g(w)$ its corresponding density. In the first period the agent still doesn't know the realization of wealth, information about wealth only arrives in the second period. Now the problem is

$$\max_{f(c,w)} \int_{0\le c\le w} \log(c(w-c)) f(c,w) dc dw$$

subject to

$$\begin{array}{ll} (1) & f(c,w) \ge 0 \\ (2) & \int_{0\le c\le w} f(c,w) dc = g(w) \quad \forall w \\ (3) & \int_{0\le c\le w} (\log f(c,w)) f(c,w) dw dc \\ & \qquad - \int_0^\infty (\log(\int_c^\infty f(c,w)dw) \int_c^\infty f(c,w) dc dw) dc \\ & \qquad - \int_0^\infty \log g(w) g(w) dw \le K \end{array}$$

In other words, constraint (3) says $I(w,c) = -H(w,c) + H(c) + H(w) \le K$.

Note that this is a reduced form model where c is the signal and the optimal decision.

Let the Lagrange multipliers associated with constraints (1)-(3) be $\psi(c,w)$, $\mu(w)$, and λ, respectively.

We can solve the problem by pointwise optimization. Fix (c,w), $c \le w$, differentiate with respect to $f(.,.)$ to get the FOC (assuming an interior solution)

$$\log c(w-c) = \psi(c,w) + \mu(w) + \lambda \left(1 + \log f(c,w) - \log \left(\int_c^\infty f(c,w) dw \right) - 1 \right).$$

Now, let $q(w|c) \equiv \frac{f(c,w)}{\int_c^\infty f(c,w)dw}$, $\alpha \equiv 1/\lambda$ and $v(w) \equiv e^{-\alpha\mu(w)}$, then

$$\log c^\alpha (w-c)^\alpha - \log q(w|c) + \frac{\mu(w)}{\lambda} + \frac{\psi(w,c)}{\lambda} = 0$$

Setting $\psi(w,c) = 0$ we get

$$\log c^\alpha (w-c)^\alpha = \log q(w|c) + \frac{\mu(w)}{\lambda}$$

or

$$q(w|c) = v(w)c^{\alpha}(w-c)^{\alpha}$$

Since $q(w|c)$ is a conditional distribution, regardless of c, we must have

$$\int v(w)c^{\alpha}(w-c)^{\alpha}dw = 1$$

but this requires,

$$v(w) \propto w^{-2\alpha-1}$$

8.3.2 Woodford (2007)

The idea in this paper is that the state of the world (a random variable) x determines the optimal price $q^* \in \arg\max V(q)$ (optimal price with full information). Price revision is costly. The decision maker wants to know if it is the time to revise the price or not.

Let K be the cost of conducting a review of the pricing policy, and let $V(q)$ be a smooth, strictly concave function. Let x denote the price gap, $x = q(i) - q^*$. Define the loss of failing to review the policy

$$L(x) \equiv V(q^*) - V(q^* + x) - K$$

In the case of full information, the optimal price-review policy is to review the price if and only if the value of x prior to the review is in the range such that $L(x) \geq 0$ (i.e., an Ss policy). The problem is that we don't know what the value of x is.

Recall that entropy is defined as $-\int f(x)\log f(x)dx$. Hence, the entropy reduction when signal s is received is given by

$$I(s) = \int \hat{f}(x|s)\log f(x|s)dx - \int f(x)\log f(x)dx$$

and let $I \equiv \mathrm{E}_s I(s)$.

Now, the theory of rational inattention posits that both the design of this signal and the decision about whether to conduct a price review conditional upon the signal observed will be optimal, in the sense of maximizing

$$\bar{L} = \mathrm{E}[\delta(s)L(x)] - \theta I$$

where, θ is the cost per unit of information and

$$\delta(s)\begin{cases} 0 & \text{price review not undertaken} \\ 1 & \text{price review is undertaken} \end{cases}$$

The optimal solution will be

Lemma 4 *Binary signal.*

This leads to a hazard function

$$\Lambda(x) \propto \pi(1|x)$$

indicating the probability of a price review occurring given that the state is x.

With two possible signals,

$$\hat{f}(x|0) = \frac{f(x)(1-\Lambda(x))}{1-\bar{\Lambda}}$$
$$\hat{f}(x|1) = \frac{f(x)\Lambda(x)}{\bar{\Lambda}}$$

where $\bar{\Lambda} = \int \Lambda(x)f(x)dx$. The information measure I is then equal to

$$\begin{aligned} I &= \bar{\Lambda}I(1) + (1-\bar{\Lambda})I(0) \\ &= \bar{\Lambda}\int \hat{f}(x|1)\log\hat{f}(x|1)dx + (1-\bar{\Lambda})\int \hat{f}(x|0)\log\hat{f}(x|0)dx - \int f(x)\log f(x)dx \end{aligned}$$

8.3.3 Reis (2006)

We have a monopolist providing a quantity y with a cost $C(y, s_t)$, where s_t is the state of the world. Let the demand faced by the monopolist be $Q(p, s_t)$, and let s_t be a stochastic process (a Brownian motion, with Markov property).

At any point in time, the firm can choose to observe (or not) the state. Observing the state is costly. Let $K \geq 0$ be the cost of observing s_t. If the firm chooses not to observe the state, it has its last observation. In addition, let the firm know $p(s_{t+1}|s_t)$.

Let t_d be the last point in time the firm chose to observe the state. Then, the flow payoff at time t given the information s_{t_d} is

$$\pi^*(s_{t_d}, t - t_d) = \max_{p_t} \mathrm{E}\Big[p_t Q(p_t, s_t) - C(Q(p_t, s_t), s_t)\Big| s_{t_d}, t - t_d\Big]$$

Assume continuous time, i.e., $t \in [0, \infty)$, and let the discount factor be $r > 0$. Let i index observations of the state variable, and let $d(i)$ denote the time at which the $i+1$-th observation occurs. Then, the firm's problem is

$$\max_{\{d(i)\}_{i=1}^{\infty}} \sum_{i=1}^{\infty} \left[\int_{d(i)}^{d(i+1)} e^{-rt} \pi^*(s_{d(i)}, t - d(i)) dt - e^{-rd(i)} K \right]$$

Note that this problem has a recursive structure, hence we can rewrite it as

$$V(s) = \mathrm{E}\left[\max_d \left\{ \int_0^d e^{-rt} \pi^*(s, t) dt + e^{-rd}[-K + V(s_d)] \right\} \right]$$

subject to

$$s_d = \psi(s, u_d)$$

where u_d are the increments in s from time 0 to d.

Taking the FOC we have

$$e^{-rd}\pi(s, d) - re^{-rd}[-K + V(s_d)] + e^{-rd} V'(s_d) \frac{\partial \psi}{\partial d} = 0$$

or

$$\pi(s, d) + rK = r\mathrm{E}V(s_d) + \mathrm{E}V'(s_d) \frac{\partial \psi}{\partial d}$$

note that the right hand side are the gains from postponing observing the state, and the left hand side, the costs.

Define $G(s, d) : R^{s+1} \to R$ as the expected differences between profits earned with full information and profits earned with the optimal plan.

Proposition 8 *A perturbation around the situation when planning is costless is*

$$d^*(s) = \sqrt{\frac{2K(s)}{G(s, 0)}}$$

Note that second order costs of planning lead to first-order long inattentiveness

Proposition 9 *With iso-elastic demand (with price elasticity $\epsilon > 1$) subject to i.i.d. multiplicative shocks, while marginal costs follow a geometric Brownian motion with variance σ^2, and planning costs a fixed share K of profits, the producer sets a plan for prices $p_t = (\frac{\theta}{\theta - .1} E[s_t]$.*

In the vicinity of $K = 0$, optimal inattentiveness approximately equals

$$d^*(s) = \sqrt{\frac{4K(s)}{\sigma^2 \theta(\theta - 1)}}$$

Note that the price rule does not depend on the state of demand.

Aggregation. Let's consider the case where inattentiveness varies randomly with changes in the profits of firms and the costs of planning. The arrival of decision dates then takes the form of a stochastic point process. Denote the arrival densities by $f_i(t)$ and assume they are mutually independent; independent across producers; and the same for all producers.

Let ρ denote the intensity of attention, defined as the long-run mean number of planning dates in a unit of time: $\rho = \frac{1}{\mathrm{E}[d(i)]}$ as $t \to \infty$.

Proposition 10 *The only stationary equilibrium distribution of inattentiveness is the exponential distribution with parameter* ρ*.*

References

Athey, S., and I. Segal (2007): "Designing Efficient Mechanisms for Dynamic Bilateral Trading Games," *American Economic Review Papers and Proceedings*, 97, 131–136.

Bergemann, D., and S. Morris (2005): "Robust Mechanism Design," *Econometrica*, 73, 1771–1813.

Bergemann, D., and J. Välimäki (2006): "Dynamic Price Competition," *Journal of Economic Theory*, 127, 232–263.

Bergemann, D., and J. Valimaki (2007): "Dynamic Marginal Contribution Mechanism," Discussion paper, Yale University and Helsinki School of Economics.

Brandenburger, A., and E. Dekel (1993): "Hierarchies of Belief and Common Knowledge," *Journal of Economic Theory*, 59, 189–198.

Clarke, E. (1971): "Multipart Pricing of Public Goods," *Public Choice*, 8, 19–33.

Dasgupta, P., P. Hammond, and E. Maskin (1979): "The Implementation of Social Choice Rules. Some General Results on Incentive Compatibility.," *Review of Economic Studies*, 66, 185–216.

Dolan, R. (1978): "Incentive Mechanisms for Prioroty Queuing Problems," *Bell Journal of Economics*, 9, 421–436.

Edelman, B., M. Ostrovsky, and M. Schwartz (2007): "Internet Advertising and the Generalized Second Price Auction: Selling Billions of Dollars Worth of Keywords," *American Economic Review*, 97, 242–259.

Ely, J. C., and M. Peski (2006): "Hierarchies of Belief and Interim Rationalizability," *Theoretical Economics*, 1.

Eso, P., and B. Szentes (2007): "Optimal Information Disclosure in Auctions," *Review of Economic Studies*, forthcoming.

Feinberg, Y., and A. Skrypacz (2005): "Uncertainty About Uncertainty and Delay in Bargaining," *Econometrica*, 73, 69–91.

Freixas, X., R. Guesnerie, and J. Tirole (1985): "Planning under Incomplete Information and the Ratchet Effect," *Review of Economic Studies*, 52, 173–191.

Green, J., and J. Laffont (1977): "Characterization of Satisfactory Mechanisms for the Revelation of the Preferences for Public Goods," *Econometrica*, 45, 427–438.

Groves, T. (1973): "Incentives in teams," *Econometrica*, 41, 617–631.

Gul, F., H. Sonnenschein, and R. Wilson (1986): "Foundations of Dynamic Monopoly and the Coase Conjecture," *Journal of Economic Theory*, 39, 155–190.

Harsanyi, J. (1967-68): "Games with Incomplete Information Played by 'Bayesian' Players," *Management Science*, 14, 159–189, 320–334, 485–502.

Hart, O., and J. Tirole (1988): "Contract Renogiation and Coasian Dynamics," *Review of Economic Studies*, 55, 509–540.

Heifetz, A., and D. Samet (1888): “Topology-Free Typology of Beliefs,” *Journal of Economic Theory*, 82, 324–341.

Ledyard, J. O. (1978): “Incentive Compatibility and Incomplete Information,” *Journal of Economic Theory*, 18, 171–189.

Mertens, J., and S. Zamir (1985): “Formalization of Bayesian Analysis for Games with Incomplete Information,” *International Journal of Game Theory*, 14, 1–29.

Reis, R. (2006a): “Inattentive Consumers,” *Journal of Monetary Economics*, 53, 1761–1800.

——— (2006b): “Inattentive Producers,” *Review of Economic Studies*, 73, 793–821.

Schmidt, K. (1990): “Commitment Through Incomplete Information in a Simple Repeated Bargaining Model,” mimeo, University of Bonn.

Sims, C. (2003): “Implications of Rational Inattention,” *Journal of Monetary Economics*, 50, 665–690.

——— (2006): “Rational Inattention: Beyond the Linear-Quadratic Case,” *American Economic Review*, 96, 158–163.

Vickrey, W. (1961): “Counterspeculation, Auctions and Competitive Sealed Tenders,” *Journal of Finance*, 16, 8–37.

Wilson, R. (1987): “Game-Theoretic Analyses of Trading Processes,” in *Advances in Economic Theory: Fifth World Congress*, ed. by T. Bewley, pp. 33–70, Cambridge. Cambridge University Press.

Woodford, M. (2007): “Information-Constrained State-Dependent Pricing,” Discussion paper, Columbia University.