THEoretical Studies on IncentiveS 2008-2009

Pierre BOYER

Toulouse School of Economics (GREMAQ) and EHESS

January 13, 2009

Today

Glazer, J. and A. Rubinstein (2004),
On Optimal Rules of Persuasion, Econometrica.

Play before reading: http://gametheory.tau.ac.il/exp5

Persuasion game

- One agent (the speaker) wishes to persuade another agent (the listener) to take a certain action. Whether or not the listener should take the action is dependent on information possessed by the speaker.
- The listener can obtain bits of relevant information but is restricted as to the total amount of evidence he can accumulate.
- The speaker can use only verbal statements to persuade.
- Whether or not the listener is persuaded is dependent on both the speaker's arguments and the hard evidence the listener has acquired.

Persuasion game

- One agent (the speaker) wishes to persuade another agent (the listener) to take a certain action. Whether or not the listener should take the action is dependent on information possessed by the speaker.
- The listener can obtain bits of relevant information but is restricted as to the total amount of evidence he can accumulate.
- The speaker can use only verbal statements to persuade.
- Whether or not the listener is persuaded is dependent on both the speaker's arguments and the hard evidence the listener has acquired.

Persuasion game

- One agent (the speaker) wishes to persuade another agent (the listener) to take a certain action. Whether or not the listener should take the action is dependent on information possessed by the speaker.
- The listener can obtain bits of relevant information but is restricted as to the total amount of evidence he can accumulate.
- The speaker can use only verbal statements to persuade.
- Whether or not the listener is persuaded is dependent on both the speaker's arguments and the hard evidence the listener has acquired.

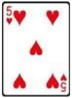
Persuasion game

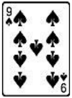
- One agent (the speaker) wishes to persuade another agent (the listener) to take a certain action. Whether or not the listener should take the action is dependent on information possessed by the speaker.
- The listener can obtain bits of relevant information but is restricted as to the total amount of evidence he can accumulate.
- The speaker can use only verbal statements to persuade.
- Whether or not the listener is persuaded is dependent on both the speaker's arguments and the hard evidence the listener has acquired.

Instructions

At the beginning of each round the computer will randomly select for you two cards, red and black. Each card will carry a value, a number between 1 and 9 for **red** and 5 and 13 for **black**. You have a **GOOD** hand if the sum of the values is at least **15**. You have a **BAD** hand if the sum of the values is 14 or below.

For Example:





Is a **bad** hand

Independently of your hand (even if you hold a bad hand), your task is to convince the other player that you hold a good hand.
The task of the other player is to guess correctly whether your hand is good or bad.

After observing the cards, you will have to send the other player a message of the form:

I have a good hand:

The number of the **black** card is:	9
The number of the **red** card is:	6

where you will have the fill each blank with a number between 1 and 9 (red) or 5 and 13 (black).

The other player will then be able to observe at most **one** of the two cards. His decision what card to observe might depend on your message.

Based on your message and his observation, the other player will make a guess

or

You will be informed about the other player's guess.

Please read the instructions again and once you are ready press START

You have completed the game

Statistics:
In the last **20** rounds,

Your hand was good **6** times,

and you succeeded in persuading the other player that your hand is good **11** times

In an optimal play of the game you would have persuaded the other player that your hand is good **11** times

The total number of points you have earned is: **11**

100% of the time your action was optimal response to the machine's strategy

Examples

Example A: Worker

- A worker wishes to be hired by an employer for a certain position.
- The worker tells the employer about his previous experience in two similar jobs.
- The employer wishes to hire the worker if his average performance in the two previous jobs was above a certain minimal level.
- However, before making the final decision the employer has sufficient time to thoroughly interview at most one of the candidate's previous employers.

Examples

Example A: Worker

- A worker wishes to be hired by an employer for a certain position.
- The worker tells the employer about his previous experience in two similar jobs.
- The employer wishes to hire the worker if his average performance in the two previous jobs was above a certain minimal level.
- However, before making the final decision the employer has sufficient time to thoroughly interview at most one of the candidate's previous employers.

Examples

Example A: Worker

- A worker wishes to be hired by an employer for a certain position.
- The worker tells the employer about his previous experience in two similar jobs.
- The employer wishes to hire the worker if his average performance in the two previous jobs was above a certain minimal level.
- However, before making the final decision the employer has sufficient time to thoroughly interview at most one of the candidate's previous employers.

Examples

Example A: Worker

- A worker wishes to be hired by an employer for a certain position.
- The worker tells the employer about his previous experience in two similar jobs.
- The employer wishes to hire the worker if his average performance in the two previous jobs was above a certain minimal level.
- However, before making the final decision the employer has sufficient time to thoroughly interview at most one of the candidate's previous employers.

Examples con't

Example B: Suspect

- A suspect is arrested on the basis of testimonies provided by two witnesses.
- The suspect's lawyer claims that their testimonies to the police have serious inconsistencies and therefore his client should be released.
- The judge's preferred decision rule is to release the suspect only if the two testimonies substantially contradict one another
- However, he is able to investigate at most one of the two witnesses

Examples con't

Example B: Suspect

- A suspect is arrested on the basis of testimonies provided by two witnesses.
- The suspect's lawyer claims that their testimonies to the police have serious inconsistencies and therefore his client should be released.
- The judge's preferred decision rule is to release the suspect only if the two testimonies substantially contradict one another
- However, he is able to investigate at most one of the two witnesses

Examples con't

Example B: Suspect

- A suspect is arrested on the basis of testimonies provided by two witnesses.
- The suspect's lawyer claims that their testimonies to the police have serious inconsistencies and therefore his client should be released.
- The judge's preferred decision rule is to release the suspect only if the two testimonies substantially contradict one another
- However, he is able to investigate at most one of the two witnesses

Examples con't

Example B: Suspect

- A suspect is arrested on the basis of testimonies provided by two witnesses.
- The suspect's lawyer claims that their testimonies to the police have serious inconsistencies and therefore his client should be released.
- The judge's preferred decision rule is to release the suspect only if the two testimonies substantially contradict one another
- However, he is able to investigate at most one of the two witnesses

Aim of the paper

What are the properties of the mechanisms that are optimal from the point of view of the listener?

Main elements of the Model

- 2 agents: a speaker and a listener
- the listener chooses between two actions *a* (accept) and *r* (reject)
- the speaker's type is the realization of two aspects (private information)
- the listener can verify at most one of the aspects

Conflict

- The speaker always prefers action a whatever his true type.
- The listener's preferred action depends on the speaker's type

The listener does not have tools to deter the speaker from cheating
$\Rightarrow$ expect that the speaker argues that his information indicates that a should be taken

Problem: Which rules the listener should follow in order to minimize the probability of making a mistake.

Mechanism

The mechanism is composed of three elements:

* A set of messages from which the speaker can choose
* A function that specifies which aspect is to be checked depending on the speaker's message
* The action the listener finally takes as a function of the message sent by the speaker and the acquired information

A- Deterministic mechanisms

* For each of the two aspects certain criteria are determined
* the speaker's preferred action is chosen if he can show that his type meets these prespecified criteria in at least one of the two aspects

Example A: Worker

A deterministic mechanism in this example would be equivalent to asking the worker to provide a reference from one of his two previous employers that meets certain criteria

B- Random mechanisms

* the speaker is asked to report his type
* one aspect is then chosen randomly and checked
* the action a is taken if and only if the speaker's report is not refute

Example A: Worker

A random mechanism in this example would involve first asking the worker to justify his application by reporting his performance in each of his previous two jobs. Based on his report, the employer then randomly selects one of the two previous employers to interview and accepts the applicant if his report is not refuted.

Results

(i) Finding an optimal mechanism can be done by solving an auxiliary linear programming problem with objective "minimize the probability of a mistake" (constraints are derived from L-principle)

(ii) An optimal mechanism with a very simple structure always exists

Results

(iii) The optimal mechanism is credible

(iv) When all types are equally likely, certain "convexity" and "monotonicity" conditions are identified under which there exists an optimal mechanism that is deterministic.

The Model

- $\{1, ..., n\}$ set of random variables called aspects
  most of the analysis with $n = 2$

- realization of aspect k is a member of a set X_k

- A problem is (X, A, p) where
  $\emptyset \neq A \subset X = \times_{k=1,...,n} X_k$
  p probability measure on X: denote $p_x = p(\{x\})$ the probability of type x
  If X is infinite p_x density function
  Assume $p_x > 0$ for all x

- A problem is finite if the set X is finite

- Two agents the *speaker* and *listener*

- A member of X called a speaker's type
- the speaker whatever is type prefers the listener to take a
- the speaker knows his type and the listener only the distribution

- the listener has to take one of two actions: a (accept) or r (reject)
- the listener is interested in taking a if speaker's type is in A
- the listener is interested in taking r if speaker's type is in $R = X - A$
- the listener can *check*, find out the realization, of at most *one* of the n aspects

Mechanism

A mechanism is (M, f) where M set of messages and $f : M \rightarrow Q$ with

- Q the set of all lotteries $< \pi_0, d_0; \pi_1, d_1; ...; \pi_n, d_n >$ where $(\pi_i)_{i=1,..,n}$ is a probability vector and $d_k : X_k \rightarrow \{a, r\}$ where $X_0 = \{e\}$ an arbitrary singleton (that is d_0 is a constant)
- An element of Q interpreted as a possible response of the listener to a message
- with probability π_0 no aspect is checked and the action $d_0 \in \{a, r\}$ is taken
- with probability π_k(k=1,...,n) aspect k is checked and if its realization is x_k the action $d_k(x_k)$ is taken
- the choice of the set Q captures the assumptions that the listener can check at most one aspect and that the aspect to be checked can be selected randomly

- A direct mechanism is one where $M = X$
- For a direct mechanism (X, f) we say that following a message m the mechanism verifies aspect k with probability π_k
  when $f(m) =< \pi_0, d_0; \pi_1, d_l; ...; \pi_n, d_n >$ is such that $d_k(x_k) = a$ iff $x_k = m_k$.

- The fair random mechanism is the direct mechanism according to which, for every $m \in A$, the listener verifies each aspect with probability $1/n$ and for every $m \in R$, he chooses the action r.
- A mechanism is deterministic if for every $m \in M$ the lottery $f(m)$ is degenerate (that is, for some k, $\pi_k = 1$).

- For every lottery $q =< \pi_0, d_0; \pi_1, d_l; ...; \pi_n, d_n >$ and every type x define $q(x)$ to be the probability that the action a is taken when the lottery q is applied to type x, that is, $q(x) = \sum_{\{k|d_k(x_k)=a\}} \pi_k$
- We assume that given a mechanism (M,f) a speaker of type x will choose a message that maximizes the probability that the action a is taken, namely he chooses a message $m \in M$ that maximizes $f(m)(x)$.
- Let μ_x be the probability that the listener takes the wrong action with respect to type x, assuming the speaker's behavior.
- For $x \in R$ we have $\mu_x = \max_{m \in M} f(m)(x)$ and for $x \in A$ we have $\mu_x = 1 - \max_{m \in M} f(m)(x)$
- We will refer to $(\mu_x)_{x \in X}$ as the vector of mistakes induced by the mechanism.
- The mistake probability induced by the mechanism is $\int_{x \in X} p_x \mu_x$.

- The mechanisms are evaluated according to the listener's interests while ignoring those of the speaker.
- Assume that the listener's loss given a mechanism is the mistake probability induced by the mechanism.
- Thus, given a problem (X, A, p), an optimal mechanism is a mechanism that minimizes the mistake probability.
- No restriction to direct mechanisms (the revelation principle not applied).

Example 1

- Let $X_1 = X_2 = [0, 1]$
- Let $A = \{(x_1, x_2) \mid x_1 + x_2 \geq 1\}$
- Let p be the uniform distribution

- If the listener chooses to ignore the speaker's message, the lowest probability of a mistake he can obtain is $1/4$.
- This mistake probability can be achieved by a mechanism in which aspect 1 is checked with probability 1 and action a is taken iff aspect 1's value is at least $1/2$ (formally, $M = \{e\}$ and $f(e)$ is the degenerate lottery where $\pi_l = 1$ and $d_l(x_l) = a$ iff $x_l \geq 1/2$).
- In this example, letting the speaker talk can improve matters.

Example 1

- Consider the following deterministic direct mechanism $(M = X)$ characterized by two numbers z_1 and z_2.
- Following the receipt of a message (m_l, m_2), the speaker verifies the value of aspect 1 if $m_1 \geq z_1$ and verifies the value of aspect 2 if $m_1 < z_1$ but $m_2 \geq z_2$.
- If $m_k < z_k$ for both k the action r is taken.
- One interpretation of this mechanism is that in order to persuade the listener, the speaker has to show that the realization of at least one of the aspects is above some threshold (which may be different for each aspect).

Example 1

- The set of types for which the listener's action will be wrong consists of the three shaded triangles shown in Figure 1A.
- One can see that the optimal thresholds are $z_l = z_2 = 2/3$ yielding a mistake probability of $1/6$.

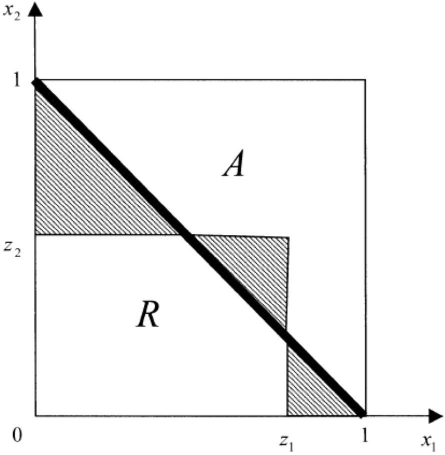


FIGURE 1A.

The L-Principle

Assume $n = 2$ and denote μ_{ij} for $\mu_{(i,j)}$.

Proposition 0: The L-Principle

Let (X, A, p) be a problem. For any mechanism and for any three types $(i, j) \in A$, $(i, s) \in R$, and $(t, j) \in R$, it must be that $\mu_{ij} + \mu_{is} + \mu_{tj} \geq 1$.

This prop. is valid for both finite and infinite X and for any p.

Example A: Worker

- If the worker's performances in each job is classified as good or bad and that the employer wishes to hire the worker only if his performance in both previous jobs was good.
- Consider the worker's three types: his performance was good in two previous jobs, good only in the first job and good only in the second job.
- The L-principle says that for any mechanism, the sum of the probabilities of a mistake conditional on each of the three worker's types is at least one.

Example A: Worker

- If the worker's performances in each job is classified as good or bad and that the employer wishes to hire the worker only if his performance in both previous jobs was good.
- Consider the worker's three types: his performance was good in two previous jobs, good only in the first job and good only in the second job.
- The L-principle says that for any mechanism, the sum of the probabilities of a mistake conditional on each of the three worker's types is at least one.

Example A: Worker

- If the worker's performances in each job is classified as good or bad and that the employer wishes to hire the worker only if his performance in both previous jobs was good.
- Consider the worker's three types: his performance was good in two previous jobs, good only in the first job and good only in the second job.
- The L-principle says that for any mechanism, the sum of the probabilities of a mistake conditional on each of the three worker's types is at least one.

Proof (1/2)

- Let (M, f) be a mechanism.
- Let m be a message optimal for type (i, j)
- Let $f(m) =< \pi_0, d_O; \pi_1, d_1; \pi_2, d_2 >$.
- For a proposition e, let δ_e be 1 if e is true and 0 if e is false. Then

$$\mu_{ij} = \pi_0 \delta_{d_0 = r} + \pi_1 \delta_{d_1(i) = r} + \pi_2 \delta_{d_2(j) = r}$$

- If type (i, s) sends the message m ("claims" that he is (i, j)) the action a will be taken with probability

$$\pi_0 \delta_{d_0 = a} + \pi_1 \delta_{d_1(i) = a}$$

  and therefore

$$\mu_{is} \geq \pi_0 \delta_{d_0 = a} + \pi_1 \delta_{d_1(i) = a}$$

Proof (2/2)

- Similarly,

$$\mu_{tj} \geq \pi_0 \delta_{d_0=a} + \pi_2 \delta_{d_2(j)=a}$$

- Therefore

$$\mu_{ij} + \mu_{is} + \mu_{tj} \geq \pi_0 \delta_{d_0=r} + \pi_1 \delta_{d_1(i)=r} + \pi_2 \delta_{d_2(j)=r}$$

$$+\pi_0 \delta_{d_0=a} + \pi_1 \delta_{d_1(i)=a} + \pi_0 \delta_{d_0=a} + \pi_2 \delta_{d_2(j)=r}$$

$$= 1 + \pi_0 \delta_{d_0=a} \geq 1$$

Q.E.D.

Intuition of the Proof

Whatever is the outcome of the randomization following a message m sent by type $(i,j) \in A$,

- Either the mistaken action r is taken,
- Or at least one of the two types (i,s) and (t,j) in R can induce the wrong action a by sending m.

- Define an L to be any set of three types $(i,j) \in A$, $(i,s) \in R$, and $(t,j) \in R$.
- The result of Proposition 0 (the sum of mistakes in every L is at least 1) refered as the L-principle.
- Extension of the L-principle to the case of $n > 2$ done in the paper

Examples

- take p to be uniform
- When X is finite: $\sum_{x \in X} \mu_x$ the number of mistakes.
- An optimal mechanism can be found by using a technique that relies on the L-principle:
  finding a mechanism that induces H mistakes and finding H disjoint L's allows to conclude that this mechanism is optimal , thus yielding a mistake probability of $H / \mid X \mid$.

Example (1/2)

- $n = 3$, $X_k = \{0, 1\}$ for $k = 1, 2, 3$
- $A = \{(x_1, x_2, x_3) \mid \sum_k x_k \geq 2\}$.
- Consider the mechanism where the speaker is asked to name two aspects, the listener checks each of them with probability $1/2$ and takes the action a if the value of the checked aspect is 1.
- This mechanism yields 1.5 mistakes (mistake probability of $3/16$) since only the three types $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$ can each mislead the listener with probability $1/2$.
- To see that this is an optimal mechanism note that the following three inequalities hold:

$$\mu_{(1,1,0)} + \mu_{(1,0,0)} + \mu_{(0,1,0)} \geq 1$$

$$\mu_{(1,0,1)} + \mu_{(1,0,0)} + \mu_{(0,0,1)} \geq 1$$

$$\mu_{(0,1,1)} + \mu_{(0,1,0)} + \mu_{(0,0,1)} \geq 1$$

Example (2/2)

- The minimum of $\sum_{x \in X} \mu_x$ subject to the constraint

$$\sum_{x | x_l + x_2 + x_3 = 2} \mu_x + 2 \sum_{x | x_l + x_2 + x_3 = 1} \mu_x \geq 3$$

  implied by summing up the three inequalities, is attained when $\mu_x = 1/2$ for any $x \in \{(1,0,0), (0,1,0), (0,0,1)\}$ and $\mu_x = 0$ for any other $x \in X$.
- the number of mistakes cannot fall below 1.5.
- The number of mistakes induced by a deterministic mechanism must be an integer for this problem, it is at least 2.
- One optimal mechanism within the set of deterministic mechanisms involves taking the action a iff the speaker can show that either aspect 1 or aspect 2 has the value 1.
- This mechanism induces two mistakes with regard to types $(1,0,0)$ and $(0,1,0)$.

Example B: Suspect (1/2)

- Situation in which a suspect's lawyer claims that the two testimonies brought against his client are inconsistent
- The judge has time to thoroughly investigate only one of them.
- Consider the problem where $X_1 = X_2 = \{1, 2, 3, ..., I\}$ and $A = \{(x_1, x_2) \mid x_1 \neq x_2\}$; p is uniform.
- Intuitively, the knowledge of the value of only one aspect by itself is not useful to the listener.
- The optimal mechanism will be shown to be nondeterministic in this case.
- The fair random mechanism (following a message in A each of the two aspects is verified with probability .5) induces $I/2$ mistakes.
- The minimal number of mistakes is $1/2$.
- The two cases $I = 2$ and $I = 3$ are illustrated in Figures 3A and 3B.
- In the case of I = 2 there is one L.

1	$R1$
$R1$	

FIGURE 3A.

$*$	$*$	$R*$
$*$	$R*$	
$R*$		

FIGURE 3B.

Example B: Suspect (2/2)

- For $I = 3$, the maximal number of disjoint L's is one; however, notice the six starred types-three in A and three in R.
- Each of the starred types in A combined with two of the starred types in R constitutes an L
- Any mechanism induces mistake probabilities $(\mu_x)_{x \in X}$ satisfying:

$$\mu_{1,3} + \mu_{1,1} + \mu_{3,3} \geq 1$$

$$\mu_{1,2} + \mu_{1,1} + \mu_{2,2} \geq 1$$

$$\mu_{2,3} + \mu_{2,2} + \mu_{3,3} \geq 1$$

- which imply that the sum of mistakes with respect to these six elements must be at least 1.5

An equivalent linear programming problem

- For finite problem: finding an optimal mechanism is equivalent to solving an auxiliary linear programming problem

- Let (X, A, p) a finite problem
- Define $P(X, A, p)$ to be the linear programming:

$$\min \sum_{x \in X} p_x \kappa_x$$

  subject to $\kappa_{ij} + \kappa_{is} + \kappa_{tj} \geq 1$ for all $(i, j) \in A$, $(i, s) \in R$ and $(t, j) \in R$ and $0 \leq \kappa_x$ for all $x \in X$
- the solution to $P((X, A, p)$ coincides with the vector of mistake probabilities induced by an optimal mechanism

Proposition 1

Let (X, A, p) be a finite problem and let $(\kappa_x)_{x \in X}$ be a solution to $P(X, A, p)$. Then, there is an optimal mechanism such the vector of mistakes induced by the mechanism is $(\kappa_x)_{x \in X}$

The structure of optimal mechanism (n=2)

Proposition 2

For every finite problem (X, A, p) there exists an optimal mechanism that is direct (i.e. $M = X$) such that:

(a) If $m \in R$ the listener takes the action r whereas if $m \in A$ the listener does one of the following:

(i) takes the action r;

(ii) takes the action r with probability $1/2$ and verifies one aspect with probability $1/2$;

(iii) verifies each aspect with probability $1/2$;

(iv) verifies one aspect with probability 1.

(b) It is optimal for type $x \in A$ to report x and for type $x \in R$ to send a message y such that for one aspect k, $x_k = y_k$

- First, the speaker is asked to report his type.
- If the speaker admits that the action r should be taken, then the listener chooses r.
- If the speaker claims that the action a should be taken, the listener tosses a fair coin where on each of its two sides one of three symbols, r, 1, or 2 appears:
  r=the listener chooses the action r;
  $i = 1, 2$:the listener checks aspect i and takes the action a if and only if the speaker's claim regarding this aspect is confirmed.

- First, the speaker is asked to report his type.
- If the speaker admits that the action r should be taken, then the listener chooses r.
- If the speaker claims that the action *a* should be taken, the listener tosses a fair coin where on each of its two sides one of three symbols, *r*, 1, or 2 appears:
  r=the listener chooses the action r;
  $i = 1, 2$:the listener checks aspect i and takes the action *a* if and only if the speaker's claim regarding this aspect is confirmed.

- First, the speaker is asked to report his type.
- If the speaker admits that the action r should be taken, then the listener chooses r.
- If the speaker claims that the action *a* should be taken, the listener tosses a fair coin where on each of its two sides one of three symbols, *r*, 1, or 2 appears:
  r=the listener chooses the action r;
  $i = 1, 2$:the listener checks aspect i and takes the action *a* if and only if the speaker's claim regarding this aspect is confirmed.

Proof

(a) A proposition due to Alon (2003) states that if $(\alpha_x)_{x\in X}$ is an extreme point of the set of all vectors satisfying the constraints in $P(X,A,p)$, then $\alpha_x \in \{0, 1/2, 1\}$ for all $x \in X$
Let $(\kappa_x)_{x\in X}$ be a solution to $P(X,A,p)$. As a solution to a linear programming problem the vector $(\kappa_x)_{x\in X}$ is an extreme point and thus $(\kappa_x) \in \{0, 1/2, 1\}$ for all $x \in X$. The construction of an optimal mechanism in Prop. 1 implies the rest of our claim since for every $i \in X_1$ and $j \in X_2$ the numbers $\min_{s|i\in R} \kappa_{is}$ and $\min_{t|j\in R} \kappa_{tj}$ are all within $\{0, 1/2, 1\}$

(b) The claim is straightforward for $x \in R$ and for $x \in A$ for which $\kappa_x = 0$. Type $(i,j) \in A$ for whom $\kappa_{ij} > 0$ can possibly obtain a positive probability of acceptance only by "cheating" about at most one aspect. If he claims to be type (t,j), then the probability that the second aspect will be verified is at most $\min_{t|j\in R} \kappa_{tj}$, which is exactly the probability that aspect 2 is verified when the speaker admits he is (i,j).

Also in this paper

- Listener credibility: an optimal mechanism build in Prop 1 and 2 is credible
- Conditions for optimality of deterministic mechanisms