ELICITING AWARENESS:
ITERATED REVELATION MECHANISMS

Evan Piermont
Royal Holloway, University of London

Universitat Pompeu Fabra

Feb 2024

Cast of Characters

Decision Maker (dm)

- Will take an action in the future
- Would like to condition the action on the resolution of uncertainty
- He is unaware/unable to express of some relevant aspects of the decision problem

Expert (ex)

- Has no right to the decision, herself
- But can reveal aspects of the environment

(we will also add a mechanism designer, later)

- Writing legislation: politician; technical advisor
- Buying a house: buyer; surveyor
- Hiring: search committee; letter writer
- **Investing: investor; financial advisor**

Contracts

- $\mathcal{A}$ is a set of actions
- Ω is a state-space
- dm must choose a **contract**:

$$\mathfrak{c} : \Omega \to \mathcal{A}$$

Contracts

- Not all contracts are feasible. dm may be
  - unable to express
  - unaware of
  - technologically unable to implement/condition on

  some actions or events in the state-space

- ex's revelations are
  - verifiable
  - ex-ante uncontactable

  effectively expand the set of feasible contracts

Why is the interesting?

- When preferences are not aligned, ex might strategically conceal her awareness
- Can dm do anything to incentivize revelation?
- A(n unaware) designer may not be able to solve the problem, if mechanisms depend on the unknowns

Literature

- Incomplete Contracting
  - Grossman and Hart (1986); Maskin and Tirole (1999); Aghion and Holden (2011); Hart (2017)
- Contracting under unawareness
  - Tirole (2009); Filiz-Ozbay (2012); Von Thadden and Zhao (2012); Auster (2013) Auster and Pavoni (2021); Piermont (2017); Lei and Zhao (2021); Francetich and Schipper (2021)
- Cheap talk
  - Crawford and Sobel (1982); Seidmann and Winter (1997); Aumann and Hart (2003);
- Robust Mechanism Design
  - Bergemann and Morris (2005); Jehiel et al., (2006); Carroll (2015, 2019).

Example

- An investor (the decision maker) is trying to invest his wealth:
  - the composition of the portfolio can be contingent on the future realized state-of-affairs, but
  - can depend only on those contingencies he is aware of
  - can only invest in assets he is aware of
- He can enlist the help of a financial advisor (the expert) who may reveal novel contingencies/assets

Example

- The true state-space is $\Omega = \{\omega, \nu\}$; equally likely
- Set of actions $\mathcal{A} = \{\alpha, \beta, \gamma\}$
- dm must choose an contract $\mathfrak{c} : \Omega \to \mathcal{A}$
- Let $V_i(\mathfrak{c})$ denote the expected utility to player i

Example

ex can distinguish the states, but dm cannot.

$$\mathscr{P}_e = \{\{\omega\}, \{\nu\}\} \qquad \mathscr{P}_d = \{\{\omega, \nu\}\}$$

$$u_d = \left\{ \begin{array}{c|c|c|c} & \alpha & \beta & \gamma \\ \hline \omega & 4 & 0 & 2 \\ \hline \nu & 0 & 6 & 0 \end{array} \right.$$

$$u_e = \left\{ \begin{array}{c|c|c|c} & \alpha & \beta & \gamma \\ \hline \omega & 0 & 2 & 4 \\ \hline \nu & 0 & 2 & 4 \end{array} \right.$$

How does dm view payoffs in coarse states?

- Assume it is aggregated via expectations
- As if he correctly assesses randomness, but condition a contract on the source of this randomness because he
  - is unaware of what causes it, or
  - does not possess language describe it in a contract, or
  - does not have the technology to condition on it

$$u_d = \left\{ \begin{array}{c|c|c|c} & \alpha & \beta & \gamma \\ \hline \omega & 4 & 0 & 2 \\ \hline \nu & 0 & 6 & 0 \end{array} \right.$$

$$u_e = \left\{ \begin{array}{c|c|c|c} & \alpha & \beta & \gamma \\ \hline \omega & 0 & 2 & 4 \\ \hline \nu & 0 & 2 & 4 \end{array} \right.$$

$$u_d = \left\{ \begin{array}{c|c|c|c} & \alpha & \beta & \gamma \\ \hline \{\omega, \nu\} & 2 & 3 & 1 \end{array} \right.$$

$$u_e = \left\{ \begin{array}{c|c|c|c} & \alpha & \beta & \gamma \\ \hline \{\omega, \nu\} & 0 & 2 & 4 \end{array} \right.$$

$$u_d = \left\{ \begin{array}{c|c|c|c} & \alpha & \beta & \gamma \\ \hline \omega & 4 & 0 & 2 \\ \hline \nu & 0 & 6 & 0 \end{array} \right.$$

$$u_e = \left\{ \begin{array}{c|c|c|c} & \alpha & \beta & \gamma \\ \hline \omega & 0 & 2 & 4 \\ \hline \nu & 0 & 2 & 4 \end{array} \right.$$

What would dm implement:

$$u_d = \left\{ \begin{array}{c|c|c|c} & \alpha & \beta & \gamma \\ \hline \omega & 4 & 0 & 2 \\ \hline \nu & 0 & 6 & 0 \end{array} \right.$$

$$u_e = \left\{ \begin{array}{c|c|c|c} & \alpha & \beta & \gamma \\ \hline \omega & 0 & 2 & 4 \\ \hline \nu & 0 & 2 & 4 \end{array} \right.$$

What would dm implement:

- Without revelation $\mathfrak{c} = \beta$
- $V_d(\mathfrak{c}) = 3,\ V_e(\mathfrak{c}) = 2$

$$u_d = \left\{ \begin{array}{c|c|c|c} & \alpha & \beta & \gamma \\ \hline \omega & 4 & 0 & 2 \\ \hline \nu & 0 & 6 & 0 \end{array} \right. \qquad u_e = \left\{ \begin{array}{c|c|c|c} & \alpha & \beta & \gamma \\ \hline \omega & 0 & 2 & 4 \\ \hline \nu & 0 & 2 & 4 \end{array} \right.$$

What would dm implement:

- With revelation: $\mathfrak{c}' : \begin{cases} \omega \mapsto \alpha \\ \nu \mapsto \beta \end{cases}$
- $V_d(\mathfrak{c}') = 5,\ V_e(\mathfrak{c}') = 1$; So ex won't reveal.

$$u_d = \left\{ \begin{array}{c|c|c|c} & \alpha & \beta & \gamma \\ \hline \omega & 4 & 0 & 2 \\ \hline \nu & 0 & 6 & 0 \end{array} \right.$$

$$u_e = \left\{ \begin{array}{c|c|c|c} & \alpha & \beta & \gamma \\ \hline \omega & 0 & 2 & 4 \\ \hline \nu & 0 & 2 & 4 \end{array} \right.$$

- But, $\mathfrak{c}^\star : \begin{cases} \omega \mapsto \gamma \\ \nu \mapsto \beta \end{cases}$ is a Pareto improvement over no revelation
- $V_d(\mathfrak{c}^\star) = 4,\ V_e(\mathfrak{c}^\star) = 3$

Example

- The Pareto improvement $\mathfrak{c}^{\star}$, requires revelation
- But revealing allows dm to exploit ex
- What if dm could commit:
    - Propose $\mathfrak{c} = \beta$ (his outside option)
    - After ex reveals, propose some other contract $\mathfrak{c}^{\dagger}$
    - $\mathfrak{c}^{\dagger}$ only get implemented if ex agrees; else $\mathfrak{c} = \beta$

Example

Internalizing this, dm solves:

$$\max_{\mathfrak{c}^\dagger:\Omega\to\mathcal{A}} V_d(\mathfrak{c}^\dagger) \qquad \text{subject to} \qquad V_e(\mathfrak{c}^\dagger) \geq V_e(\mathfrak{c}) \qquad \text{(IC)}$$

Example

Internalizing this, dm solves:

$$\max_{\mathfrak{c}^\dagger:\Omega\to\mathcal{A}} V_d(\mathfrak{c}^\dagger) \qquad \text{subject to} \qquad V_e(\mathfrak{c}^\dagger) \geq V_e(\mathfrak{c}) \qquad \text{(IC)}$$

- The solution is $\mathfrak{c}^\star : \begin{cases} \omega \mapsto \gamma \\ \nu \mapsto \beta \end{cases}$

So a two stage game with commitment to not revoke the prior proposal resulted in

- full revelation
- an efficient contract

So a two stage game with commitment to not revoke the prior proposal resulted in

- full revelation
- an efficient contract

Does this always work?

So a two stage game with commitment to not revoke the prior proposal resulted in

- full revelation
- an efficient contract

Does this always work? No

Example

What if dm is initially unaware of action β?

$$u_d = \left\{ \begin{array}{c|c|c|c} & \alpha & \beta & \gamma \\ \hline \omega & 4 & 0 & 2 \\ \hline \nu & 0 & 6 & 0 \end{array} \right.$$

$$u_e = \left\{ \begin{array}{c|c|c|c} & \alpha & \beta & \gamma \\ \hline \omega & 0 & 2 & 4 \\ \hline \nu & 0 & 2 & 4 \end{array} \right.$$

$$u_d = \left\{ \begin{array}{c|c|c|c} & \alpha & & \gamma \\ \hline \{\omega, \nu\} & 2 & & 1 \end{array} \right.$$

$$u_e = \left\{ \begin{array}{c|c|c|c} & \alpha & & \gamma \\ \hline \{\omega, \nu\} & 0 & & 4 \end{array} \right.$$

$$u_d = \left\{ \begin{array}{c|c|c|c} & \alpha & \beta & \gamma \\ \hline \omega & 4 & 0 & 2 \\ \hline \nu & 0 & 6 & 0 \end{array} \right.$$

$$u_e = \left\{ \begin{array}{c|c|c|c} & \alpha & \beta & \gamma \\ \hline \omega & 0 & 2 & 4 \\ \hline \nu & 0 & 2 & 4 \end{array} \right.$$

What would dm implement:

$$u_d = \left\{ \begin{array}{c|c|c|c} & \alpha & \beta & \gamma \\ \hline \omega & 4 & 0 & 2 \\ \hline \nu & 0 & 6 & 0 \end{array} \right.$$

$$u_e = \left\{ \begin{array}{c|c|c|c} & \alpha & \beta & \gamma \\ \hline \omega & 0 & 2 & 4 \\ \hline \nu & 0 & 2 & 4 \end{array} \right.$$

What would dm implement:

- Without revelation $\mathfrak{c}^* = \alpha$
- $V_d(\mathfrak{c}^*) = 2,\ V_e(\mathfrak{c}^*) = 0$

$$u_d = \left\{ \begin{array}{c|c|c|c} & \alpha & \beta & \gamma \\ \hline \omega & 4 & 0 & 2 \\ \hline \nu & 0 & 6 & 0 \end{array} \right.$$

$$u_e = \left\{ \begin{array}{c|c|c|c} & \alpha & \beta & \gamma \\ \hline \omega & 0 & 2 & 4 \\ \hline \nu & 0 & 2 & 4 \end{array} \right.$$

What would dm implement:

- Under full revelation (same as before): $\mathfrak{c}' : \begin{cases} \omega \mapsto a \\ \nu \mapsto \beta \end{cases}$
- $V_d(\mathfrak{c}') = 5$, $V_e(\mathfrak{c}') = 1$; this satisfies the incentive constraint.

$$u_d = \left\{ \begin{array}{c|c|c|c} & \alpha & \beta & \gamma \\ \hline \omega & 4 & 0 & 2 \\ \hline \nu & 0 & 6 & 0 \end{array} \right. \qquad u_e = \left\{ \begin{array}{c|c|c|c} & \alpha & \beta & \gamma \\ \hline \omega & 0 & 2 & 4 \\ \hline \nu & 0 & 2 & 4 \end{array} \right.$$

- But, revealing only β leads to $\mathfrak{c} = \beta$
- $V_d(\mathfrak{c}) = 3,\ V_e(\mathfrak{c}) = 2$; partial revelation is preferred

$$u_d = \left\{ \begin{array}{c|c|c|c} & \alpha & \beta & \gamma \\ \hline \omega & 4 & 0 & 2 \\ \hline \nu & 0 & 6 & 0 \end{array} \right.$$

$$u_e = \left\{ \begin{array}{c|c|c|c} & \alpha & \beta & \gamma \\ \hline \omega & 0 & 2 & 4 \\ \hline \nu & 0 & 2 & 4 \end{array} \right.$$

- As before, $\mathfrak{c}^\star : \begin{cases} \omega \mapsto \gamma \\ \nu \mapsto \beta \end{cases}$ is a Pareto improvement over $\mathfrak{c} = \beta$

What if the procedure was repeated?

(1) Initially, dm commits (to $\mathfrak{c}^* = \alpha$)

(1) Initially, dm commits (to $\mathfrak{c}^* = \alpha$)

(2) ex partially reveals, (the action β)

(1) Initially, dm commits (to $\mathfrak{c}^* = \alpha$)

(2) ex partially reveals, (the action β)

(3) dm proposes a new contract ($\mathfrak{c} = \beta$); ex accepts

(1) Initially, dm commits (to $\mathfrak{c}^* = \alpha$)

(2) ex partially reveals, (the action β)

(3) dm proposes a new contract ($\mathfrak{c} = \beta$); ex accepts

- This solves the IC-constrained maximization problem with constraint $\mathfrak{c}^* = \alpha$

(1) Initially, dm commits (to $\mathfrak{c}^* = \alpha$)

(2) ex partially reveals, (the action β)

(3) dm proposes a new contract ($\mathfrak{c} = \beta$); ex accepts

- This solves the IC-constrained maximization problem with constraint $\mathfrak{c}^* = \alpha$
- $\mathfrak{c} = \beta$ becomes the new IC constraint

(1) Initially, dm commits (to $\mathfrak{c}^* = \alpha$)

(2) ex partially reveals, (the action β)

(3) dm proposes a new contract ($\mathfrak{c} = \beta$); ex accepts

- This solves the IC-constrained maximization problem with constraint $\mathfrak{c}^* = \alpha$
- $\mathfrak{c} = \beta$ becomes the new IC constraint

(4) ex reveals more (the partition $\{\{\omega\}, \{\nu\}\}$)

(1) Initially, dm commits (to $\mathfrak{c}^* = \alpha$)

(2) ex partially reveals, (the action β)

(3) dm proposes a new contract ($\mathfrak{c} = \beta$); ex accepts

- This solves the IC-constrained maximization problem with constraint $\mathfrak{c}^* = \alpha$
- $\mathfrak{c} = \beta$ becomes the new IC constraint

(4) ex reveals more (the partition $\{\{\omega\}, \{\nu\}\}$)

(5) dm solves

$$\max_{\mathfrak{c}^\dagger:\Omega\to\mathcal{A}} V_d(\mathfrak{c}^\dagger) \qquad \text{subject to} \qquad V_e(\mathfrak{c}^\dagger) \geq V_e(\mathfrak{c}) \qquad \text{(IC)}$$

(1) Initially, dm commits (to $\mathfrak{c}^* = \alpha$)

(2) ex partially reveals, (the action β)

(3) dm proposes a new contract ($\mathfrak{c} = \beta$); ex accepts

- This solves the IC-constrained maximization problem with constraint $\mathfrak{c}^* = \alpha$
- $\mathfrak{c} = \beta$ becomes the new IC constraint

(4) ex reveals more (the partition $\{\{\omega\}, \{\nu\}\}$)

(5) dm solves

$$\max_{\mathfrak{c}^\dagger : \Omega \to \mathcal{A}} V_d(\mathfrak{c}^\dagger) \qquad \text{subject to} \qquad V_e(\mathfrak{c}^\dagger) \geq V_e(\mathfrak{c}) \qquad \text{(IC)}$$

- $\mathfrak{c}^\star$ is the solution

Does this always work?

Does this always work? Yes, but what is ‘always’?

Model

The environment is described by

$\mathcal{A}$ — a set of actions

Ω — a state-space

π — a probability over Ω

(u_d, u_e) — state-dependent utility functions

Let V_i denote expectation operator w.r.t u_i and π

Types

- A **type** $t = (\mathscr{P}^t, A^t)$ is described by

$$\mathscr{P}^t \;\text{ — }\; \text{a partition of } \Omega$$
$$A^t \;\text{ — }\; \text{a subset of } \mathcal{A}$$

- Say that type t is **more expressive** than type t', written $t' \preccurlyeq t$, if
  - $\mathscr{P}^t$ refines $\mathscr{P}^{t'}$ and $A^{t'} \subseteq A^t$
- $\mathbf{T}$ is the set of all types and $\mathbf{T}(t)$ those more expressive than t.
- Fix types t_d and t_e, and assume $t_e \in T(t_d)$.

Types

For $t \in \mathbf{T}$, let

$$C(t) = \{\mathfrak{c} : \Omega \to \mathcal{A}^t \mid \mathfrak{c} \text{ is } \mathscr{P}^t \text{ measurable}\}$$

denote those contracts expressible to type t.

- $C(t) \subseteq C(t')$ if and only if $t \preccurlyeq t'$.

Outcome Profiles

An **outcome profile** is a function from types to contracts:

$$
\begin{array}{rccc}
f: & t & \mapsto & \mathfrak{c} \\
 & \pitchfork & & \pitchfork \\
 & \mathbf{T}(t_d) & \rightarrow & \mathcal{A}^{\Omega}
\end{array}
$$

such that $f(t) \in C(t)$.

Call f **incentive compatible** if ex's payoff is monotone in her type

$$V_e(f(h)) \leq V_e(f(h'))$$

whenever $h \preccurlyeq h'$

- There need not be any 'strategic' way of enacting an outcome profile.
- But if there is, it *must* be incentive compatible.

Iterated Revision Mechanisms

An **iterated revelation mechanism** (IRM), is parameterized by a function from *sequences of types* to contracts:

$$\mathfrak{m} : (t_0 \ldots t_n) \mapsto \mathfrak{c} \in C(t_n)$$

STEP 1 — Set $n = 0$. dm announces $t_0 = t_d$, and proposes $\mathfrak{m}(t_0)$.

STEP 2 — ex reveals t_{n+1}.

- If $t_n \prec t_{n+1}$, dm proposes $\mathfrak{m}(t_{n+1})$, goto STEP 3.
- Otherwise, the mechanism is over and $\mathfrak{m}(t_n)$ get implemented.

STEP 3 — ex can accept or reject the proposal, $\mathfrak{m}(t_{n+1})$:

- If she accepts, set $n = n + 1$ and goto STEP 2.
- If she rejects, the mechanism is over and $\mathfrak{m}(t_n)$ get implemented.

Importantly: the contracts proposed by an IRM are *jointly* expressible at the time of proposal

Full Revelation

Theorem

The following are equivalent for an outcome profile f

(1) f can be induced by an IRM

(2) f is a incentive compatible

- Induced: $f(t) = \mathfrak{m}(\sigma)$ where σ is a *best response* over all expressible sequences for type t.

The can be seen as an impossibility result:

- Without commitment to leave proposed contracts on the table, full revelation cannot be guaranteed.

Each proposed contract in an IRM specifies:

(1) The outcome should the game end
   - dm wants to maximize his own payoff

(2) The implicit incentive constraint should the game continue
   - dm wants to minimize ex's payoff

In the examples, contracts solved (1) ignoring (2)

If dm cannot conceive of what ex is aware of it seems prudent to consider *robust* strategies:

- those that maximize the worst case outcome
- this is belief free: does not require conjecturing about probability of types
- Robust strategies turn out to be exactly those that follow the principle of myopic optimization

Robustness

Call an IRM, $\mathfrak{m}$, **robust** if at every sequence of (possibly partial) revelations σ, $\mathfrak{m}$ maximizes the worst case payoff over

- all best responses that extend σ.
- for types for which σ would have been rational
- compared to any other $\mathfrak{m}'$ that coincides with $\mathfrak{m}$ over σ

Robustness

Theorem

The following are equivalent (up to the implemented outcome profile)

(1) $\mathfrak{m}$ is robust

(2) $\mathfrak{m}$ is myopically optimal: at each sequence $(t_0, \ldots, t_n)$,

$$\mathfrak{m}(t_0, \ldots, t_n) \in \underset{\mathfrak{c} \in C(t_n)}{\text{argmax}}\ V_d(\mathfrak{c}) \qquad \text{subject to}$$

$$V_e(\mathfrak{c}) \geq V_e(\mathfrak{m}(t_0 \ldots t_{n-1}))$$

The Designers Problem

- A designer wants the decision maker to take some action
- The designer knows *neither* dm's nor ex's type
- A **mechanism** elicits types and returns an contract

Mechanism

A **mechanism** is a mapping from pairs of types into contracts:

$$\mathcal{M} : (t_d, t_e) \mapsto \mathcal{M}(t_d, t_e)$$

where $\mathcal{M}(t_d, t_e) \in C(h_e)$

- It common knowledge that $t_d \preccurlyeq t_e$

Desiderata:

INDIVIDUAL RATIONALITY: dm can not do better alone (there is no constraint for ex)

INCENTIVE COMPATIBILITY: i prefers to report t_i than any $t \prec t_i$

PARETO OPTIMALITY: there is no feasible contract that dominates the outcome of the mechanism

These are all **ex-post** restrictions — they must hold for all type realizations

Fixing t_d, a mechanism determines an outcome profile:

$$f^{t_d} : t \mapsto \mathcal{M}(t_d, t)$$

By incentive compatibility, f^{t_d} can be induced by an appropriate IRM.

Consider the mechanism, $\mathcal{M}^{\text{MO}}$, that implements a myopically optimal IRM:

(1) first, the decision maker reveals $t \in \mathbf{T}$

(2) then we run a myopically optimal IRM, $\mathfrak{m}^t$:

- starting from t
- multiple m.o. contracts $\Rightarrow$ break ties in favor of the expert

Theorem

The mechanism $\mathcal{M}^{\text{MO}}$

- is individually rational, incentive compatible, and Pareto optimal, and,
- for any other such mechanism $\mathcal{M}$,

$$V_d(\mathcal{M}^{\text{MO}}(t, t')) \geq V_d(\mathcal{M}(t, t'))$$

for all $t, t' \in \mathbf{T}$ with $t \preccurlyeq t'$.

- There is a 'dual' IRM that implements the V_e-dominant mechanism

Distributed Awareness

What if we relax the assumption that $t_d \preccurlyeq t_e$?

Theorem

Allowing for distributed awareness, there exists no incentive compatible and Pareto optimal mechanism.

Let $\Omega = \{\omega\}$ and everything else defined by

	α	β	γ
u_d	0	1	2
u_e	0	2	1

- for $A \subseteq \{\alpha, \beta, \gamma\}$, let t^A denote the type aware of A: $t^A \preccurlyeq t^B$ iff $A \subseteq B$.
- Let $\mathcal{M}$ be any Pareto optimal mechanism. This requires

$$\mathcal{M}(t^{\{\alpha\}}, t^{\{\alpha,\gamma\}}) = \gamma \qquad \mathcal{M}(t^{\{\alpha,\beta\}}, t^{\{\alpha\}}) = \beta \qquad \mathcal{M}(t^{\{\alpha,\beta\}}, t^{\{\alpha,\gamma\}}) \in \{\beta, \gamma\}$$

- if $\mathcal{M}(t^{\{\alpha,\beta\}}, t^{\{\alpha,\gamma\}}) = \beta$, then dm of type $t^{\{\alpha,\beta\}}$ misreports as $t^{\{\alpha\}}$,
- if $\mathcal{M}(t^{\{\alpha,\beta\}}, t^{\{\alpha,\gamma\}}) = \gamma$, then ex of type $t^{\{\alpha,\gamma\}}$ misreports as $t^{\{\alpha\}}$,

Thank You!