# Implementation with Evidence

Navin Kartik & Olivier Tercieux

May 2011

# Introduction

- General goal of implementation theory:

  Design a game form in which agents' strategic behavior leads to desirable outcomes

  - Contracts, Taxation, Elections, Legal systems, Auctions, ...

- The game form specifies
  - messages (or actions) available to players
  - outcomes selected for each profile of messages
  - ...

- Crucially: messages are typically assumed to be cheap talk
  - Available to an agent in all states of the world
  - Are not intrinsically payoff relevant

- **Restrictive and precludes important class of problems**

# Introduction: Motivation

Suppose a decision must be made about how an organization must allocate its scarce resources

- If agents can only send cheap-talk messages, scope for unrestrained manipulation
- But in reality, agents may submit supporting documentation, data, verifiable claims, etc.
  - If an agent can withhold but not falsify

    $\Longrightarrow$ setting with hard/state-contingent evidence
  - If agents can falsify or fabricate at some cost and/or there is a cost of disclosure

    $\Longrightarrow$ setting with costly signaling

# Introduction: Our Contribution

- state-contingent evidence / costly evidence fabrication is introduced into a standard implementation setting (à la Maskin, 1977/99)

Three main issues of interest:

1. given some evidentiary structure, what social objectives can be fully implemented? (This paper concerns complete information)

   We provide a necessary and largely-sufficient condition

2. given a social objective, what minimal evidentiary structure is needed for implementation?

   For hard evidence, derive an appropriate notion of distinguishability

3. relationship between "alienable" and "inalienable" evidence

   Unexpectedly, we find a bridge between the two problems

# Related Literature

- full implementation
  - no evidence: Maskin (1977/99)
  - hard evidence: Ben-Porath and Lipman (2011)
- partial implementation
  - hard evidence: Green and Laffont (1986), Bull and Watson (2007), Deneckere and Severinov (2007), ...
  - costly evidence provision: Bull (2008)
- communication games
  - hard evidence: Milgrom (1981), Lipman and Seppi (1995), ... , Glazer and Rubinstein (2001/4/6), ...
  - costly signaling: Spence (1973), ...
  - costly evidence fabrication: Kartik, Ottaviani, & Squintani (2007), Kartik (2009), Emons and Fluet (2010), ...

# Outline

# Model: Basics

- Finite set of players, $I = \{1, \cdots, n\}$
- Set of outcomes / allocations, $A$ $\quad (|A| > 1)$
- Set of states, $\Theta$ $\quad (|\Theta| > 1)$
- A Social Choice Function (SCF) is $f : \Theta \to A$
  - can extend to correspondences, but notation gets messy

# Model: Evidence Structure

- Each agent $i$ has a set of evidence, $E_i$
  - document, receipt, legal record, verbal proof, collateral, ...
- Preferences over $A \times E_i \times \Theta$ represented by

$$U_i(a, e_i, \theta) := u_i(a, \theta) - c_i(e_i, \theta)$$

  - Assume payoffs are bounded
  - Ordinal vs. vNM preferences
  - Separability is for simplicity
- Assume that for any $i$, $\theta$, there is some least-cost evidence:

$$E_i^\ell(\theta) := \arg\min_{e_i} c_i(e_i, \theta) \neq \emptyset.$$

- Interpretation:
  - At $\theta$, any $e_i \in E_i^\ell(\theta)$ is costless for $i$ (wlog, $c_i(e_i, \theta) = 0$)
  - At $\theta$, any $e_i \notin E_i^\ell(\theta)$ imposes some production or fabrication cost $c_i(e_i, \theta) > 0$

# Model: Evidence Structure

- Important special case is that of hard evidence, where

$$e_i \notin E_i^{\ell}(\theta) \quad \Longrightarrow \quad c_i(e_i, \theta) > \sup_a u_i(a, \theta) - \inf_a u_i(a, \theta).$$

- Instead of prohibitive costs, can also have infeasibility (modulo inessential differences)

- $e_i$ is cheap-talk evidence for agent $i$ if and only if

$$e_i \in \bigcap_{\theta \in \Theta} E_i^{\ell}(\theta)$$

- Special case: the standard setting without evidence

$$\forall i, \forall e_i \in E_i : e_i \text{ is cheap talk}$$

# Model: Mechanisms

- A player's evidentiary choice is "inalienable"
- Thus, a mechanism is a pair $(M, g)$ where
  - $M = M_1 \times \cdots \times M_n$ is the (cheap-talk) message space
  - Let $E := E_1 \times \cdots \times E_n$
  - $g : M \times E \to A$ is an outcome function
- A mechanism induces a strategic-form game in each state $\theta$
  - A pure strategy for player $i$ is $s_i \in M_i \times E_i$
  - For each $(m, e) \in M \times E$, player $i$'s payoff is $U_i(g(m, e), e_i, \theta)$
  - $NE(M, g, \theta)$ is the set of pure strategy Nash equilibria
  - $O(M, g, \theta) := \{a : a = g(m, e) \text{ and } (m, e) \in NE(M, g, \theta)\}$

# Model: Implementation

- A mechanism $(M, g)$ implements a SCF $f$ if
  1. $\forall \theta: \ f(\theta) = O(M, g, \theta)$
  2. $\forall \theta$: if $(m, e) \in NE(M, g, \theta)$, then for all $i$, $e_i \in E_i^{\ell}(\theta)$
- Comments:
  - 1st condition equality is full implementation
    - precludes a revelation principle
  - 2nd condition $\Leftrightarrow$ no costly evidence production in equilibrium
    - wlog in a setting of hard evidence
- A SCF $f$ is implementable if there exists a mechanism that implements it

# Model: Comments

1. Distinguish states from preference profiles

2. A planner can always "ignore" evidence

   $\Longrightarrow$ evidence can only weaken implementation constraints

3. Results extend to mixed NE

4. Without loss of generality:

   - Planner knows evidence structure
   - Every player has some evidence that he may submit
   - Submit exactly *one* piece of evidence
   - Static mechanisms (unless dynamics are used to "change" evidence structure)

# Outline

# An Example

- $N = \{1, 2\}$ (partners)
- $\Theta = \{\theta_1 < \cdots < \theta_K\}$ (joint output)
- $A = \mathbb{R}^2_+$, so $a = (a_1, a_2)$ (each agent's pay)
- Preferences: $u_i(a, \theta) = u_i(a_i)$, str. increasing

REMARK.

*Without any evidence, a SCF $f$ is implementable if and only if it is constant.*

- Equivalently, only constant SCFs are Maskin-monotonic
- Intuition: Full implementation and state-independent prefs

# An Example: Adding Evidence

honesty

- Suppose now that player 1 can provide hard evidence $e_1 \in E_1 = \Theta$
- Assume $E_i^\ell(\theta) = \{\theta_1, \ldots, \theta\}$
  - i.e., can reveal any subset of the true output, but prohibitively costly to fabricate
  - $e_1 = \theta_1$ is the only cheap-talk evidence, can be interpreted as silence
- Player 2 has "no evidence": $E_2^\ell(\theta) = E_2$ for all $\theta$

# An Example: Implementation with Evidence

Define $\mathcal{F} := \{f : range[f] \subseteq \mathbb{R}^2_{++}\}$

**Claim.** Any $f \in \mathcal{F}$ can be implemented with the given evidentiary structure.

**Proof.** Let $M_2 = \Theta$ and use the outcome function

$$g(e_1, m_2) = \begin{cases} f(e_1) & \text{if } e_1 = m_2 \\ (\text{“}+\infty\text{”}, 0) & \text{if } e_1 > m_2 \\ (0,0) & \text{if } e_1 < m_2. \end{cases}$$

Comments:

- Simple and well-behaved “direct” mechanism
- Rationalizability is enough ($\implies$ no bad MSNE)
- Agent 2's cheap-talk message is important (not an unraveling result)

# Outline

# Maskin-Monotonicity: A Review

Suppose there is no evidence, or all evidence is cheap-talk evidence.

## Definition (Maskin-monotonicity)

A SCF $f$ is Maskin-monotonic provided that for all $\theta$ and $\theta'$, if

$$\forall i, a : \left[u_i\left(f(\theta), \theta\right) \geq u_i\left(a, \theta\right) \Rightarrow u_i\left(f(\theta), \theta'\right) \geq u_i\left(a, \theta'\right)\right]$$

then $f(\theta) = f(\theta')$.

- Well-known that this is a stringent requirement
  - Any M-monotonic SCF defined on unrestricted domain of preferences is constant (Saijo, 1987)
  - If preferences are state-independent, a M-monotonic SCF is constant

# Maskin-Monotonicity: A Review

## Theorem (Maskin)

*Without evidence, a SCF is implementable only if it is Maskin-monotonic.*

## Proof.

- Pick a mechanism $(M, g)$ that implements $f$
- Suppose that $\theta$ and $\theta'$ satisfy

$$\forall i, a : \left[u_i\left(f(\theta), \theta\right) \geq u_i\left(a, \theta\right) \Rightarrow u_i\left(f(\theta), \theta'\right) \geq u_i\left(a, \theta'\right)\right]$$

- Pick any $s^* \in NE(M, g, \theta)$

$$\begin{aligned} &\Longrightarrow \quad g(s^*) = f(\theta) \\ &\Longrightarrow \quad s^* \in NE(M, g, \theta') \\ &\Longrightarrow \quad g(s^*) = f(\theta') = f(\theta) \end{aligned}$$

□

# Maskin-Monotonicity: A Review

More remarkably, Maskin-monotonicity is also "largely" sufficient.

## Condition (No Veto Power)

$\forall \theta, a$: *if* $\left|\left\{ i : a \in \arg\max_{b \in A} u_i(b, \theta) \right\}\right| \geq n - 1$, *then* $a = f(\theta)$.

## Theorem (Maskin)

*Assume* $n \geq 3$ *and that* $f$ *satisfies NVP. Then if* $f$ *is Maskin-monotonic, it is implementable.*

# Outline

# Evidence-Monotonicity

▸ honest cor.

## Definition (Evidence-monotonicity)

A SCF $f$ is evidence-monotonic if there exists $e^* : \Theta \to E$ s.t.

(i) for all $\theta$, $i$: $e_i^*(\theta) \in E_i^\ell(\theta)$; and

(ii) for all $\theta$ and $\theta'$,

if

$$\forall i, a, e_i : \left[ \begin{array}{c} U_i\left(f(\theta), e_i^*(\theta), \theta\right) \geq U_i\left(a, e_i, \theta\right) \\ \Rightarrow U_i\left(f(\theta), e_i^*(\theta), \theta'\right) \geq U_i\left(a, e_i, \theta'\right) \end{array} \right]$$

then $f(\theta) = f(\theta')$.

# Characterization: Necessity

Illustrate some applications of the condition later

- The existential qualifier on $e^*$ suggests that, in principle, may be cumbersome to verify
- But often simple, particularly because $E_i^\ell(\theta)$ may be very small (possibly singleton)

Can show that $f$ being evidence-monotonic

- is strictly weaker than $f$ being Maskin-monotonic
- is equivalent to existence of a certain kind of SCC that is Maskin-monotonic on the extended space $A \times E$

## Theorem

*If $f$ is implementable then it is evidence-monotonic.*

# Characterization: Sufficiency with $n \geq 3$

## Definition

There is disagreement if for all $\theta$ and $a$,

$$\left| \left\{ i : a \in \arg\max_b u_i(b, \theta) \right\} \right| < n - 1.$$

- If $n \geq 3$, satisfied in any "economic" environment
  - $\exists$ a private good that the SCF never gives all of to one agent
- For any $n$, guaranteed if planner has an *open* set of off-path transfers (can be only $\varepsilon$)
- Stronger than usual no veto power

## Theorem

*Assume disagreement and $n \geq 3$. Then any evidence-monotonic SCF is implementable.*

# Characterization: Sufficiency with $n = 2$

Moore and Repullo (1990) introduced the following condition:

**Definition**

A SCF $f$ has a bad outcome, $z$, if for all $\theta$, $i$, and $a \in f(\Theta)$, $u_i(z, \theta) < u_i(a, \theta)$.

- In economic applications, there is often a bad outcome: no trade, zero allocation, etc.

**Theorem**

*Assume disagreement and $n = 2$. If $f$ is evidence-monotonic and has a bad outcome, then $f$ is implementable.*

# Outline

# Normal Hard Evidence

## Definition

A hard evidence structure satisfies normality if $\forall i, \theta$, $\exists \bar{e}_i(\theta) \in E_i^\ell(\theta)$ s.t.

$$\bar{e}_i(\theta) \in E_i^\ell(\theta') \implies E_i^\ell(\theta) \subseteq E_i^\ell(\theta')$$

Interpretation:

- At $\theta$, $\bar{e}_i(\theta)$ is a "maximal" evidence for $i$
- The earlier example had normal hard evidence: $\bar{e}_1(\theta) = \theta$
- Normality holds if there are no time/space constraints on presenting evidence
- Most models with hard evidence assume normality

# Normal Hard Evidence

## Definition

A SCF $f$ satisfies non-satiation if for all $i$, $\theta$, and $a \in f(\Theta)$, there exists $\tilde{a}$ such that $u_i(\tilde{a}, \theta) > u_i(a, \theta)$.

- Intuitively, it is always be possible to reward a player
- Again, satisfied in "economic" environments or if there are off-path transfers available

# Normal Hard Evidence

PROPOSITION.

*Assume normal hard evidence and let $f$ satisfy non-satiation. Then $f$ is evidence-monotonic if for any $\theta$, $\theta'$:*

$$f(\theta) \neq f(\theta') \implies E^{\ell}(\theta) \neq E^{\ell}(\theta).$$

- I.e., measurability with respect to players' joint evidence
- Immediate corollaries, in particular because **every** $f$ is evidence-monotonic (given non-satiation) if there is universal distinguishability:

$$\theta \neq \theta' \implies E^{\ell}(\theta) \neq E^{\ell}(\theta)$$

- Implies implementation in the example (it has non-satiation, disagreement, and bad outcome)

▸ e.g.

# Preferences for Honesty

Suppose that players have a small preference for honesty when asked for a direct message about the state.

Formally, for each $i$, $E_i = \Theta$ and

$$c_i(\theta', \theta) = \begin{cases} 0 \text{ if } \theta' = \theta \\ \varepsilon \text{ if } \theta' \neq \theta \end{cases}$$

where $\varepsilon > 0$ can be arbitrarily small.

## Proposition.

*Under honesty preferences, every SCF is evidence-monotonic.*

Proof: use $e_i^*(\theta) = \theta$ and verify definition. ▸ EM

## Corollary.

*Assume honesty preferences. If there is disagreement and either $n \geq 3$ or [$n = 2$ and $f$ has a bad outcome] then $f$ is implementable.*

# Preferences for Honesty

Suppose that players have a small preference for honesty when asked for a direct message about the state.

Formally, for each $i$, $E_i = \Theta$ and

$$c_i(\theta', \theta) = \begin{cases} 0 \text{ if } \theta' = \theta \\ \varepsilon \text{ if } \theta' \neq \theta \end{cases}$$

where $\varepsilon > 0$ can be arbitrarily small.

## Proposition.

*Under honesty preferences, every SCF is evidence-monotonic.*

Proof: use $e_i^*(\theta) = \theta$ and verify definition. ▸ EM

## Corollary.

*Assume honesty preferences. If there is disagreement and either $n \geq 3$ or [$n = 2$ and $f$ has a bad outcome] then $f$ is implementable.*

# Preferences for Honesty: A Mechanism

An implementing mechanism due to Dutta and Sen (2010):

For each $i$, $M_i = A \times \mathbb{N}$

- **Rule 1:** If there is an $i$ s.t. $\forall j \neq i$, $e_j = \theta$ and $m_j = (f(\theta), 0)$ $\implies$ outcome is $f(\theta)$
- **Rule 2:** Otherwise, outcome announced by player with highest integer

Proof: Suppose true state is $\theta'$.

1. "Truthtelling" is an eqm.
2. Pick any equilibrium. Cannot fall into Rule 2, because of disagreement.
3. But in Rule 1, if $e_k \neq \theta'$ for some $k$, $k$ can profitable deviate.

# Preferences for Honesty: Large Fines

As with the previous mechanism, our general sufficiency results

- use an "integer" game (also called "tail-chasing" mechanisms)
- this is because they are canonical mechanisms that work for *all* applications
- nevertheless, from an applied view, may be questioned for this reason

Next result responds to this.

# Preferences for Honesty: Large Fines

In many applications, typically focus on quasi-linear preferences in money and assume mechanism can impose large off-path fines.

**Proposition.**
*In above setting, assume $n \geq 2$ and player 1 has honesty preferences. Then any SCF is implementable in a direct mechanism using only two rounds of IDSDS.*

**Proof.**
$E_1 = \Theta$. Let $M_2 = \Theta$, $F \gg 0$ be large enough, and use

$$g(e_1, m_2) = \begin{cases} f(m_2) + (0,0) & \text{if } e_1 = m_2 \\ f(m_2) + (0,-F) & \text{if } e_1 \neq m_2 \end{cases} \quad \square$$

- Only need mutual knowledge of rationality (and honesty prefs)

# Conclusion

- Main message: non-cheap-talk evidence can dramatically increase scope for implementation
  - Applications to hard evidence, preferences for honesty
  - Mechanisms can be remarkably simple in certain classes of problems
  - Contribute to the critique of "just" using non-verifiability as foundations for incomplete contracts
- Characterization uses complete information assumption substantially
- But the themes carry over to incomplete information
  - e.g. small preference for honesty mechanism readily extends
  - more generally, weakening of Jackson's (1992) Bayesian-monotonicity condition

**Thank you!**

# On Inalienability of Evidence

- Hypothetical problem where evidence is alienable: planner can choose both allocation $a$ and compel the submission of an evidence vector $e$
- But still require that in equilibrium, only costless evidence is compelled
- Given $f : \Theta \to A$, a costless extension is $\hat{f} : \Theta \rightrightarrows A \times E$ such that for all $\theta$, if $(a, e) \in \hat{f}(\theta)$ then $a = f(\theta)$ and $e \in E^{\ell}(\theta)$

# On Inalienability of Evidence

**Theorem**

*Assume $n \geq 3$ and disagreement. Then $f$ is implementable with inalienable evidence if and only if a costless-extension of $f$ is implementable with alienable evidence.*

Now suppose that when evidence is inalienable, we allow the planner to forbid evidence, i.e. he can restrict any $i$'s evidence choice to lie in $\hat{E}_i \subseteq E_i$, s.t. $\forall \theta : E_i^{\ell}(\theta) \cap \hat{E}_i \neq \emptyset$.

**Theorem**

*Assume $n \geq 3$. Then $f$ is forbid-implementable with inalienable evidence if and only if a costless-extension of $f$ is implementable with alienable evidence.*

# Introduction: Relevance to Incomplete Contracting

- Foundations of incomplete contract theory debate
- Observable but not verifiable information
- Critique: if the information is observable it can effectively be *made* verifiable
  - use a Moore-Repullo mechanism to induce revelation in subgame perfect equilibrium of a sequential mechanism
- As a matter of theory, this looks devastating

# Introduction: Relevance to Incomplete Contracting

- Two main responses
  1. Renegotiation: but even here, can sometimes be circumvented by clever design (e.g. Aghion et al. 1994) and/or if there is risk aversion; or debate why cannot commit to not renegotiate
  2. Critique the critique: MR mechanism is not "robust", is too complex, requires excessive rationality
     - The MR mechanism is not a direct mechanism, and has the same player acting at multiple stages
- We find that under an (arguably mild) "behavioral" assumption, the original critique comes back with quite dramatic force
  - Simple direct mechanism induces truthful revelation with iterated deletion of strictly dominated strategies