ExpEcon Methods: The Theory of Incentive Compatible Experiments

ECON 8877

P.J. Healy

Updated 2023-11-01 at 17:39:33

Outline

1. Pay One Randomly vs. Pay All
    - History
    - Azrieli et al. (2018,2020)
    - Experimental Evidence
2. Dynamic methods
    - ACH appendix
    - Luke's student
    - Manu's student
    - Jim Cox's student
    - DOSE and Ian's paper

Pay One Randomly vs. Pay All

Savage's Hot Man Example

 Leonard "Jimmy" Savage

1. Eminem of statistics (genius from Detroit)
2. Wayne State $\rightarrow$ Michigan BS & PhD in math (1941)
3. IAS Princeton, then Chicago. Milton Friedman & W. Allen Wallis mentors
4. WWII: assistant to John von Neumann
5. *The Foundation of Statistics* (1954)
   - Subjective expected utility without objective lotteries

Savage’s Hot Man Example

Pay All: “Suppose, for example, that a hot man actually prefers a swim, a shower, and a glass of beer, in that order. Once he decides on, and thereby becomes entitled to, the swim, he can no longer appropriately be asked to decide between shower and beer... [because he would be] deciding between a swim and shower... and a swim and a beer.”

Pay One Randomly: “W. Allen Wallis has mentioned to me an interesting an very general device... (I have since seen this same device used by M. Allais.) Suppose that the hot man is instructed to rank the three acts in order, subject to the consideration that two of them wil be drawn at random... and that he is then to have whichever of those two acts he has assigned a lower rank.”

Early uses: Allais (1953), Yaari (1965)

Savage's Hot Man Example

Savage says this requires two things:

1. The hot man thinks each pair is drawn with positive probability
2. Hot man's preferences over $\{f, g\}$ are the same as over $\{f, g, h\}$

But it creates a lottery... what about preferences over lotteries?

Later authors: expected utility is required

A More Relevant Example

1. Play the following game:

	L	R
U	1, 1	0, 0
D	0, 0	1, 1

2. Guess which strategy your opponent will pick.
    - Paid $1 if right, $0 if wrong.

Paying for both decisions creates a hedging problem:

Truth: $2 if right, $0 if wrong
Hedge: $1 for sure

Another Problem

Experiment: Correlate dictator-game giving with risk preferences

1. High-Stakes Dictator Game
   - Each subject given \$100
   - Paired with another subject (anonymously)
   - Asked how much he wants to give to the other subject (Dollar increments)
2. Holt-Laury (2002) procedure for estimating risk preferences.

#	Safe Lottery	Risky Lottery
1	(0.1, \$2.00; 0.9, \$1.60)	(0.1, \$3.85; 0.9, \$0.10)
2	(0.3, \$2.00; 0.7, \$1.60)	(0.3, \$3.85; 0.7, \$0.10)
3	(0.5, \$2.00; 0.5, \$1.60)	(0.5, \$3.85; 0.5, \$0.10)
4	(0.7, \$2.00; 0.3, \$1.60)	(0.7, \$3.85; 0.3, \$0.10)
5	(0.9, \$2.00; 0.1, \$1.60)	(0.9, \$3.85; 0.1, \$0.10)

Another Problem

Suppose paying for all 6 decisions:

- Wealth effect: Earning $90 in dictator game may reduce risk aversion
- Portfolio effect: The 5 risky lotteries as a portfolio aren't that risky

A Proposed Solution

Proposed solution: Pay for one randomly-selected decision

Names used for this mechanism:

1. Random Problem Selection (RPS) mechanism
2. Pay One Randomly (POR)
3. Random Lottery Incentive Mechanism (RLIM)
4. ⋮

A Problematic Example (Holt 1986, Cox et al 2011)

Let $L = (0.5, \$0; 0.5, \$3)$.

- Decision 1: L vs. \$1 for sure
- Decision 2: L vs. \$2 for sure
- Each decision chosen for payment w/ 50% probability

- Suppose $\$2 \succ L \succ \$1$
- Picking $\{L, \$2\}$ gives lottery $(0.25, \$0; 0.5, \$2; 0.25, \$3)$ (TRUTH)
- Picking $\{\$1, \$2\}$ gives lottery $(0.5, \$1; 0.5, \$2)$ (LIE)
- $\exists$ RDU preferences where $\$2 \succ L \succ \$1$ and LIE $\succ$ TRUTH

$$U(x_1, p_1; \ldots; x_n, p_n) = \sum_{i=1}^{n} u(x_i) \left[w(\sum_{j=1}^{i} p_j) - w(\sum_{j=1}^{i-1} p_j) \right]$$

Karni & Safra (1987)

Preference reversal literature:

- Lichtenstein & Slovic (1971): binary choice & WTP, inconsistent
- Grether & Plott (1979): more careful design & BDM incentives
    - State valuation \$$m$
    - Random dollar amount \$$d$ is drawn
    - Get item if $d < m$
    - Get \$$d$ if $d \geq m$

Karni & Safra:

- Write out the two-stage lottery
- Assume rank-dependent utility (non-EU)

$$U(x_1, p_1; \ldots; x_n, p_n) = \sum_{i=1}^{n} u(x_i) \left[w(\sum_{j=1}^{i} p_j) - w(\sum_{j=1}^{i-1} p_j) \right]$$

- Finds an example where $A \succ B$ but announce $v(A) < v(B)$ in BDM

Karni & Safra (1987)

Theorem: There are never any preference reversals in any such experiment if and only if preferences satisfy expected utility

Following this paper, many authors believed “RPS is incentive compatible if and only if people satisfy EU”

Unseen problem: Karni & Safra (1987) implicitly assumed ROCL.
And CompIND + ROCL $\Rightarrow$ MixIND
So it it MixIND or CompIND that’s needed?

Some authors were on the right track (Harrison & Swarthout 2014, e.g.) but necessary and sufficient conditions for RPS were still unclear.

As of 2011, things weren’t nailed down

Mechanism Usage as of 2011

Mechanism:	Only 1 Task	None Paid	One Random	Some Random	All Paid	Rank-Based	Total
	Individual Choice Experiments						
‘Top 5’	7	0	3	1	3	0	14
ExpEcon	3	0	1	0	2	0	6
	Muti-Person (Game) Experiments						
‘Top 5’	9	0	1	0	8	0	18
ExpEcon	8	1	3	3	5	1	21
Totals	27	1	8	4	18	1	59

1. Experimenters lack a convention.
2. Theory is unclear. Is expected utility needed for RPS??

ACH 2018

Goals of Azrieli et al. 2018

1. Describe an abstract model of experiment
2. Define a notion of incentive compatibility of the payment mechanism ("each decision is made as if in isolation")
3. Understand under what conditions the RPS mechanism is incentive compatible (answer: 'monotonicity')
4. Characterize the set of incentive compatible payment mechanisms (assuming monotonicity)
5. Perform 3 & 4 for the Pay All mechanism as well

An Abstract Model of Experiment

- X: A finite set of 'objects' (no structure).
- $D = (D_1, \ldots, D_k)$: A finite list of decision problems, where each $D_i \subseteq X$. Assume $D_i \neq D_j$ and $|D_i| > 1$ for every i (can be easily relaxed).

- $\succeq$ over X (complete & transitive)
- $\mu_i(\succeq) = \{x \in D_i : (\forall y \in D_i)\, x \succeq y\}$
- $\mu(\succeq) = \times_i \mu_i(\succeq)$ ('optimal choices in isolation')

- Messages: $M = \times_i D_i$ ('announced choice')
- Payment mechanism: Maps M to 'payments'

Static/simultaneous framework. We'll discuss dynamics later.

The Example

- First decision: dictator game
  $D_1 = \{(\$100, \$0), (\$99, \$1), \ldots, (\$0, \$100)\}$. $m_1 = (\$90, \$10)$
- Next: 5-question Holt-Laury elicitation
  $D_2 = \{(0.1, \$2; \$1.60), (0.1, \$3.85; \$0.10)\}$. $m_2 = (0.1, \$2; \$1.60)$
  $D_3 = \{(0.3, \$2; \$1.60), (0.3, \$3.85; \$0.10)\}$. $m_3 = (0.3, \$2; \$1.60)$
  $D_4 = \{(0.5, \$2; \$1.60), (0.5, \$3.85; \$0.10)\}$. $m_4 = (0.5, \$2; \$1.60)$
  $D_5 = \{(0.7, \$2; \$1.60), (0.7, \$3.85; \$0.10)\}$. $m_5 = (0.7, \$3.85; \$0.10)$
  $D_6 = \{(0.9, \$2; \$1.60), (0.9, \$3.85; \$0.10)\}$.
  $m_6 = (0.9, \$3.85; \$0.10)$
- Payment: RPS Mechanism
  - Roll a 6-sided die.
  - Roll a 1: pay m_1
  - Roll a 2: pay m_2

    $\vdots$
  - Roll a 6: pay m_6

Application to Games

A decision problem:

	Red ball	Green ball
U	2A, 1O	3A, 2O
D	1A, 3O	2A, 3O

A game:

	L	R
U	\$2,\$1	\$3,\$2
D	\$1,\$3	\$2,\$3

$\succeq$ are over S_i, not dollar payments.
($\succeq$ represented by u and p)

Two-Stage Lotteries

Like Karni & Safra, we need to analyze preferences over two-stage lotteries

- Preferences we're studying ($\succeq$) are over X
- "Payment objects" are different! $\mathcal{P}(X) \neq X$ (unless $k = 1$)
  - Random choice of what's paid $\Rightarrow \mathcal{P}(X) = \Delta(X)$ (lotteries or acts)
  - Pay all: $\mathcal{P}(X) = 2^X$ (bundles from X)
- Need to extend $\succeq$ to $\mathcal{P}(X)$
- Call this $\succeq^*$
- Incentive compatibility *must* be an analysis of $\succeq^*$, not $\succeq$

Payments: Acts vs Lotteries

The researcher may use a randomization device (say, roll a die) to determine which element of X is chosen for payment

Two possible approaches regarding how the subject views this uncertainty:

1. Savage (1954): Payment based on a die roll is an *act*
    - Finite state space $\Omega = \{\omega_1, \ldots, \omega_n\}$
    - A payment $f(\omega) \in X$ for each $\omega \in \Omega$
    - The set of all acts is $\mathcal{F} = X^\Omega$. So $\mathcal{P}(X) \subseteq X^\Omega$
    - Each $m \in M$ is mapped to some act $\phi(m) \in \mathcal{F}$

2. Payment based on a die roll is an *objective lottery*
    - $\Delta(X)$ – the set of lotteries on X. So $\mathcal{P}(X) \subseteq \Delta(X)$
    - Each $m \in M$ is mapped to some lottery $\varphi(m) \in \Delta(X)$

Incentive Compatibility (Acts)

- Each $\succeq$ over X extends to $\succeq^*$ over $\mathcal{F}$
- $\succeq^*$ agrees with $\succeq$ on constant acts
    - "Consistency" (Barbera 1977)
- Let $\mathcal{E}(\succeq)$ be the set of admissible extensions of $\succeq$

Definition
An experiment (D, ϕ) is **incentive compatible** with respect to $\mathcal{E}$ if, for every $\succeq$ and extension $\succeq^* \in \mathcal{E}(\succeq)$, every $m^* \in \mu(\succeq)$ and every $m \in M$,

$$\phi(m^*) \succeq^* \phi(m)$$

and

$$\phi(m^*) \succ^* \phi(m)$$

whenever $m \notin \mu(\succeq)$.

Strict incentive compatibility.

IC Experiments vs. IC Mechanisms

How does this differ from classical mechanism design?

Mechanism Design:

- Trying to implement a particular SCF/SCC
- IC is only important because IC $\iff$ implementable
    - We don't really *care* about truth-telling *per se*
- Often, weak IC is fine
- Usually use deterministic mechanisms
    - Exceptions: Gibbard (1977), Barbera (1977)

IC Experiments:

- Don't directly care what outcomes are paid (SCF/SCC)
    - (Except budget considerations, etc.)
- IC is important b/c we want to observe true $\succeq$!
- We will demand strict IC to ensure $\succeq$ is observed
- Need to rely on random mechanisms

Preliminary Observation

Proposition
If no restrictions are placed on $\mathcal{E}(\succeq)$ (other than consistency), then there is an IC payment mechanism if and only if there is only one decision problem ($k = 1$).

Monotonicity

What restrictions on $\succeq^*$?

- (Subjective) expected utility representation
- Probabilistic sophistication
- Uncertainty aversion (say, maxmin expected utility)

  $\vdots$

(Statewise) Monotonicity (Savage's P3):

$$f(\omega) \succeq g(\omega) \; \forall \omega \Rightarrow f \succeq^* g$$
$$\text{and } f(\omega) \succ g(\omega) \text{ for some } \omega \Rightarrow f \succ^* g$$

$\mathcal{E}^{\mathrm{mon}}(\succeq)$ = set of all monotonic extensions of $\succeq$

Monotonicity

	States of the World					
Act	1	2	3	4	5	6
f	\$1	\$25	pizza	\$0	\$1	Twix
g	\$1	\$24	pizza	\$0	\$1	Mars

\$25 $\succ$ \$24 and Twix$\succ$Mars $\Rightarrow f \succ^* g$

Monotonicity and Dominance

Lemma
*An experiment (D, ϕ) is incentive compatible w.r.t. $\mathcal{E}^{\mathrm{mon}}$ if and only if it has the "**Truth Dominates Lies**" property:*

For every $\succeq$, $m^* \in \mu(\succeq)$, $m \in M$ and $\omega \in \Omega$,

$$\phi(m^*)(\omega) \succeq \phi(m)(\omega).$$

If $m \notin \mu(\succeq)$ then there is $\omega \in \Omega$ such that

$$\phi(m^*)(\omega) \succ \phi(m)(\omega).$$

The RPS Mechanism

Definition
ϕ is an RPS mechanism if $\exists$ a partition $\{\Omega_1, \ldots, \Omega_k\}$ of Ω into non-empty sets such that

$$\omega \in \Omega_i \Rightarrow \phi(m)(\omega) = m_i.$$

(Assume each Ω_i is non-null.)

Die roll example: $\Omega = \{\omega_1, \ldots, \omega_6\}$ and each $\Omega_i = \{\omega_i\}$

RPS and Monotonicity

Proposition
If only monotonic extensions are admissible ($\mathcal{E} \subseteq \mathcal{E}^{\mathrm{mon}}$) then any RPS mechanism is incentive compatible.

Sketch of Proof:

Suppose each $D_i = \{x_i, y_i, z_i, \ldots\}$
Suppose $x_i = \mu_i(\succeq)$ for each i

	States of the World					
Act	1	2	3	4	$\cdots$	k
$\phi(x_1, x_2, x_3, \ldots, x_k)$	x_1	x_2	x_3	x_4	$\cdots$	x_k
$\phi(x_1, y_2, x_3, \ldots, x_k)$	x_1	y_2	x_3	x_4	$\cdots$	x_k
$\phi(x_1, y_2, z_3, \ldots, x_k)$	x_1	y_2	z_3	x_4	$\cdots$	x_k

Now apply previous lemma.

Monotonicity (on a restricted domain) is also necessary for incentive compatibility of the RPS mechanism.

Monotonicity

Is monotonicity strong?

Suppose X is a space of lotteries (vNM).
Monotonicity + reduction $\Rightarrow$ independence (vNM EU)

Suppose X is a space of acts (AA).
Monotonicity + order-reversal $\Rightarrow$ ambiguity neutrality (SEU)

So, problematic if subject is non-EU but satisfies ROCL or OR.

Empirical evidence??

Halevy (2007)

Four urns:

1. **K:** Known 5 Red, 5 Black
2. **U:** Unknown 10 marbles, Red or Black
3. **C1:** Compound urn, each urn composition equally likely
4. **C2:** Compound urn, either all Red or all Black

BDM used to elicit CE for a $2 bet on each urn.
All four are paid.
But is BDM IC? Halevy argues yes for theories considered.

Halevy (2007)

1. **K:** Known 5 Red, 5 Black
2. **U:** Unknown 10 marbles, Red or Black
3. **C1:** Compound urn, each urn composition equally likely
4. **C2:** Compound urn, either all Red or all Black

Values: V_K, V_U, V_{C1}, V_{C2}

Ambiguity Neutral: $V_K = V_U$
ROCL: $V_{C1} = V_{C2} = V_K$

Halevy (2007)

Ambiguity Neutral: $V_K = V_U$
ROCL: $V_{C1} = V_{C2} = V_K$

Worry for experiments: ROCL but not Amb. Neutral:

	Not ROCL	ROCL
Not Amb.Neutral	77%	**1%**
Amb. Neutral	6%	16%

Overall conclusion: ROCL and Ambiguity Neutrality highly correlated.

But was it IC to pay for all four valuations? Stay tuned...

“Acceptable” Theories

Theories under which RPS is IC:

- EU + ROCL
- Original Prospect Theory (1977)
  - Editing phase: “isolation effect”
- “Substitution of certainty equivalents” theories
  - Each second stage lottery is replaced with its CE
  - Then first stage becomes a simple lottery
  - Loomes & Sugden (1986) Disappointment Aversion
  - Segal (1988)
- Regret Theory
  - Bell (1982), Fishburn (1982,1987), Loomes & Sugden (1982,1987)

Other IC Mechanisms?

Maintaining the monotonicity assumption ($\mathcal{E} = \mathcal{E}^{\text{mon}}$), what is the class of all incentive compatible mechanisms?

From now on, assume only strict $\succeq$ are admissible:

- A unique maximal element in each decision problem ($\mu(\succeq)$ is a singleton).
- There may be $m \in M$ that cannot be rationalized:
  $D_1 = \{x, y\}$, $D_2 = \{y, z\}$, $D_3 = \{x, z\}$
  $m = (x, y, z)$ is not *rationalizable*

  M_R=rationalizable messages
  M_{NR}=non-rationalizable messages

Surely Identified Sets

Example: $D_1 = \{x, y\}$, $D_2 = \{y, z\}$, $D_3 = \{x, z\}$
Consider $E = \{x, y, z\}$
If $m \in M_R$, then we know your favorite thing in E.

Definition
A set $E \subseteq X$ is surely identified if, for every $\succeq$, the choices $m = \mu(\succeq)$ reveal the $\succeq$-maximal element of E. Let $SI(D)$ be the family of surely identified sets for D.

Lemma

$$E \in SI(D) \Leftrightarrow \forall x, y \in E \;\; \exists D_i \in D, \;\; \{x, y\} \subseteq D_i \subseteq E$$

In practice, usually not much overlap in D_i sets.
In that case, $SI(D) = \{D_i\}_{i=1}^{k} \bigcup \{x\}_{x \in X}$.

RSS Mechanisms

Given ϕ, denote $P^{\phi}(\omega) = \{\phi(m)(\omega)\}_{m \in M}$.
Things you could get paid in state ω as you vary m
In RPS, $P^{\phi}(\omega_i) = D_i$

Definition
ϕ is a Random Set Selection (RSS) Mechanism if, for each $\omega \in \Omega$, $P^{\phi}(\omega) \in SI(D)$ and for every $m \in M_R$,

$$\phi(m)(\omega) = \max(P^{\phi}(\omega)|m).$$

Interpretation: I roll a die and pay you either for a real decision you made, or for a fake decision where I can *always* figure out what you would have chosen. Note: RPS $\subset$ RSS

One known example of RSS that's not RPS : Krajbich (2011)

Characterization

Theorem
(D, ϕ) is incentive compatible w.r.t. $\mathcal{E}^{\text{mon}}$ if and only if

1. *ϕ is an RSS mechanism;*
2. *Each D_i is surely identified by the sets $\{P^\phi(\omega)\}_{\omega \in \Omega}$;*
3. *$m \in M_{NR}$ implies $\phi(m) \notin \phi(M_R)$.*

Idea of Proof:

1. At each ω you get the revealed-best possible element $\phi(m)(\omega) = \max(P^\phi(\omega)|m)$; thus, RSS
2. Each D_i matters for the outcome
3. Non-rationalizable messages give you something from each payment set (by definition), but it shouldn't your favorite in all sets.

Almost-Characterizing RPS

Usually $SI(D) = \{D_i\}_{i=1}^{k} \bigcup \{x\}_{x \in X}$.
(For example, if each D_i is disjoint.)

In this case, RSS = RPS + "singleton payments"
Like a random show-up fee paid instead of a chosen object
Example: $\Omega = \{\omega_1, \ldots, \omega_6, \omega_7\}$
$\phi(m)(\omega_i) = m_i$ if $i \leq 6$
$\phi(m)(\omega_7) = \$10$

Thus, in practice, IC $\iff$ RPS + singleton payments

Dynamic Settings

How to apply this to dynamic settings? Topics:

1. Repeated games
2. Updating beliefs/preferences
3. Adaptive designs: D_i depends on m_{i-1}
4. Experimentation incentives

Repeated Games

- Example: repeated Prisoners' Dilemma
- Fixed opponent
    - That's not *k* decisions! That's one huge decision
    - RPS: pay one random *supergame*
    - How to pay within the supergame? Stay tuned
- Different opponents
    - Sequence of one-shot PDs
    - Now you can pay one randomly
    - But what about updating/learning?

Updating & Learning

- Preferences $\succeq$ might change during the course of an experiment
    - Updated beliefs
    - Other reasons
- That's okay if subject treats past choices as "sunk"
- RPS is incentive compatible "going forward"

Dynamically-Generated Decisions

- In general, it's not IC to have D_i depend on m_{i-1}
- Example: Learn $\succ$ via $D_1 = \{a, b, c\}$ and $D_2 = D_1 \setminus \{m_1\}$
  - Suppose $a \succ b \succ c$
  - Truth: $m^* = (a, b) \mapsto (a, 1/2; b, 1/2)$.
  - Lie: $m = (b, a) \mapsto (b, 1/2; a, 1/2)$
  - Not strictly IC
  - Not even weakly IC if D_2 is paid with $> 1/2$ chance
- Solution: Pay based on an equivalent hypothetical static-choice experiment
- Can save a lot of time/questions!
- Only known usage: Krajbich

Experimentation Incentives

- Example: Two plays of binary ultimatum game: $(9, 1)$ or $(5, 5)$
- Opponent can only reject if $(9, 1)$.
- Might prefer $(5, 5)$ but $(9, 1)$ in $t = 1$ gives more info!
- Similar to multi-armed bandit problem
- Problem: *Experimentation incentives*
- Solution: Full feedback via the strategy method
    1. Player 2: Will you accept or reject if $(9, 1)$?
    2. Player 1: pick $(9, 1)$ or $(5, 5)$ without knowing P2's choice
    3. If $(9, 1)$ then P2's prior choice applies
    4. Feedback: P1 is told P2's choice *even if* they played $(5, 5)$
- Eliminates informational differences across choices

Paying in Bundles

$\succ$ only defined over single choice objects. Payments can be bundles. Need to extend $\succ$ to $\succeq^* \in \mathcal{P}(X) = 2^X$ (bundles).

No Complementarities at the Top (NCaT): If x_i is the true favorite in each decision problem, then for any $(y_1, \ldots, y_k)$,

$$\{x_1, x_2, \ldots, x_k\} \succ^* \{y_1, y_2, \ldots, y_k\}.$$

NCaT Violations: fairness, wealth effects, portfolio effects, hedging effects...

Theorem

1. *If $\succ^*$ satisfies NCaT then the Pay-All mechanism is IC.*
2. *If we assume nothing else (and $\succ$ is strict), it is essentially unique.*

Different versions of NCaT if paying 3 random decisions, *e.g.*

Testing IC

How to test IC of payment mechanism:

	D_1	D_2
Treatment 1:	$\{\$4, (\frac{1}{2}, \$10)\}$	
Treatment 2:	$\{\$4, (\frac{1}{2}, \$10)\}$	$\{\$3, (\frac{1}{2}, \$12)\}$

If we observe differences on D_1, it could be

- the mechanism was not IC, or
- the presence of D_2 altered preferences (e.g., decoy effect).

- Cubitt Starmer Sugden (1998 Exp.1)
- Beattie & Loomes (1997)
- Cubitt Starmer Sugden (1998 Exp.2)
- Harrison & Swarthout (2014)
- Cox Sadiraj & Schmidt (2015)

Tests Without Framing Confound

Replace Treatment 1 with a "Framed Control" treatment:

	D_1	D_2	Mechanism
Treatment 1:	$\{\$4, (\frac{1}{2}, \$10)\}$	$\{\$3, (\frac{1}{2}, \$12)\}$	Pay only D_1
Treatment 2:	$\{\$4, (\frac{1}{2}, \$10)\}$	$\{\$3, (\frac{1}{2}, \$12)\}$	RPS

LESSON: Proper test of IC must show all subjects same choices.

Incentive Compatibility (Lotteries)

- Each $\succeq$ over X extends to $\succeq^*$ over $\Delta(X)$
- $\succeq^*$ agrees with $\succeq$ on degenerate lotteries
- Let $\mathcal{E}(\succeq)$ be the set of admissible extensions of $\succeq$

Definition
An experiment (D, φ) is **incentive compatible** with respect to $\mathcal{E}$ if, for every $\succeq$ and extension $\succeq^* \in \mathcal{E}(\succeq)$, every $m^* \in \mu(\succeq)$ and every $m \in M$,

$$\varphi(m^*) \succeq^* \varphi(m)$$

and

$$\varphi(m^*) \succ^* \varphi(m)$$

whenever $m \notin \mu(\succeq)$.

Monotonicity (Lotteries)

Definition
Fix $\succeq$. The lottery f *First Order Stochastically Dominates (FOSD)* the lottery g with respect to $\succeq$ if, for every $x \in X$,

$$\sum_{\{x' \in X : x' \succeq x\}} f(x') \geq \sum_{\{x' \in X : x' \succeq x\}} g(x').$$

If there is strict inequality for at least one x then we say f *strictly FOSD* g with respect to $\succeq$.

Definition
An extension $\succeq^*$ of $\succeq$ is monotonic if $f \succeq^* g$ whenever f FOSD g w.r.t. $\succeq$ and $f \succ^* g$ whenever f strictly FOSD g w.r.t. $\succeq$.

$\mathcal{E}^{\mathrm{mon}}(\succeq)$ = The set of all monotonic extensions of $\succeq$.

Monotonicity and Dominance (Lotteries)

Lemma
*A mechanism φ is incentive compatible with respect to $\mathcal{E}^{\text{mon}}$ if and only if, for every $\succeq$ and every $m \neq \mu(\succeq)$, $\varphi(\mu(\succeq))$ FOSD $\varphi(m)$ w.r.t. $\succeq$. (**Truth FOSD's Lies**)*

The RPS Mechanism (Lotteries)

Definition
A mechanism φ is an RPS mechanism if there exists a full-support probability distribution λ over $D = (D_1, \ldots, D_k)$ such that for every alternative $x \in X$,

$$\varphi(m)(x) = \sum_{\{i \,:\, m_i = x\}} \lambda(D_i).$$

RPS and Monotonicity (Lotteries)

Proposition
If only monotonic extensions are admissible ($\mathcal{E} \subseteq \mathcal{E}^{\text{mon}}$) then any RPS mechanism is incentive compatible.

Sketch of Proof:

- Lying in any decision problem shifts probability from more to less desired objects, hence any lottery that can be obtained by lying is FOSD by the lottery obtained by truth-telling
- Now apply previous lemma

What else is IC (with $\mathcal{E}^{\text{mon}}$)?

Example:

- $D_1 = \{x, y\},\ D_2 = \{x, z\},\ D_3 = \{y, z\}$
- Consider the mechanism φ that puts probability of 0.8 on the revealed most preferred object and 0.2 on the revealed second-best (for $m \in M_R$)
- φ is IC but not an RPS mechanism (even when restricted to M_R)
- $E = \{x, y, z\}$ is SI
- $\lambda(D_1) = \lambda(D_2) = \lambda(D_3) = 0.2,\ \lambda(E) = 0.4$ generates φ

Lesson: We may put weight on surely identified sets outside of D

What else is IC (with $\mathcal{E}^{\mathrm{mon}}$)?

Example:

- $D_1 = \{x, y\},\ D_2 = \{x, z\},\ D_3 = \{y, z\}$
- Consider the mechanism φ that puts probability of 0.6 on the revealed most preferred object and 0.4 on the revealed second-best (for $m \in M_R$)
- φ is IC but not an RPS mechanism (even when restricted to M_R)
- $E = \{x, y, z\}$ is SI
- $\lambda(D_1) = \lambda(D_2) = \lambda(D_3) = 0.4,\ \lambda(E) = -0.2$ generates φ

Lesson: We may put *negative* weights on surely identified sets

Note: $\lambda(D_1) = \lambda(D_2) = \lambda(D_3) = 0.6,\ \lambda(E) = -0.8$ generates a non-IC mechanism. $Pr(best) = 0.4 < Pr(2nd\ best) = 0.6$

WSS Mechanisms

Definition
A mechanism $\varphi : M \to \Delta(X)$ is a *weighted set-selection (WSS) mechanism* if there exists some $\lambda : SI(D) \to \mathbb{R}$ such that for every rationalizable $m \in M_R$ and every $x \in X$,

$$\varphi(m)(x) = \sum_{\{E \in SI(D) \,:\, \max(E|m) = x\}} \lambda(E).$$

RPS $\subset$ WSS

Switch Positivity

Table 1 Examples to demonstrate the switch positivity condition

	$\lambda^1(\cdot)$	$\lambda^2(\cdot)$	$\lambda^3(\cdot)$	$\lambda^4(\cdot)$		$\succ^x$	$\succ^y$	$\succ^a$	$\succ^b$
$D_1 = \{x, y\}$	0	1/4	0	1/2		x	y	a	b
$D_2 = \{x, a\}$	1/3	1/4	1/4	1/4		y	x	b	a
$D_3 = \{y, a\}$	1/3	1/4	1/4	1/4		a	a	c	c
$D_4 = \{b, c\}$	1/3	1/4	1/4	1/4		b	b	x	x
$E_1 = \{x, y, a\}$	0	0	1/4	– 1/4		c	c	y	y
$E_2 = \{a, b\}$	–	–	–	–	$\mu(\cdot) =$	(x, x, y, b)	(y, x, y, b)	(x, a, a, b)	(x, a, a, b)

Four different WSS mechanisms λ^1 through λ^4

λ^2 is RPS

$E_1 \in SI(D)$ but $E_2 \notin SI(D)$

Four example preferences on the right, with $\mu(\succ)$ given

Switch Positivity

Table 1 Examples to demonstrate the switch positivity condition

	$\lambda^1(\cdot)$	$\lambda^2(\cdot)$	$\lambda^3(\cdot)$	$\lambda^4(\cdot)$		$\succ^x$	$\succ^y$	$\succ^a$	$\succ^b$
$D_1 = \{x, y\}$	0	1/4	0	1/2		x	y	a	b
$D_2 = \{x, a\}$	1/3	1/4	1/4	1/4		y	x	b	a
$D_3 = \{y, a\}$	1/3	1/4	1/4	1/4		a	a	c	c
$D_4 = \{b, c\}$	1/3	1/4	1/4	1/4		b	b	x	x
$E_1 = \{x, y, a\}$	0	0	1/4	– 1/4		c	c	y	y
$E_2 = \{a, b\}$	–	–	–	–	$\mu(\cdot) =$	(x, x, y, b)	(y, x, y, b)	(x, a, a, b)	(x, a, a, b)

How to differentiate $\succ^x$ and $\succ^y$?

Need x vs. y, which is given by D_1 or E_1. "Switch test sets" for $\{x, y\}$

Thus, λ^1 is not IC because $\lambda^1(D_1) = \lambda^1(E_1) = 0$

λ^2 (RPS) and λ^3 are IC

$\varphi^3(\succ^x)(x) = 1/2$ (from $E_1 + D_2$) $> \varphi^3(\succ^x)(y) = 1/4$ (from D_3)

$\varphi^3(\succ^y)(y) = 1/2$ (from $E_1 + D_3$) $> \varphi^3(\succ^y)(x) = 1/4$ (from D_2)

Switch Positivity

Table 1 Examples to demonstrate the switch positivity condition

	$\lambda^1(\cdot)$	$\lambda^2(\cdot)$	$\lambda^3(\cdot)$	$\lambda^4(\cdot)$		$\succ^x$	$\succ^y$	$\succ^a$	$\succ^b$
$D_1 = \{x, y\}$	0	1/4	0	1/2		x	y	a	b
$D_2 = \{x, a\}$	1/3	1/4	1/4	1/4		y	x	b	a
$D_3 = \{y, a\}$	1/3	1/4	1/4	1/4		a	a	c	c
$D_4 = \{b, c\}$	1/3	1/4	1/4	1/4		b	b	x	x
$E_1 = \{x, y, a\}$	0	0	1/4	– 1/4		c	c	y	y
$E_2 = \{a, b\}$	–	–	–	–	$\mu(\cdot) =$	(x, x, y, b)	(y, x, y, b)	(x, a, a, b)	(x, a, a, b)

Don't even need $\lambda(E_1)$ to be positive! See λ^4:
λ^4 is still IC for $\succ^x$ and $\succ^y$ since:

$$\varphi^4(\mu(\succ^x))(x) = \underbrace{\lambda^4(D_1)}_{1/2} + \underbrace{\lambda^4(E_1)}_{-1/4} + \underbrace{\lambda^4(D_2)}_{1/4} > \underbrace{\lambda^4(D_3)}_{1/4} = \varphi^4(\mu(\succ^x))(y)$$

$$\varphi^4(\mu(\succ^y))(y) = \overbrace{\lambda^4(D_1)} + \overbrace{\lambda^4(E_1)} + \overbrace{\lambda^4(D_3)} > \overbrace{\lambda^4(D_2)} = \varphi^4(\mu(\succ^y))(x)$$

Switch Positivity

Table 1 Examples to demonstrate the switch positivity condition

	$\lambda^1(\cdot)$	$\lambda^2(\cdot)$	$\lambda^3(\cdot)$	$\lambda^4(\cdot)$		$\succ^x$	$\succ^y$	$\succ^a$	$\succ^b$
$D_1 = \{x, y\}$	0	1/4	0	1/2		x	y	a	b
$D_2 = \{x, a\}$	1/3	1/4	1/4	1/4		y	x	b	a
$D_3 = \{y, a\}$	1/3	1/4	1/4	1/4		a	a	c	c
$D_4 = \{b, c\}$	1/3	1/4	1/4	1/4		b	b	x	x
$E_1 = \{x, y, a\}$	0	0	1/4	– 1/4		c	c	y	y
$E_2 = \{a, b\}$	–	–	–	–	$\mu(\cdot) =$	(x, x, y, b)	(y, x, y, b)	(x, a, a, b)	(x, a, a, b)

But we don't need switch test sets for *all* $\{x, y\}$ pairs.

Note $\mu(\succ^a) = \mu(\succ^b)$ since $\{a, b\}$ are not in any SI sets together

This experiment isn't designed to distinguish $a \succ b$ vs. $b \succ a$!

We only need weight on switch test set E for $\{x, y\}$ if $\{x, y\} \subseteq E \in SI(D)$

Switch Positivity

Table 1 Examples to demonstrate the switch positivity condition

	$\lambda^1(\cdot)$	$\lambda^2(\cdot)$	$\lambda^3(\cdot)$	$\lambda^4(\cdot)$		$\succ^x$	$\succ^y$	$\succ^a$	$\succ^b$
$D_1 = \{x, y\}$	0	1/4	0	1/2		x	y	a	b
$D_2 = \{x, a\}$	1/3	1/4	1/4	1/4		y	x	b	a
$D_3 = \{y, a\}$	1/3	1/4	1/4	1/4		a	a	c	c
$D_4 = \{b, c\}$	1/3	1/4	1/4	1/4		b	b	x	x
$E_1 = \{x, y, a\}$	0	0	1/4	– 1/4		c	c	y	y
$E_2 = \{a, b\}$	–	–	–	–	$\mu(\cdot) =$	(x, x, y, b)	(y, x, y, b)	(x, a, a, b)	(x, a, a, b)

Tentative necessary condition:
For every $x, y \in X$ it holds that

$$\sum_{\{E \in SI(D)\,:\,\{x,y\} \subseteq E\}} \lambda(E) > 0$$

Switch Positivity

Table 1 Examples to demonstrate the switch positivity condition

	$\lambda^1(\cdot)$	$\lambda^2(\cdot)$	$\lambda^3(\cdot)$	$\lambda^4(\cdot)$		$\succ^x$	$\succ^y$	$\succ^a$	$\succ^b$
$D_1 = \{x, y\}$	0	1/4	0	1/2		x	y	a	b
$D_2 = \{x, a\}$	1/3	1/4	1/4	1/4		y	x	b	a
$D_3 = \{y, a\}$	1/3	1/4	1/4	1/4		a	a	c	c
$D_4 = \{b, c\}$	1/3	1/4	1/4	1/4		b	b	x	x
$E_1 = \{x, y, a\}$	0	0	1/4	– 1/4		c	c	y	y
$E_2 = \{a, b\}$	–	–	–	–	$\mu(\cdot) =$	(x, x, y, b)	(y, x, y, b)	(x, a, a, b)	(x, a, a, b)

But wait... we may need to distinguish things not at the top of $\succ$

Suppose z is added to the top of both $\succ^x$ and $\succ^y$

Consider $E_3 = \{x, y, z, a\}$. Now E_3 isn't a helpful switch test set!

Both would pick z

Need to require that $E \subseteq \{a, b, c, x, y\}$ so that $z \notin E$

Switch Positivity

Definition
A WSS mechanism φ (with associated weighting vector λ) satisfies *switch positivity* if, for every $x, y \in X$ and $A \subseteq X \setminus \{x, y\}$ it holds that

$$\sum_{\{E \in SI(D) \,:\, \{x,y\} \subseteq E \subseteq A \cup \{x,y\}\}} \lambda(E) > 0$$

(provided the sum is not empty).

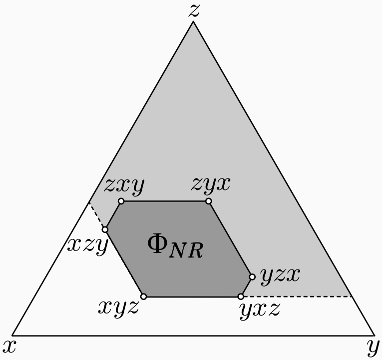
Here, A is the stuff ranked lower than x and y for some possible pair $\succ^x$ and $\succ^y$

Almost There...

So we need a WSS mechanism that satisfies switch positivity...
but we also need to make sure $m \in M_{NR}$ is never optimal

Non-Rationalizable Messages

Example: $D_1 = \{x, y\},\ D_2 = \{x, z\},\ D_3 = \{y, z\}$. $\Delta(X)$ is shown:



Light gray: lotteries dominated by $\varphi(\mu(\succ))$ for $x \succ y \succ z$

Dark gray: lotteries dominated by $\varphi(\mu(\succ))$ for all possible $\succ$

So if $m \in M_{NR}$, pay something in Φ_{NR}!!

Characterization (Lotteries)

Theorem
(D, φ) is incentive compatible w.r.t. $\mathcal{E}^{\mathrm{mon}}$ if and only if

1. *φ is a WSS mechanism;*
2. *φ satisfies switch positivity;*
3. *if $m \in M_{NR}$ then $\varphi(m) \in conv(\varphi(M_R)) \setminus \varphi(M_R)$.*

‘Proof’

$D_1 = \{x, y\},\ D_2 = \{x, z\},\ D_3 = \{y, z\}$

There is a normalized and convex ‘capacity’ $v : 2^{\{x,y,z\}} \to [0, 1]$ that ‘represents’ φ:

$$\begin{array}{rcl} \varphi(\mu(\succeq))(a_1) & = & v(a_1, a_2, a_3) - v(a_2, a_3) \\ \varphi(\mu(\succeq))(a_2) & = & v(a_2, a_3) - v(a_3) \\ \varphi(\mu(\succeq))(a_3) & = & v(a_3) \end{array}$$

$\{a_1, a_2, a_3\} = \{x, y, z\}$ and $\succeq$ ranks $a_1 \succeq a_2 \succeq a_3$.

‘Proof’ (cont.)

Each v can be represented uniquely by the ‘unanimity capacities’:

$$v(A) = \sum_{E \subseteq A} \lambda(E)$$

$$\begin{aligned}
\varphi(\mu(\succeq))(a_1) &= v(a_1, a_2, a_3) - v(a_2, a_3) = \sum_{a_1 \in E} \lambda(E) \\
\varphi(\mu(\succeq))(a_2) &= v(a_2, a_3) - v(a_3) = \sum_{a_2 \in E \subseteq \{a_2, a_3\}} \lambda(E) \\
\varphi(\mu(\succeq))(a_3) &= v(a_3) = \sum_{E \subseteq \{a_3\}} \lambda(E)
\end{aligned}$$

But this is exactly the required representation...

Note: v convex $\Leftrightarrow$ λ satisfies “switch positivity”

IC Mechanisms: Acts vs. Lotteries

- The lotteries framework can be seen as a restriction of the set of possible extensions $\succeq^*$
- The subject is indifferent between any two acts that generate the same lottery
- Incentive compatibility becomes a weaker requirement
- ‘More’ mechanisms are IC

IC Mechanisms: Acts vs. Lotteries

Imagine the Savage framework with subjective belief μ on Ω

Definition
Say that $((\Omega, \mu), \phi)$ *generates* φ if, for each $m \in M$ and $x \in X$,

$$\varphi(m)(x) = \mu\left(\{\omega \in \Omega : \phi(m)(\omega) = x\}\right).$$

IC Mechanisms: Acts vs. Lotteries

Proposition
If ϕ is an IC act-mechanism (defined on some state space Ω), and μ is a full-support probability distribution on Ω, then the lotteries-mechanism φ generated by $((\Omega, \mu), \phi)$ is IC.

Proposition
Assume that φ is an IC lotteries-mechanism.

1. *If the associated weighting vector λ of φ is non-negative, then there exists an IC acts-mechanism ϕ (on some Ω) and a probability μ on Ω such that $((\Omega, \mu), \phi)$ generates φ on rationalizable messages.*
2. *If the associated weighting vector λ of φ contains negative elements, then φ cannot be generated by any IC acts-mechanism ϕ (even when restricted to rationalizable messages).*

Summary

- If paying all, need to assume no complementarities.
    - Fairness, portfolio, hedging, wealth, ...
- If RPS, need to assume monotonicity. Weak, unless 2-stage gambles.
    - Reduction & non-expected utility
    - Order Reversal & ambiguity aversion
- Other mechanisms may be IC for certain models.
- **Experimenter needs to decide for themselves!**

My (current) opinion:

- Use RPS
- Separate decisions as much as possible.
- Use separate, physical randomizing devices.

Other Issues

Other Monotonicity Violations:

- Decision Overload w/ Easy/Default Option (NCaT also questionable)
- Ex-Ante Fairness (NCaT: ex-post fairness)
- Irrational Diversification (NCaT also violated)
- Repeated decision problems (randomization)
- List format (Brown-Healy)

Issues Besides IC:

- Payment Inequality & Variance (matching pennies)
- Payment Size ($1/k$; same for NCaT)
- Random Choice

The End