

# **ExpEcon Methods:**

## **A Theory of Testing Theories**

---

ECON 8877

P.J. Healy

Updated 2026-01-28 at 01:56:05

# Introduction

A principal wants to learn *something* about the preferences of an agent, but *not* the whole ordering  
(Why not? complexity, costs, privacy, etc.)

**Example:** NYC school match: only list favorite 12 schools

**Which properties of preferences can be elicited in an incentive compatible way?**

## Leading Example:

$X = \{x, y, z\}$ . Let  $xyz$  denote  $x \succ y \succ z$ , e.g. Assume strict prefs.

All orderings:

$$\{xyz, xzy, zxy, zyx, yzx, yxz\}$$

A simple elicitation mechanism:

Pick from  $\{x, y\}$

Paid what you choose

## Leading Example:

$X = \{x, y, z\}$ . Let  $xyz$  denote  $x \succ y \succ z$ , e.g. Assume strict prefs.

All orderings:

$$\underbrace{\{xyz, xzy, zxy\}}_{\text{pick } x}, \underbrace{\{zyx, yzx, yxz\}}_{\text{pick } y}$$

A simple elicitation mechanism:

Pick from  $\{x, y\}$   
Paid what you choose

## Leading Example:

$X = \{x, y, z\}$ . Let  $xyz$  denote  $x \succ y \succ z$ , e.g. Assume strict prefs.

All orderings:

$$\underbrace{\{\textcolor{red}{xyz, xzy, zxy}\}}_{\text{a "type"}, \text{red}} , \underbrace{\{\textcolor{blue}{zyx, yzx, yxz}\}}_{\text{a "type", blue}}$$

A simple elicitation mechanism:

Pick from  $\{x, y\}$

Paid what you choose

## Leading Example:

$X = \{x, y, z\}$ . Let  $xyz$  denote  $x \succ y \succ z$ , e.g. Assume strict prefs.

All orderings:

a “type space” or “model” or “theory”  
 $\overbrace{\{\{xyz, xzy, zxy\}, \{zyx, yzx, yxz\}\}}^{\text{a ‘type’}}$   
a “type”

A simple elicitation mechanism:

Pick from  $\{x, y\}$   
Paid what you choose

## Leading Example:

$X = \{x, y, z\}$ . Let  $xyz$  denote  $x \succ y \succ z$ , e.g. Assume strict prefs.

All orderings:

a “type space” or “model” or “theory”  
 $\overbrace{\{\{xyz, xzy, zxy\}, \{zyx, yzx, yxz\}\}}^{\text{a ‘type’}}$   
a “type”

A simple elicitation mechanism:

Pick from  $\{x, y\}$   
Paid what you choose

This type space is *elicitable*. Truth FOSD’s lie.

$\{xyz, xzy, zxy, zyx, yzx, yxz\}$

Mechanism:

Pick from  $\{x, y\}$  and from  $\{x, z\}$

We randomly pick one of your answers and pay it to you

$$\{\underbrace{\{xyz, xzy\}}_{\text{pick } x,x}, \underbrace{\{zxy\}}_{\text{pick } x,z}, \underbrace{\{zyx, yzx\}}_{\text{pick } y,z}, \underbrace{\{yxz\}}_{\text{pick } y,x}\}$$

Mechanism:

Pick from  $\{x, y\}$  and from  $\{x, z\}$

We randomly pick *one* of your answers and pay it to you

This type space is *elicitable*. Truth FOSD's lie.

$$\left\{ \underbrace{\{xyz, yxz\}}_{\text{dislike z}}, \underbrace{\{xzy, zxy\}}_{\text{dislike y}}, \underbrace{\{yzx, zyx\}}_{\text{dislike x}} \right\}$$

$$\left\{ \underbrace{\{xyz, yxz\}}_{\text{dislike z}}, \underbrace{\{xzy, zxy\}}_{\text{dislike y}}, \underbrace{\{yzx, zyx\}}_{\text{dislike x}} \right\}$$

There are **no** menus that generate this type space.  
Generated by top two elements of  $X$

$$\left\{ \underbrace{\{xyz, yxz\}}_{\text{dislike z}}, \underbrace{\{xzy, zxy\}}_{\text{dislike y}}, \underbrace{\{yzx, zyx\}}_{\text{dislike x}} \right\}$$

There are **no** menus that generate this type space.  
Generated by top two elements of  $X$

But it is elicitable!

$$\left\{ \underbrace{\{xyz, yxz\}}_{\text{dislike z}}, \underbrace{\{xzy, zxy\}}_{\text{dislike y}}, \underbrace{\{yzx, zyx\}}_{\text{dislike x}} \right\}$$

There are **no** menus that generate this type space.  
Generated by top two elements of  $X$

But it is elicitable!

Mechanism:  
Announce least favorite,  
get paid 50-50 lottery over the other two options.

# Results

## Preview of Main Results:

Generated by top  $k$  elements  $\Rightarrow$  elicitable  $\Rightarrow$  “convex”

# Results

## Preview of Main Results:

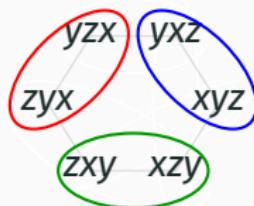
Generated by top  $k$  elements  $\Rightarrow$  elicitable  $\Rightarrow$  “convex”



# Results

## Preview of Main Results:

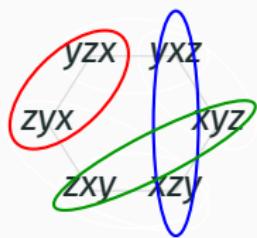
Generated by top  $k$  elements  $\Rightarrow$  elicitable  $\Rightarrow$  “convex”



# Results

## Preview of Main Results:

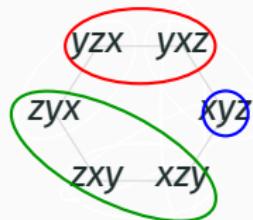
Generated by top  $k$  elements  $\Rightarrow$  elicitable  $\Rightarrow$  “convex”



# Results

## Preview of Main Results:

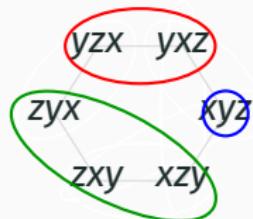
Generated by top  $k$  elements  $\Rightarrow$  elicitable  $\Rightarrow$  “convex”



# Results

## Preview of Main Results:

Generated by top  $k$  elements  $\Rightarrow$  elicitable  $\Rightarrow$  “convex”



We get complete characterization when:

1. Restrict to neutral type spaces, or
2. Pay in acts, not lotteries (no objective probabilities)

## Some related literature

- Osband (1985), Lambert-Pennock-Shoham (2008), Lambert (2018): Scoring rules to elicit properties of beliefs
- Gibbard (1977), Bahel and Sprumont (2019): Characterizing strategy-proof random mechanisms
- Carroll (2012), Saito (2013): Sufficiency of local IC constraints
- ACH (2018, 2020): Characterizing IC mechanisms in experimental environments

# The General Model

# Framework

- $X$  - a finite set of alternatives
  - Typical elements:  $x, y, z, w, \dots$
- $O$  - the set of strict orders over  $X$ 
  - Typical elements:  $\succeq, \succeq', \dots$

# Framework

- $X$  - a finite set of alternatives
  - Typical elements:  $x, y, z, w, \dots$
- $O$  - the set of strict orders over  $X$ 
  - Typical elements:  $\succeq, \succeq', \dots$

## Definition

A type space  $T = \{t_1, \dots, t_k\}$  is a partition of  $O$ .

- A type is any  $t \in T$ , so  $t = \{\succeq, \succeq', \dots, \succeq''\}$
- Example:  $t = \{\text{all } \succeq \text{ satisfying the Independence axiom}\}$
- Notation:  $t(\succeq) \in T$  is the type containing  $\succeq$

# Examples

$$X = \{x, y, z\}$$

- Entire ranking:

$$T = \{\{xyz\}, \{yxz\}, \{yzx\}, \{zyx\}, \{zxy\}, \{xzy\}\}$$

- First-best:

$$T = \{\{xyz, xzy\}, \{yxz, yzx\}, \{zyx, zxy\}\}$$

- Top-2:

$$T = \{\{xyz, yxz\}, \{xzy, zxy\}, \{zyx, yzx\}\}$$

- Best from  $\{x, y\}$ :

$$T = \{\{xyz, xzy, zxy\}, \{zyx, yzx, yxz\}\}$$

- Where you rank x:

$$T = \{\{xyz, xzy\}, \{zxy, yxz\}, \{yzx, zyx\}\}$$

(This type space is not “neutral”. Labels matter.)

# Mechanisms

$\Delta(X)$  is the set of lotteries on  $X$

## Definition

A  $T$ -mechanism is any  $g : T \rightarrow \Delta(X)$ .

- Why random payments?
  - Allows use of the RPS mechanism (and more)
  - With deterministic mechanisms very little can be elicited

## Elicitable type spaces

Recall that  $p$  strictly FOSD  $q$  relative to  $\succeq$  (written  $p \succ^* q$ ) if

$$\forall x \in X \quad p(\{y : y \succeq x\}) \geq q(\{y : y \succeq x\})$$

with strict inequality for at least one  $x$

## Elicitable type spaces

Recall that  $p$  strictly FOSD  $q$  relative to  $\succeq$  (written  $p \succ^* q$ ) if

$$\forall x \in X \quad p(\{y : y \succeq x\}) \geq q(\{y : y \succeq x\})$$

with strict inequality for at least one  $x$

### Definition

$g$  is IC if for every  $\succeq \in O$  and every  $t \neq t(\succeq)$

$$g(t(\succeq)) \succ^* g(t).$$

# Elicitable type spaces

Recall that  $p$  strictly FOSD  $q$  relative to  $\succeq$  (written  $p \succ^* q$ ) if

$$\forall x \in X \quad p(\{y : y \succeq x\}) \geq q(\{y : y \succeq x\})$$

with strict inequality for at least one  $x$

## Definition

$g$  is IC if for every  $\succeq \in O$  and every  $t \neq t(\succeq)$

$$g(t(\succeq)) \succ^* g(t).$$

## Definition

A type space  $T$  is *elicitable* if there exists an IC  $T$ -mechanism.

**Goal:** Characterize elicitable type spaces (spoiler: we can't)

## Top elements of menus

“What’s your favorite thing from  $X'$ ? ”

- Every menu  $X' \subseteq X$  corresponds to a type space:

$\succeq, \succeq' \in t \iff \succeq, \succeq'$  have the same favorite item in  $X'$

Examples:

$$X' = \{x, y\} \implies T = \{\{xyz, xzy, zxy\}, \{zyx, yzx, yxz\}\}$$

$$X' = \{x, y, z\} \implies T = \{\{xyz, xzy\}, \{yxz, yzx\}, \{zyx, zxy\}\}$$

## Top elements of menus

“What’s your favorite thing from  $X'$ ?”

- Every menu  $X' \subseteq X$  corresponds to a type space:

$$\succeq, \succeq' \in t \iff \succeq, \succeq' \text{ have the same favorite item in } X'$$

Examples:

$$X' = \{x, y\} \implies T = \{\{xyz, xzy, zxy\}, \{zyx, yzx, yxz\}\}$$

$$X' = \{x, y, z\} \implies T = \{\{xyz, xzy\}, \{yxz, yzx\}, \{zyx, zxy\}\}$$

- The (deterministic) mechanism that pays the revealed top element in  $X'$  is IC

## RPS mechanisms

- One can elicit top elements of several menus  $X_1, \dots, X_l \subseteq X$

Examples:

$$X_1 = \{x, y, z\}, X_2 = \{x, y\}$$

$$\implies T = \{\{xyz, xzy\}, \{yzx, yxz\}, \{zxy\}, \{zyx\}\}$$

$$X_1 = \{x, y\}, X_2 = \{x, z\}, X_3 = \{y, z\}$$

$$\implies T = \{\{xyz\}, \{yxz\}, \{yzx\}, \{zyx\}, \{zxy\}, \{xzy\}\}$$

- The corresponding IC mechanism randomly chooses a menu and pays the announced top element
- This is widely used in experimental economics

## RPS mechanisms

- One can elicit top elements of several menus  $X_1, \dots, X_l \subseteq X$

Examples:

$$X_1 = \{x, y, z\}, X_2 = \{x, y\}$$

$$\implies T = \{\{xyz, xzy\}, \{yzx, yxz\}, \{zxy\}, \{zyx\}\}$$

$$X_1 = \{x, y\}, X_2 = \{x, z\}, X_3 = \{y, z\}$$

$$\implies T = \{\{xyz\}, \{yxz\}, \{yzx\}, \{zyx\}, \{zxy\}, \{xzy\}\}$$

- The corresponding IC mechanism randomly chooses a menu and pays the announced top element
- This is widely used in experimental economics

**What else is elicitable?**

## Top sets of menus

The top-2 type space  $T = \{\{xyz, yxz\}, \{xzy, zxy\}, \{zyx, yzx\}\}$  does not reveal top elements of menus but is elicitable

## Top sets of menus

The top-2 type space  $T = \{\{xyz, yxz\}, \{xzy, zxy\}, \{zyx, yzx\}\}$  does not reveal top elements of menus but is elicitable

- How? If they announce “ $x$  and  $y$ ” pay  $x$  and  $y$  with equal probability, and  $z$  with less probability.
- Every  $X' \subseteq X$  and  $k$  defines a type space by

$$\succeq, \succeq' \in t \iff \succeq, \succeq' \text{ have the same top } k \text{ elements of } X'$$

- This is elicitable by paying the uniform lottery over the set of announced top- $k$  elements
- Can elicit the top- $k_i$  elements of  $X_i \subseteq X, i = 1, \dots, l$

## Top sets of menus

The top-2 type space  $T = \{\{xyz, yxz\}, \{xzy, zxy\}, \{zyx, yzx\}\}$  does not reveal top elements of menus but is elicitable

- How? If they announce “ $x$  and  $y$ ” pay  $x$  and  $y$  with equal probability, and  $z$  with less probability.
- Every  $X' \subseteq X$  and  $k$  defines a type space by

$$\succeq, \succeq' \in t \iff \succeq, \succeq' \text{ have the same top } k \text{ elements of } X'$$

- This is elicitable by paying the uniform lottery over the set of announced top- $k$  elements
- Can elicit the top- $k_i$  elements of  $X_i \subseteq X, i = 1, \dots, l$

Anything else??

## Example (based on Shapley, 1971)

$$X = \{x, y, z, w\}$$

Type space:

$$\{xyzw, yxzw, xywz, yxwz\}$$

$$\{xzyw\}, \{xwyz\}, \{xzwy, xwzy\}$$

$$\{ywxz\}, \{yzxw\}, \{yzwx, ywzx\}$$

$$\{zxyw, zyxw\}, \{zywx, zwyx\}, \{zxwy, zwxy\}$$

$$\{wxyz, wyxz\}, \{wyzx, wzyx\}, \{wxzy, wzxy\}$$

**Claim**

$\exists$  IC mechanism, but type space is not generated by top sets.

There is a close connection between IC mechanisms and convex TU cooperative games...

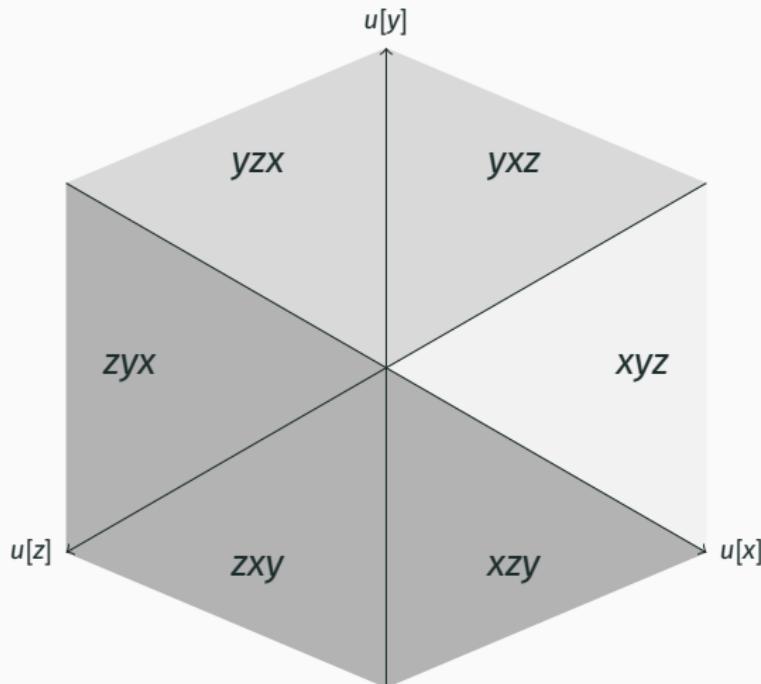
## So far...

$$\begin{aligned} & \{\text{all } T\} \\ & \cup \\ & \{T : \text{elicitable}\} \\ & \cup \\ & \{T : \text{generated by top sets}\} \\ & \cup \\ & \{T : \text{generated by top elements}\} \end{aligned}$$

# A convex type space - example

Necessary condition: **convex** type space

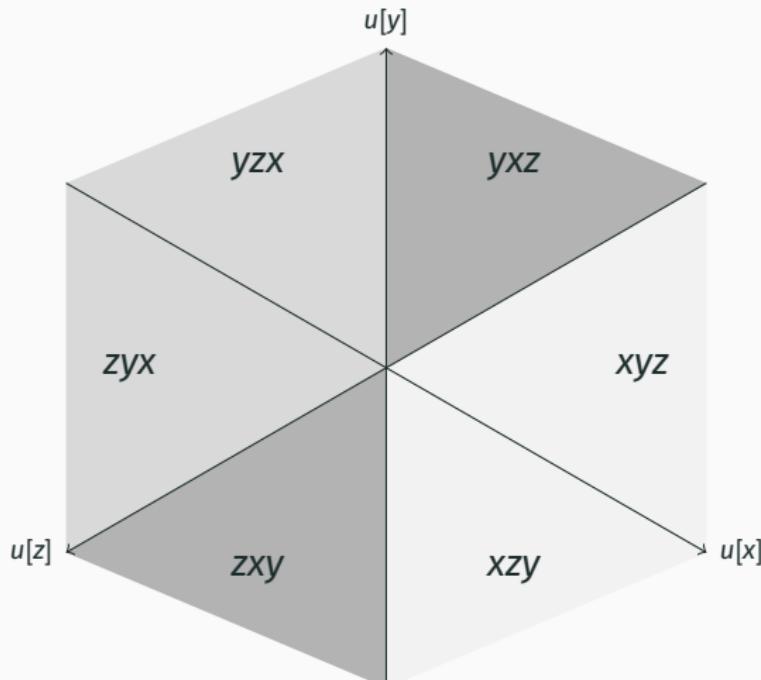
Example:  $T = \{\{xyz\}, \{yxz, yzx\}, \{zyx, zxy, xzy\}\}$



# A non-convex type space - example

Example of a non-convex type space:

$$T = \{\{xyz, xzy\}, \{yxz, zxy\}, \{zyx, yzx\}\}$$



# Convexity is necessary

## Proposition

If  $T$  is elicitable then it is convex.

Ex: Where do you rank  $x$ ?

$t = x$  is 2nd (dark gray)

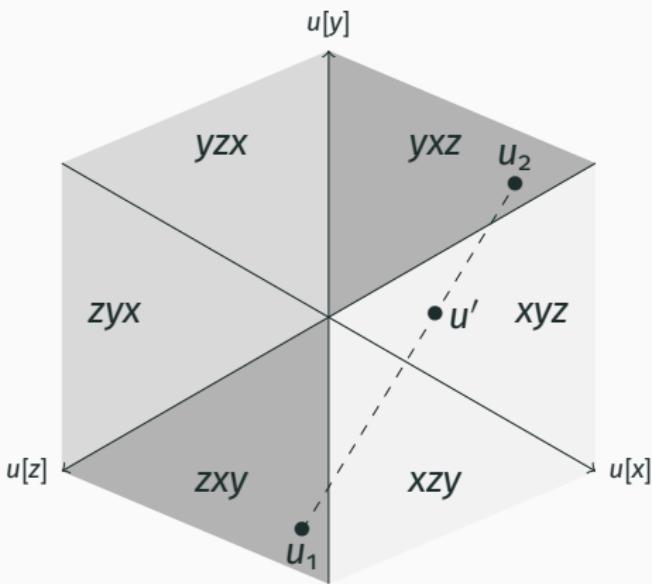
$t' = x$  is 1st (off-white)

IC requires:

$$\mathbb{E}_{g(t)}(u_1) > \mathbb{E}_{g(t')}(u_1)$$

$$\mathbb{E}_{g(t)}(u_2) > \mathbb{E}_{g(t')}(u_2)$$

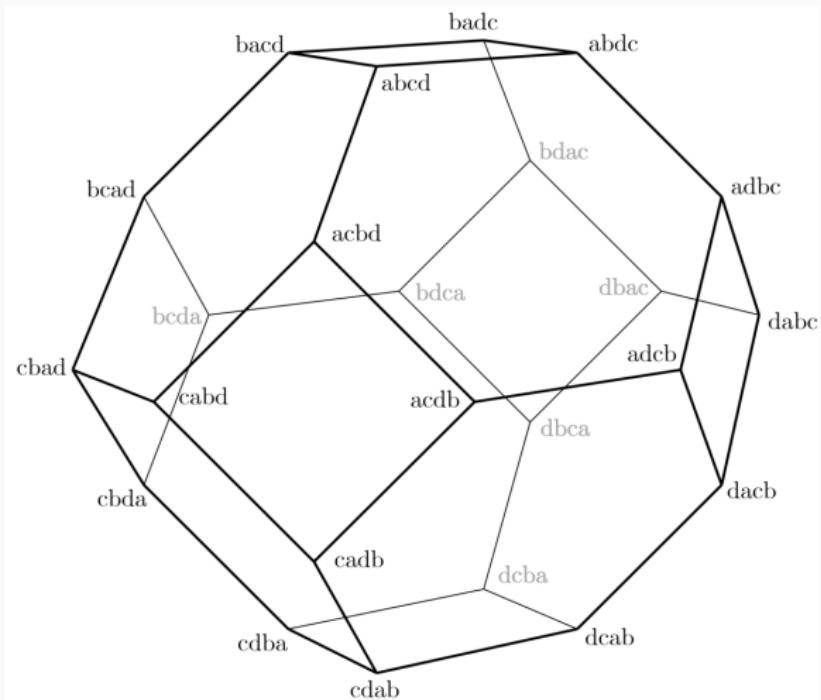
$$\implies \mathbb{E}_{g(t)}(u') > \mathbb{E}_{g(t')}(u')$$



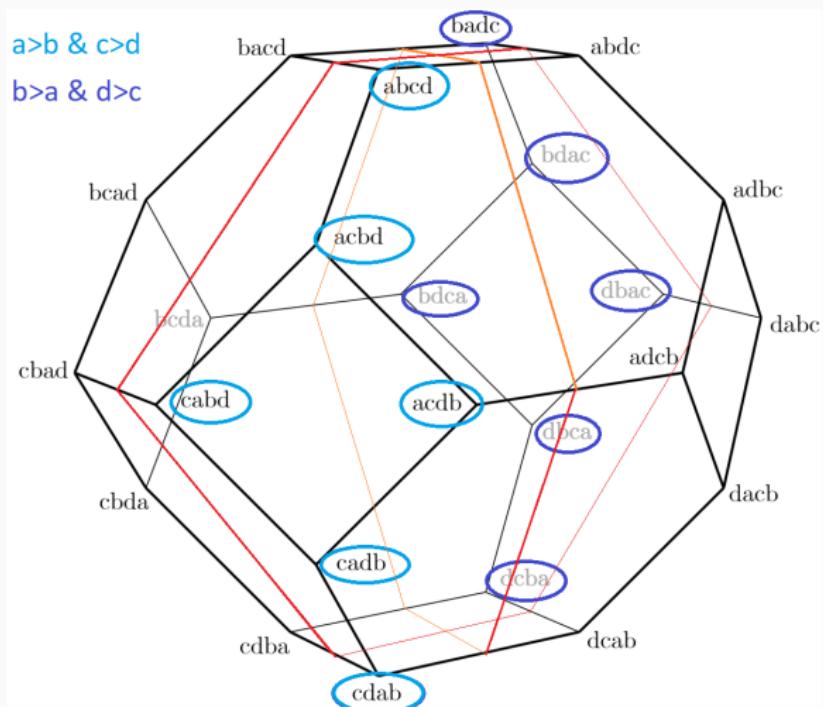
## Some non-convex type spaces

- Where do you rank  $x$ ? (with  $|X| \geq 3$ )
- What is the  $k$ th ranked alternative for  $1 < k < |X|$  (e.g. median)
- Any binary  $T = \{t_1, t_2\}$ , except  $T = \{\{x \succeq y\}, \{y \succeq x\}\}$ .  
In particular, tests of most axioms of preferences!  
Usually: “If  $x \succeq y$  then  $w \succeq z$  (and  $y \succeq x \Rightarrow z \succeq w$ )”

# Visualizing Convexity: The Permutohedron



# Visualizing Convexity: The Permutohedron



# Convexity is not sufficient

$$T = \{t_1 = \{xyz\}, t_2 = \{yxz, yzx\}, t_3 = \{zyx, zxy, xzy\}\}$$

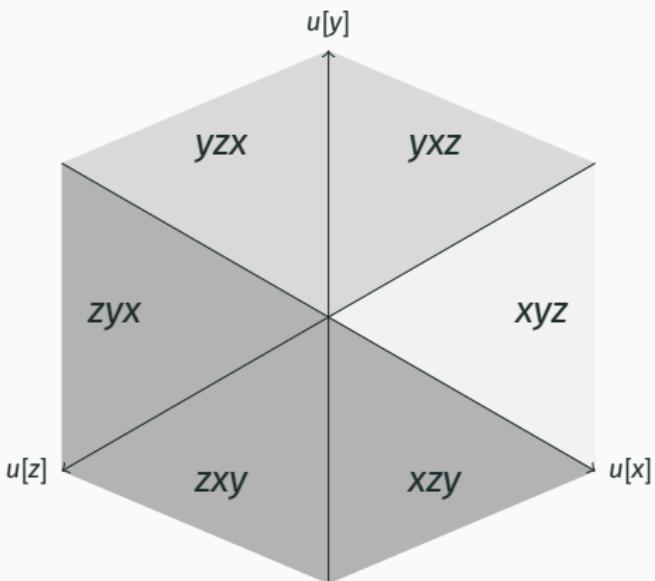
IC requires:

$$g(t_1)(x) > g(t_2)(x)$$

$$g(t_2)(x) = g(t_3)(x)$$

$$g(t_3)(x) = g(t_1)(x)$$

$$\implies g(t_1)(x) > g(t_1)(x)$$



# Summary

$$\begin{aligned} & \{ \text{all } T \} \\ & \cup \$ \\ & \{ T : \text{convex} \} \\ & \cup \$ \\ & \{ T : \text{no bad cycles} \} \\ & \cup | \\ & \{ T : \text{elicitable} \} \\ & \cup \$ \\ & \{ T : \text{generated by top sets} \} \\ & \cup \$ \\ & \{ T : \text{generated by top elements} \} \end{aligned}$$

## Neutral type spaces

- Permutation:  $\pi : X \rightarrow X$
- Let  $\pi T$  be  $T$ , but with every  $\succeq$  permuted by  $\pi$

### Definition

$T$  is *neutral* if  $\pi T = T$  for every  $\pi$ .

Neutral: “What do you rank 3rd?”

Not: “Where do you rank x?”

# Neutral type spaces

- Permutation:  $\pi : X \rightarrow X$
- Let  $\pi T$  be  $T$ , but with every  $\succeq$  permuted by  $\pi$

## Definition

$T$  is *neutral* if  $\pi T = T$  for every  $\pi$ .

Neutral: “What do you rank 3rd?”

Not: “Where do you rank x?”

## Proposition

Suppose  $T$  is neutral. Then the following are equivalent:

- (1)  $T$  is elicitable
- (2)  $T$  is convex
- (3)  $T$  is generated by top sets

# Neutral type spaces

$$\begin{aligned} & \{\text{all } T\} \\ & \cup \text{\texttt{\$}} \\ & \{T : \text{convex}\} \\ & \parallel \\ & \{T : \text{elicitable}\} \\ & \parallel \\ & \{T : \text{generated by top sets}\} \\ & \cup \text{\texttt{\$}} \\ & \{T : \text{generated by top elements}\} \end{aligned}$$

## Robust elicitation

What if the agent has subjective beliefs about the likelihood of realizations of the randomization device (or other kinds of preferences over acts)?

## Robust elicitation

What if the agent has subjective beliefs about the likelihood of realizations of the randomization device (or other kinds of preferences over acts)?

- Now payments are Savage acts, not lotteries.
- Can't pay a uniform lottery b/c can't control beliefs
- This kills our ability to elicit top sets

## Robust elicitation

What if the agent has subjective beliefs about the likelihood of realizations of the randomization device (or other kinds of preferences over acts)?

- Now payments are Savage acts, not lotteries.
- Can't pay a uniform lottery b/c can't control beliefs
- This kills our ability to elicit top sets

### Proposition

*T is elicitable with acts iff it is generated by top elements.*

# Elicitation under acts

{all  $T$ }

$\cup \$$

{ $T$  : convex}

$\cup \$$

{ $T$  : elicitable with lotteries}

$\cup \$$

{ $T$  : generated by top sets}

$\cup \$$

{ $T$  : generated by top elements}

||

{ $T$  : elicitable with acts}

## Multiple agents

- $N = \{1, \dots, n\}$  - agents
- $T_i$  - agent's  $i$  type space
- $T = (T_1, \dots, T_n)$  - a profile of type spaces
- $g : T \rightarrow \Delta(X)$  - a mechanism

## Multiple agents

- $N = \{1, \dots, n\}$  - agents
- $T_i$  - agent's  $i$  type space
- $T = (T_1, \dots, T_n)$  - a profile of type spaces
- $g : T \rightarrow \Delta(X)$  - a mechanism

### Proposition

$T = (T_1, \dots, T_n)$  is dominant-strategy-elicitable iff each  $T_i$  is elicitable.

# Conclusion

- We formulate a notion of elicability for properties of preferences
- Some necessary conditions and some sufficient conditions for elicability, but no characterization
- We do have a characterization for neutral type spaces and for robust elicitation (acts)
- Potential extensions: Weak orders, infinite sets of alternatives, domain restrictions,...

# Thank You!