DISENTANGLING STRICT AND WEAK CHOICE IN RANDOM EXPECTED UTILITY MODELS

Evan Piermont & Roee Teper

September 2019

Economic choice data is usually aggregated across

- many subjects, or,
- many different points in time, or both

For a choice problem: $D = \{x, y\}$

The analyst observes a **Random Choice Rule** (RCR): $\mu_D(x)$ representing the frequency x is chosen from D.

Economic choice data is usually aggregated across

- many subjects, or,
- many different points in time, or both

For a choice problem: $D = \{x, y\}$

The analyst observes a **Random Choice Rule** (RCR): $\mu_D(x)$ representing the frequency x is chosen from D.

Can we still falsify individual rationality / expected utility?

Random Utility Models are a way of dealing with aggregated choice data:

Random Utility Models are a way of dealing with aggregated choice data:

Let $\mathcal{U}$ denote a set of (linear) utility functions. Then $\xi \in \Delta(\mathcal{U})$ is a **Random Expected Utility Model** (REU) representing μ if

$$\mu_D(x) = \xi\{u \in \mathcal{U} \mid \arg\max_D u = x\}$$

Random utility models do not deal well with ties:

- If with positive ξ-probability $u(x) = u(y)$, what is $\mu_D(x)$?
- Depends on how indifferences are broken.

Gul and Pesendorfer (2006) (and the following literature) assume ties occur with probability 0.

- Arises from the **extremeness axiom**: only the extreme points of a decision problem are ever chosen.
- Implicit that $\mathcal{U}$ is infinite, ξ is non-atomic, full dimensional.
- μ uniquely determines ξ and vice-versa.

In this paper, we consider RCRs without any axiomatic restrictions and

1. characterize the set of REUs that could have generated them under *some* tie-breaking procedure,
2. outline the relation between properties of tie-breaking rules and properties of the RCRs, and
3. study the REUs that rely on indifference the least in order to explain the observed data.

We do this via a the theory of **Choice Capacities**, the non-additive counterpart to RCRs.

- There is a bijection between CCs and (all) REUs.
- The CCs that dominate μ correspond to those that could have generated it.
- There is a natural ordering on CCs representing “how much tie-breaking.”

Explaining data with less tie-breaking is always better for allocation problems:

- Modelers assumes $x \succ y$ when $x \sim y$,
  - the provision of x rather than y is not harmful.
- Modelers assumes $x \sim y$ when $y \succ x$,
  - allocating x is costly to the agents welfare.

Literature

- Lu (2016) extends GP methodology to allow ties with probability 0 or probability 1
- Frick et al. (2017) entertain a dynamic environment; assume indifference is a null event.
- Ahn and Sarver (2013) ; Gul and Pesendorfer (2006) consider tie-breaking:
  - the tie breaking procedure is often assumed to be an regular REU with 0-probability ties
  - that the resulting choice data is as in GP
- Lin (2018) studies random set valued choice—equivalent to CCs.

Set Up

$\mathcal{D}$ collects **decision problems**, finite non-empty subsets of R^n.

A **choice capacity** is a family of functions

$$\rho = \{\rho_D : 2^D \to [0, 1]\}_{D \in \mathcal{D}}$$

that are

- **grounded**: $\rho_D(\emptyset) = 0$.
- **normalized**: $\rho_D(D) = 1$.
- **monotone**: $\rho_D(A \cup B) \geq \rho_D(A)$.

An *additive* CC is called a **Random Choice Rule** and is denoted by μ.

Random Expected Utility

Call ξ, a probability measure over $\mathbb{R}^n$, a **random expected utility model** (REU).

- ξ is **regular** if $\xi(\{u \in \mathbb{R}^n \mid \#(\arg\max_{y \in D} u \cdot y) = 1\}) = 1$

Say that ρ **maximizes** ξ if

$$\rho_D(A) = \xi(\{u \in \mathbb{R}^n \mid A \cap (\arg\max_{y \in D} u \cdot y) \neq \emptyset\})$$

for all (D, A).

Theorem

Every REU has a unique maximizer and every ρ maximizes at most one REU. Moreover, ρ maximizes a regular ξ iff ρ is additive.

- Let ρ^{ξ} denote the CC that maximizes ξ.

GP axioms

A1. **Monotonicity**: $D \subseteq D' \implies \rho_D(a) \geq \rho_{D'}(a)$.

A2. **Extremeness**: $\text{ext}(D) = \text{ext}(D')$, implies $\rho_D(A) = \rho_{D'}(A)$.

A3. **Linearity**: $\rho_{\lambda D+z}(\lambda A + z) = \rho_D(A)$ for $\lambda > 0$.

A4. **Mixture Cont**: For $D, D' \in \mathcal{D}$, $\rho_{\lambda D+\lambda' D'}$ is continuous in λ, λ' for $\lambda, \lambda' \geq 0$.

A4'. **U-Cont**: For $\{D_n\}_{n\in\mathbb{N}} \to D$, $\limsup \rho_{D_n}(C) \leq \rho_D(C)$ for closed $C \subseteq \mathbb{R}^n$

Theorem (GP)

An additive RCR μ satisfies **Mon, Ext, Lin, MxCont** if and only if it maximizes a finitely additive regular REU ξ.

Moreover, μ additionally satisfies **U-Cont** if and only if ξ is countably-additive.

Tie Breaking Rules

A **tie-breaking rule** is a set of measures $\tau = \{\tau^D_{D'}\}_{D \in \mathcal{D}, D' \subseteq D}$ where $\text{supp}(\tau^D_{D'}) = D^{'}$.

- $\tau^D_{D'}(x)$ frequency x is chosen when $\arg\max_D = D^{'}$.

We say μ is **consistent** with (ξ, τ) if

$$\mu_D = \int_{R^n} \tau^D_{\arg\max_D u} \xi(du)$$

and consistent with ξ is there is such a τ.

Example

$D = \{a, b\}$ and $D^{'} = \{a, b, c = \frac{1}{2}a + \frac{1}{2}b\}$

$$\mu_D : \begin{cases} a & \mapsto \frac{2}{3} \\ b & \mapsto \frac{1}{3} \end{cases} \qquad \mu_{D'} : \begin{cases} a & \mapsto \frac{1}{2} \\ b & \mapsto \frac{1}{4} \\ c & \mapsto \frac{1}{4} \end{cases}$$

This RCR does not satisfy GP's extremeness axiom and therefore does not maximize any *regular* random expected utility model.

Example

- Let ξ be given by

$$\xi([1,0]) = \frac{1}{2}, \qquad \xi([-1,0]) = \frac{1}{4}, \qquad \xi([0,0]) = \frac{1}{4}$$

- Let τ be a tie breaking rule—itself a random choice rule—such that $\tau_{\{a,b\}}(a) = \frac{2}{3}$ and $\tau_{\{a,b,c\}}(c) = 1$.
- μ is consistent with (ξ, τ). For example

$$\mu_D(a) = \frac{2}{3} = \frac{1}{2} + \frac{1}{4} \cdot \frac{2}{3} = \xi([1,0]) + \xi([0,0]) \cdot \tau_{\{a,b\}}(a)$$

Theorem

An RCR μ satisfies any subset of $\{$**Mon, Ext, Lin, MxCont**$\}$ if and only if it is consistent with some (ξ, τ) where τ itself an RCR that satisfies the same subset.

This result helps establish falsification, but does not guide the construction of REUs from observable data.

For this, we turn to general CCs.

Example, Still

Let ξ be from from earlier

$$\xi([1,0]) = \frac{1}{2}, \qquad \xi([-1,0]) = \frac{1}{4}, \qquad \xi([0,0]) = \frac{1}{4}$$

We can construct ρ^{ξ}, according to the definition of maximization:

$$\rho^{\xi}_{D}(A) = \xi(\{u \in \mathbb{R}^n \mid A \cap (\arg\max_{y \in D} u \cdot y) \neq \emptyset\})$$

Example, Still

$$\arg\max_{D'}(u_1) = \{a\}, \quad \arg\max_{D'}(u_2) = \{b\}, \quad \arg\max_{D'}(u_3) = \{a, b, c\}$$

$$\begin{aligned}
\rho^{\xi}_{D'}(\{a, b\}) = \rho^{\xi}_{D'}(\{a, b, c\}) &= 1 \\
\rho^{\xi}_{D'}(a) = \rho^{\xi}_{D'}(\{a, c\}) &= \frac{3}{4} \\
\rho^{\xi}_{D'}(b) = \rho^{\xi}_{D'}(\{b, c\}) &= \frac{1}{2} \\
\rho^{\xi}_{D'}(c) &= \frac{1}{4}
\end{aligned}$$

For any $A \subseteq D$, we have $\rho^{\xi}_{D}(A) = \rho^{\xi}_{D'}(A)$.

A5. **Convex Modularity:** Let $A, B \subseteq D$ be such that $\alpha A + (1-\alpha)B \subseteq D$ for $\alpha \in (0,1)$. Then

$$\rho_D(\alpha A + (1-\alpha)B) = \rho_D(A) + \rho_D(B) - \rho_D(A \cup B).$$

- Controls 'how' non-additive ρ can be.
- $\alpha A + (1-\alpha)B$ chosen iff indifferent between A and B.
- Satisfied by all additive RCRs.

Example, Still

We can verify that ρ^ξ satisfies convex modularity:

$$\rho^\xi_{D'}(a) + \rho^\xi_{D'}(b) - \rho^\xi_{D'}(\{a, b\}) =$$
$$\frac{3}{4} + \frac{1}{2} - 1 =$$
$$\frac{1}{4} =$$
$$\rho^\xi_{D'}(c)$$

Example, Still

We can verify that ρ^ξ satisfies convex modularity:

$$\xi(\{u_1, u_3\}) + \xi(\{u_2, u_3\}) - \xi(\{u_1, u_2, u_3\}) =$$
$$\frac{3}{4} + \frac{1}{2} - 1 =$$
$$\frac{1}{4} =$$
$$\xi(\{u_3\})$$

Theorem

An CC ρ satisfies **Mon, Ext, Lin, MxCont** and **CvxMod** if and only if it maximizes a finitely additive REU ξ.

Moreover, ρ additionally satisfies **U-Cont** if and only if ξ is countably-additive.

Why do we care about CCs, if they are not observable?

Example, Still

Notice also that $\rho^{\xi}_{D'}(A) \geq \mu_{D'}(A)$:

$$\rho^{\xi}_{D'}(\{a, b, c\}) = 1 \geq 1 = \mu^{\xi}_{D'}(\{a, b, c\})$$

$$\rho^{\xi}_{D'}(\{a, b\}) = 1 \geq \frac{3}{4} = \mu^{\xi}_{D'}(\{a, b\})$$

$$\rho^{\xi}_{D'}(\{a, c\}) = \frac{3}{4} \geq \frac{3}{4} = \mu^{\xi}_{D'}(\{a, c\})$$

$$\rho^{\xi}_{D'}(\{b, c\}) = \frac{1}{2} \geq \frac{1}{2} = \mu^{\xi}_{D'}(\{b, c\})$$

$$\rho^{\xi}_{D'}(\{a\}) = \frac{3}{4} \geq \frac{1}{2} = \mu^{\xi}_{D'}(\{a\})$$

$$\rho^{\xi}_{D'}(\{b\}) = \frac{1}{2} \geq \frac{1}{4} = \mu^{\xi}_{D'}(\{b\})$$

$$\rho^{\xi}_{D'}(\{c\}) = \frac{1}{4} \geq \frac{1}{4} = \mu^{\xi}_{D'}(\{c\})$$

- μ is consistent with (ξ, τ):
    - x was a maximizer of D **at least** $\mu_D(x)$.
- ρ is is the CC maximizing ξ:
    - x was a maximizer of D **exactly** $\rho_D(x)$.

$$\Gamma(\xi, D) = \left\{ \int_{R^n} \tau_u(\cdot)\xi(du) \mid \tau_u \in \Delta\Big(\arg\max_{y\in D} u(y)\Big) \right\}.$$

- the set of all possible choice rules constructed by first choosing a utility u according to ξ, and subsequently choosing among the maximizers in D according to some tie breaking procedure.

Theorem

$\rho^{\xi}_{D}(A) = \sup_{\gamma\in\Gamma(\xi,D)} \gamma(A)$ for all D.

Theorem

An RCR μ is consistent with ξ if and only if $\rho_D^{\xi}(A) \geq \mu_D(A)$ for every (D, A).

Further, the set of REUs consistent with μ is non-empty, convex, and compact.

- To ensure consistency with ξ, we only need to check $\rho^{\xi} \geq \mu$.
- This characterizes all REUs consistent with the data.
- This can be done constructively (our axiomatic representation result is constructive).

The set of consistent REUs is generally large. How do we make sense of it?

- Set is ‘upwards closed.’
- Consider the minimal elements.

For an RCR μ, call ξ **μ-minimal** if (i) ξ is consistent with μ and (ii) ρ^{ξ} does not point-wise dominate ρ^{ζ} for any ζ consistent with μ.

- μ-minimal CCs explain the data with the least reliance on indifference.
- The set of μ-minimal REUs is non-empty (follows from compactness).
- May not be unique.

Let $c = \frac{1}{2}a + \frac{1}{2}b$. Consider $D = \{a, b, c\}$. Then

- $\rho^{\xi}_{D}(c)$ is the ξ-probability of a u such that $u(a) = u(b)$.
- So if $\rho^{\xi}_{D}(c) \geq \rho^{\zeta}_{D}(c)$, then ξ yields indifference between a and b more often than ζ.

When ξ is the unique μ-minimal REU, then ξ explains the observed choices and any other explanation must realize every kind of indifference weakly more often.

Remark

If ξ is the unique μ-minimal REU and ζ is consistent with μ, then ρ^ζ dominates ρ^ξ.

Example, Still, Continued

Taking μ from earlier in the example:

$$\mu_D : \begin{cases} a & \mapsto \frac{2}{3} \\ b & \mapsto \frac{1}{3} \end{cases} \qquad \mu_{D'} : \begin{cases} a & \mapsto \frac{1}{2} \\ b & \mapsto \frac{1}{4} \\ \frac{1}{2}a + \frac{1}{2}b & \mapsto \frac{1}{4} \end{cases}$$

Consider the CC, ρ, that assigns to each element the **minimal** probability it was a maximizer consistent with all the observed data.

For example, what is the minimal probability that a *could be* chosen from $D = \{a, b\}$? It is $\frac{3}{4}$, despite the fact that it is not chosen this frequently.

Example, Still, Continued

We have

$$
\begin{aligned}
&\rho_D(a) = \rho_{D'}(a) = \mu_{D'}(a) + \mu_{D'}(c) = \frac{3}{4},\\
&\rho_D(b) = \rho_{D'}(b) = \mu_{D'}(b) + \mu_{D'}(c) = \frac{1}{2},\\
&\rho_{D'}(c) = \mu_{D'}(c) = \frac{1}{4},\\
&\rho_D(\{a, b\}) = \rho_{D'}(\{a, b\}) = \mu_D(\{a, b\}) = 1
\end{aligned}
$$

ρ coincides with ρ^{ξ} constructed earlier.

Example, Still, Continued

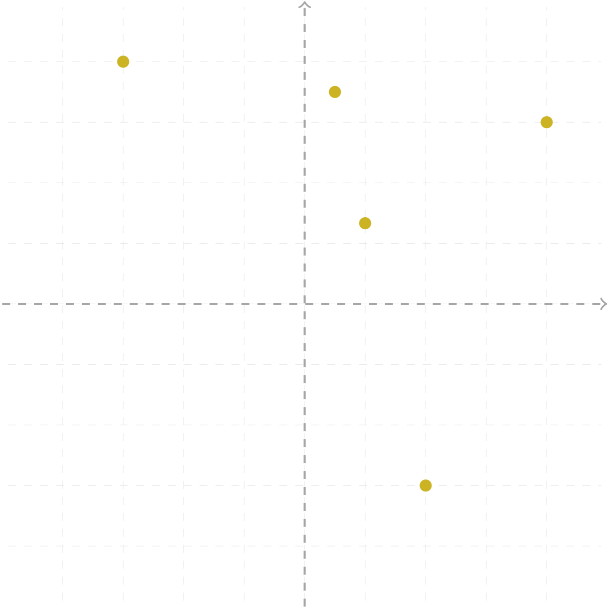
Notice that this also shows that ξ is uniquely μ-minimal, since $\rho^{\xi}_{D'}(c) = \mu_{D'}(c)$, so any ξ which yields ties between a and b less often would not be consistent with μ.

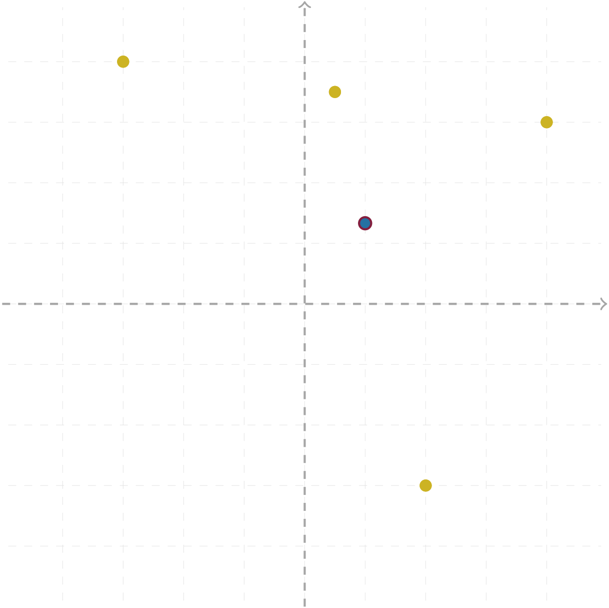
Let $\mathsf{pi}(D, A) =$

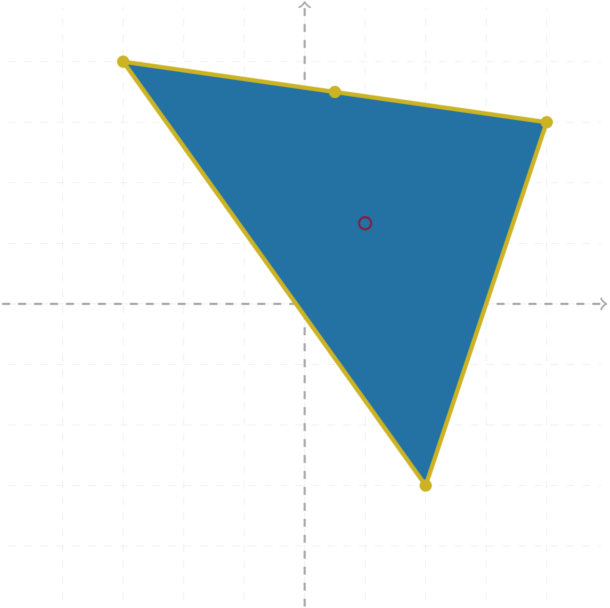
$$\{x \in \mathsf{conv}(D) \mid x = \alpha a + (1-\alpha) y, a \in A, y \in \mathsf{conv}(D), \alpha \in (0, 1]\}$$

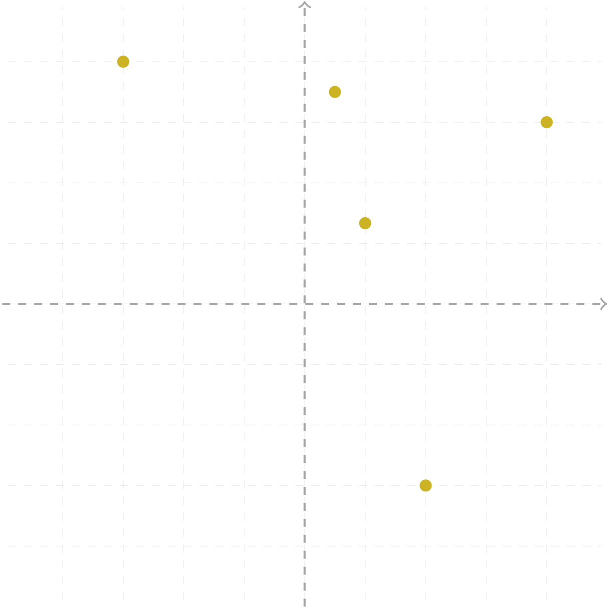
denote the **projective interior** of A in D.

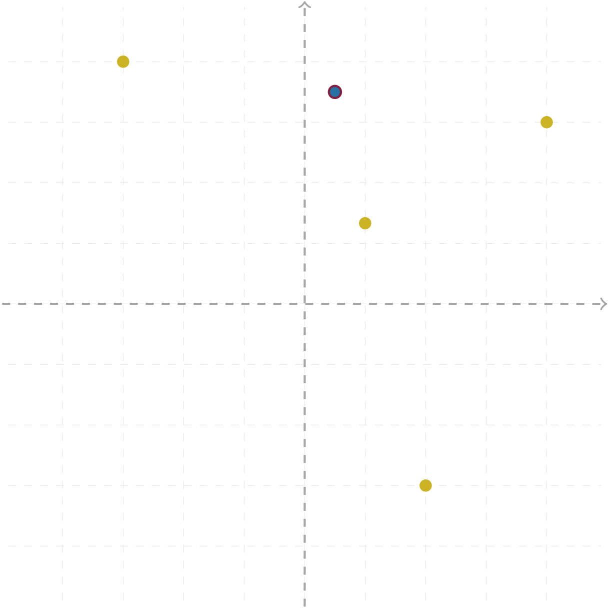
- $\mathsf{pi}(D, A)$ is the union of the relative interiors of all faces intersecting A.
- If $x \in \mathsf{pi}(D, A)$ is chosen, then something in A is maximal.

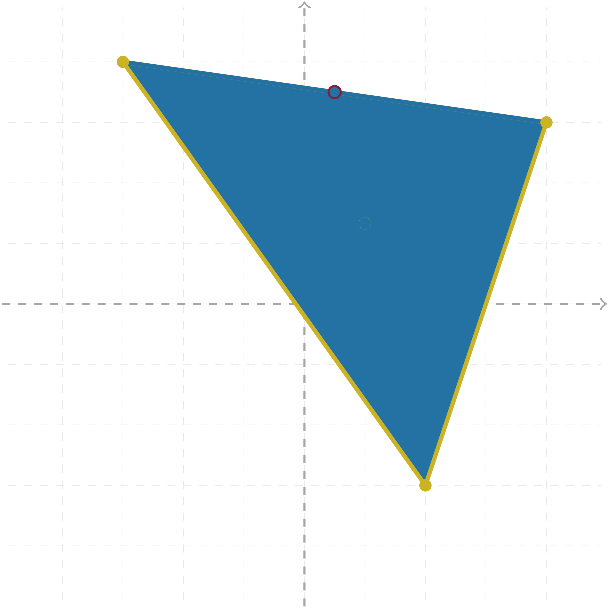


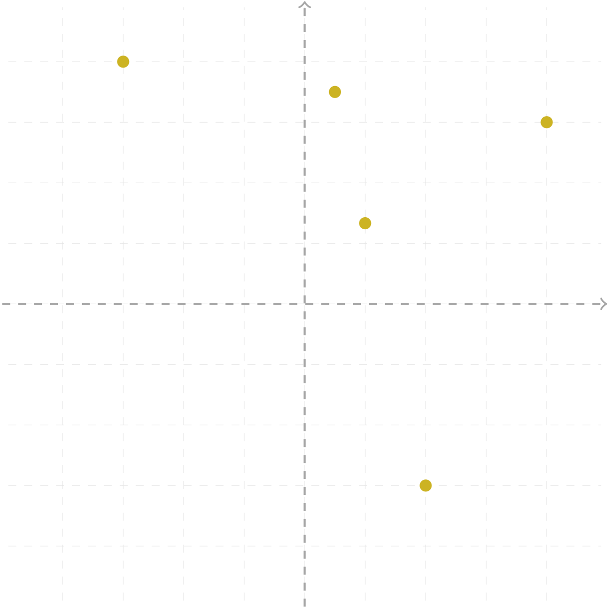


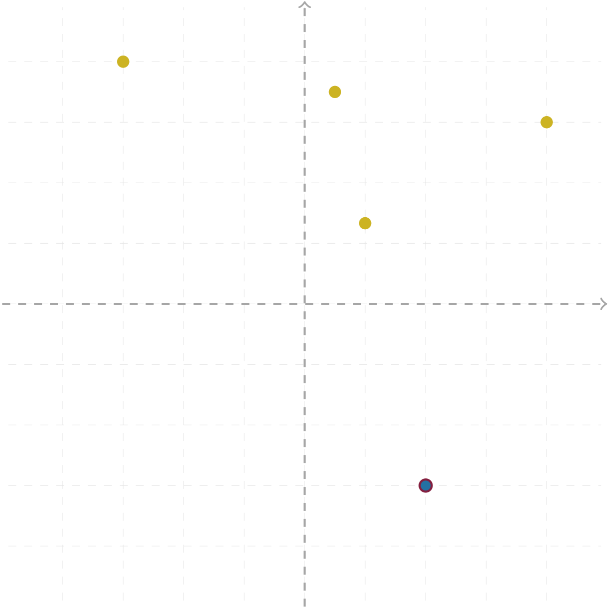


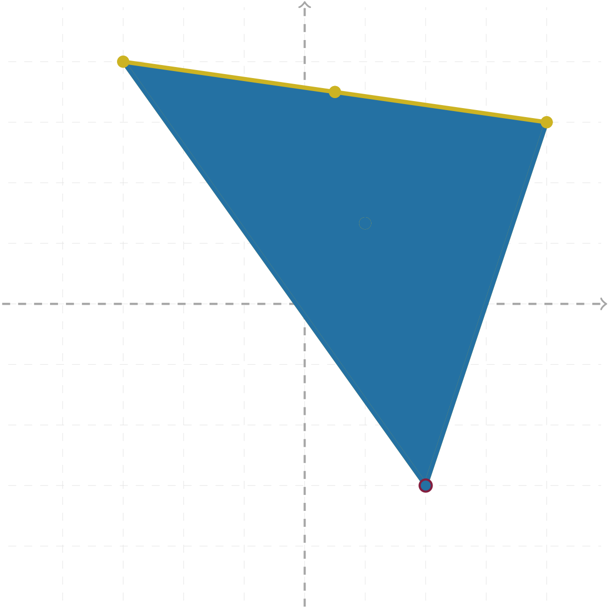












Let $CV(D)$ denote the set of decision problems with the same convex hull as D. Then set:

$$\rho_D^\mu(A) = \sup_{D' \in CV(D)} \mu_{D'}(\mathsf{pi}(D, A)).$$

- This was the procedure from the example.
- The *least* upper bound on A being a maximizer.

Remark

ρ^μ satisfies **Mon, Ext, Lin, MxCont, CvxMod** if and only if it maximizes the unique μ-minimal REU.

- But when does this happen?

Call τ **lexicographically risk averse** if for all $\epsilon > 0$, $D \in \mathcal{D}$ and $A \subseteq D$, there exists a $D^{'} \in CV(D)$ such that

$$\tau_{D^{''}}^{D^{'}}(\mathrm{pi}(D, A)) \geq 1 - \epsilon,$$

for all $D^{''} \subseteq D^{'}$ such that $D^{''} \cap A \neq \emptyset$.

- If the elements of A are all maximizers, there is some decision problem where indifference is broken in favor of pi(A).

Call τ **strongly lexicographically risk averse** if for all $D \in \mathcal{D}$ and $D^{'} \subseteq D$,

$$\tau_{D'}^{D}(\mathrm{ri}(D^{'})) = 1,$$

whenever $\mathrm{ri}(D^{'}) \cap D \neq 0$.

Call τ **uniform** if

$$\tau_{D'}^{D}(x) = \frac{1}{\#(D')}$$

for all $x \in D^{'}$.

Remark

If τ is strongly lexicographically risk averse or uniform it is lexicographically risk averse.

Theorem

An RCR μ is consistent with ξ and some lexicographically risk averse tie breaking rule, τ, if and only if ρ^{μ} is the CC that maximizes ξ.

In such cases ξ is the unique μ-minimal REU.