UNDERSTANDING INFORMATION ACQUISITION THROUGH F-INFORMATIVITY AND DUALITY

Luciano Pomatto, Caltech

Choice and Decision Workshop, BSE Summer Forum --- June 2025

~~UNDERSTANDING INFORMATION ACQUISITION THROUGH F-INFORMATIVITY AND DUALITY~~

MODELING THE MODELER:

A NORMATIVE THEORY OF EXPERIMENTAL DESIGN

~~Luciano Pomatto, Caltech~~

Fernando Payró
UAB, BSE, & CODE

Evan Piermont
Royal Holloway, University of London

Choice and Decision Workshop, BSE Summer Forum --- June 2025

- We propose three normative principles for experimental design
    - minimal rationality properties, independent of specific motivations
    - We will specifically think about revealed preference experiments
- We show that they imply a particular representation
    - Relates a experiment to the expected value of identification
    - Unifies many distinct models of experimentation
    - Axiomatic characterization for Bayesian Experimental Design
    - Test for analyst to make sure they do not have an “agenda”

Normative Principles

Structural Invariance: Two experiments that identify the sets of parameters are equally valued

Information Monotonicity: Experiments that induce sharper identification are (weakly) better

Identification Separability: The value of identifying a set of parameters should *not* depend on counterfactuals

Model

An analyst (she) wishes to infer a subject's (he) utility function over Z:

- Revealed Preference: she should offer a menu $A \subseteq Z$ and observe the subject's choice
- Different menus offer different "inference" opportunities
- Ranking over menus will depend on the goals for the analyst

Experimental Environment

- Z set of alternatives
- $\mathcal{U} \subseteq \{u : Z \to \mathbb{R}\}$ set of utility functions over Z
- Ω algebra of measurable sets of $\mathcal{U}$
- μ prior over $(\mathcal{U}, \Omega)$

The tuple $(Z, \mathcal{U}, \Omega, \mu)$ constitutes a theory for a Bayesian experimenter

Expected Identification Value

These principles characterize ranking according to *expected identification value*

- Exists some τ: for $W \subseteq \mathcal{U}$, $\tau(W)$ is the value of identifying W
- Let $\mathcal{W}$ be the sets that might get identified
- Experiments are valued according to:

$$\sum_{W \in \mathcal{W}} \tau(W)\mu(W)$$

- where μ is the (exogenous) prior probability

Special Case: Entropy

$$\tau(W) = -\log(\mu(W))$$

- Value of experiment is expected reduction in entropy

Special Case: Hypothesis Testing

$$\tau(W) = \begin{cases} 1 & \text{if } W \subseteq W^* \text{ or } W^* \subseteq W^c \\ 0 & \text{otherwise .} \end{cases}$$

- Hypothesis: the true utility lies in W^*
- Value of exp is the probability the hypothesis can be accepted or rejected

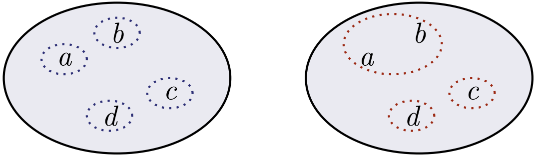
Special Case: Actions

$$\tau(W) = \max_{a \in \mathbb{A}} \int_W \xi(a, u)\, d\mu.$$

- The analyst will take action $a \in \mathbb{A}$
- Utility of outcome depends on the utility: $\xi(a, u)$
- Value of exp is expected value of conditionally optimal action

An **experiment** $e = (A, \mathscr{P})$ is a pair:

- $A \subseteq Z$ is finite decision problem
- $\mathscr{P}$ is a partition of A



- Represents observability constraints
- Allows for dynamic experiments, non-lab settings, etc

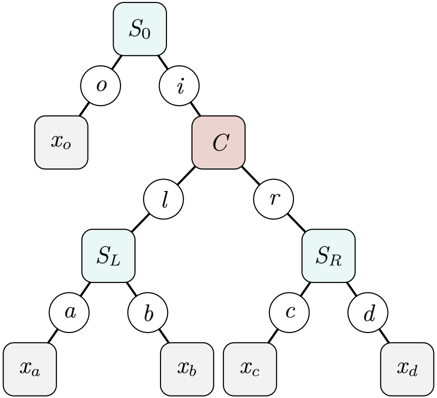
Example: Dynamic Games

Consider a dynamic game:

- Computer randomizes 50-50
- Subject's strategies:

$$A = \left\{ \begin{array}{l} (i, a, c), (i, b, c), (i, a, d), (i, b, d), \\ (o, a, c), (o, a, d), (o, b, c), (o, b, d) \end{array} \right\}$$

- There are observable paths: (o), (i, a), (i, b), (i, c), and (i, d)

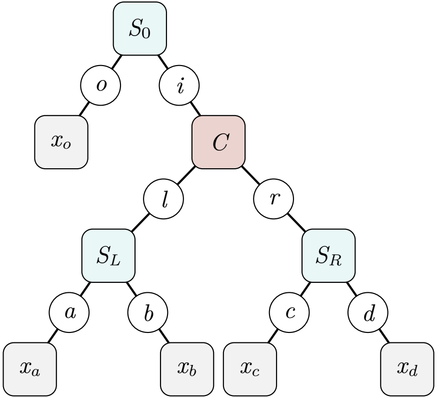


Example: Dynamic Games

Now consider the following partitions of A

$$\mathscr{P}_L = \left\{ \begin{array}{c} \left\{ \begin{array}{c} (o,a,c),(o,b,c), \\ (o,a,d),(o,b,d) \end{array} \right\}, \\ \left\{ \ (i,a,c),(i,a,d) \ \right\}, \\ \left\{ \ (i,b,c),(i,b,d) \ \right\} \end{array} \right\}$$

$$\mathscr{P}_R = \left\{ \begin{array}{c} \left\{ \begin{array}{c} (o,a,c),(o,b,c), \\ (o,a,d),(o,b,d) \end{array} \right\}, \\ \left\{ \ (i,a,c),(i,b,c) \ \right\}, \\ \left\{ \ (i,a,d),(i,b,d) \ \right\} \end{array} \right\}$$



Given an experiment, $(A, \mathscr{P})$, define the ***identified set***:

$$W_{A,P} = \{u \in \mathcal{U} \mid P \cap \text{argmax}_{x \in A} u(x) \neq \varnothing\}$$

- Observing $P \in \mathscr{P}$ identifies that the subject's utility is in $W_{A,P}$
- We require for an experiment $(A, \mathscr{P})$ that for any $P, Q \in \mathscr{P}$

  (1) $W_{A,P} \in \Omega$ — measurability

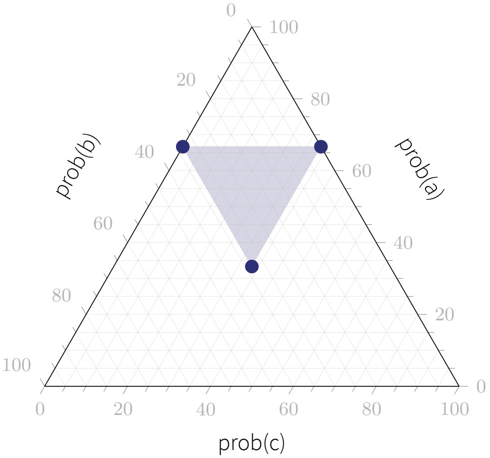
  (2) $\mu(W_{A,P} \cap W_{A,Q}) = 0$ — zero μ-prob of ties

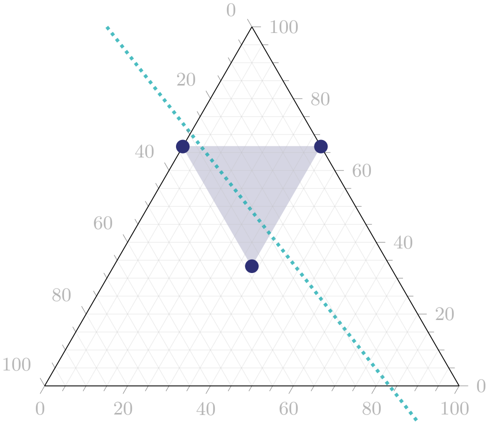
Example: EU Preferences

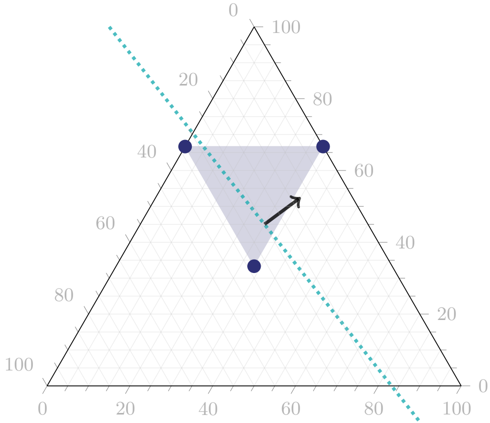
We can example identifying EU preferences as an example:

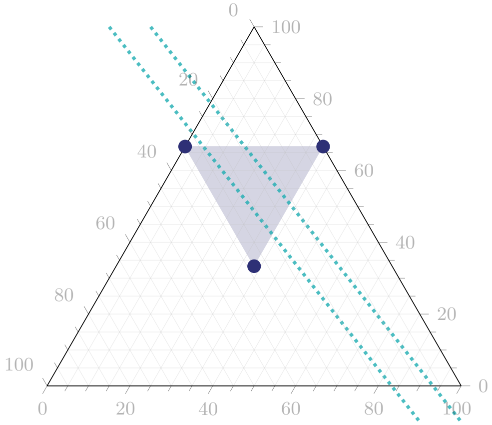
- Z is lotteries over $\{a, b, c\}$
- $\mathcal{U}$ is affine functions

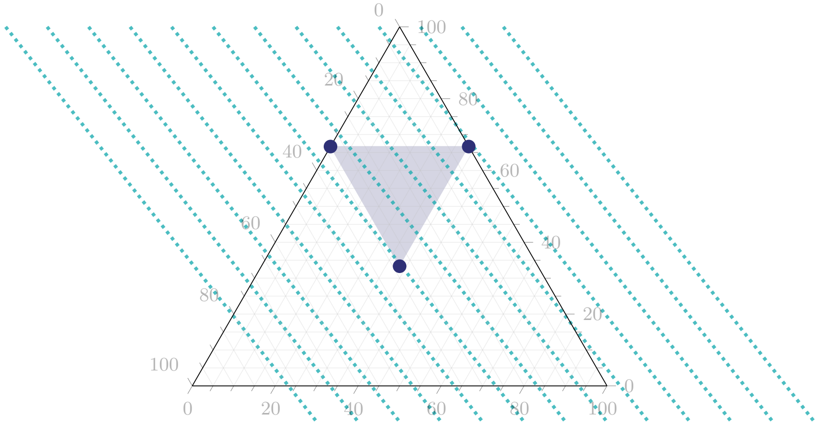
$$\{\tfrac{2}{3}a + \tfrac{1}{3}b, \tfrac{2}{3}a + \tfrac{1}{3}c, \tfrac{1}{3}a + \tfrac{1}{3}b + \tfrac{1}{3}c\}$$



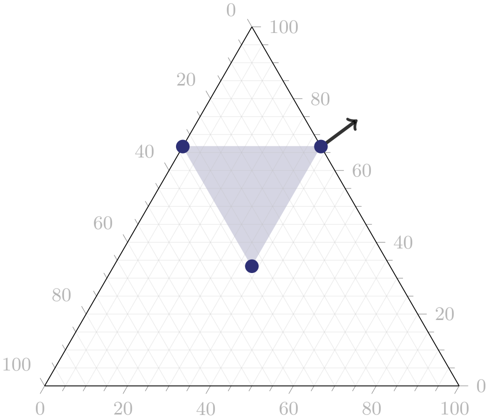

0
20
40
60
80
100
100
80
60
40
20
0
0
20
40
60
80
100

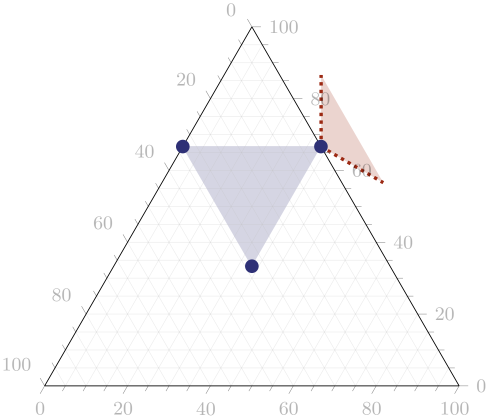

0
20
40
60
80
100
100
80
60
40
20
0
0
20
40
60
80
100

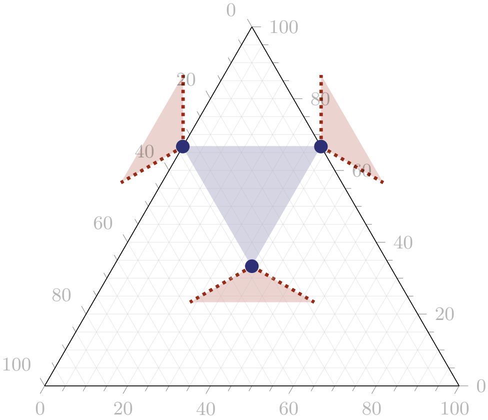

0
20
40
60
80
100
100
80
60
40
20
0
0
20
40
60
80
100



0
20
40
60
80
100
100
80
60
40
20
0
0
20
40
60
80
100

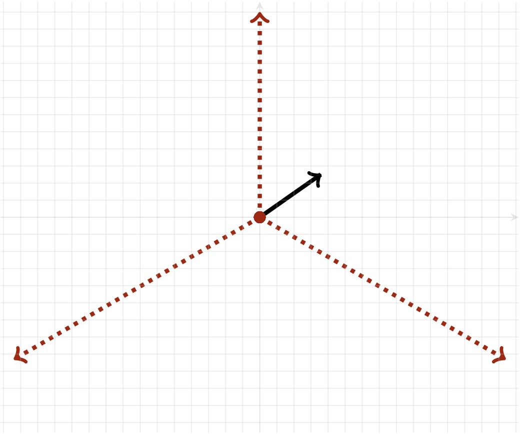

0
20
40
60
80
100
100
80
60
40
20
0
0
20
40
60
80
100


0
20
40
60
80
100
100
80
60
40
20
0
0
20
40
60
80
100



0
20
40
60
80
100
100
80
60
40
20
0
0
20
40
60
80
100

$\mathcal{U} \cong \mathbb{R}^2$



μ-equivalence

Call $\{W_1, \ldots, W_n\}$ and $\{V_1, \ldots, V_m\}$, families of subsets of $\mathcal{U}$, *μ-equivalent* if for all W_i and V_j:

$$\mu(W_i) = \mu(W_i \cap V_j) \text{ for some } j \quad \text{and,}$$
$$\mu(V_j) = \mu(W_i \cap V_j) \text{ for some } i$$

- Such collections identify the same sets of utilities up to a measure zero
- Take $[0, 1]$ with λ the Lebesgue measure. The following are λ-equivalent:
  - $\{[0, \frac{1}{2}), (\frac{1}{2}, 1]\}$; $\{[0, \frac{1}{2}), \{\frac{1}{2}\}, (\frac{1}{2}, 1]\}$, $\{[0, \frac{1}{2}], [\frac{1}{2}, 1]\}$

Rich Experimental Settings

We say a set of experiments $\mathbb{E}$ is ***rich*** if

(1) $(A, \mathscr{P}) \in \mathbb{E} \to (A, \mathscr{Q}) \in \mathbb{E}$ whenever $\mathscr{Q}$ is a coarsening of $\mathscr{P}$

(2) For any finite Ω-measurable partition of $\mathcal{U}$, there exists an experiment $(A, \mathscr{P})$ such that $\{W_{A,P}\}_{P \in \mathscr{P}}$ is μ equivalent

- Any partition can be approximated up to 0 probability events
- For the EU model, the set of all experiments is rich for any "regular" μ

Primitive

- Our primitive is a ranking $\succcurlyeq$ over the set of all random experiments
- A ***random experiment*** is a lottery over some (fixed) rich set $\mathbb{E}$

Normative Principles

"Two experiments that identify the sets of parameters are equally valued"

(P1) - Structural Invariance

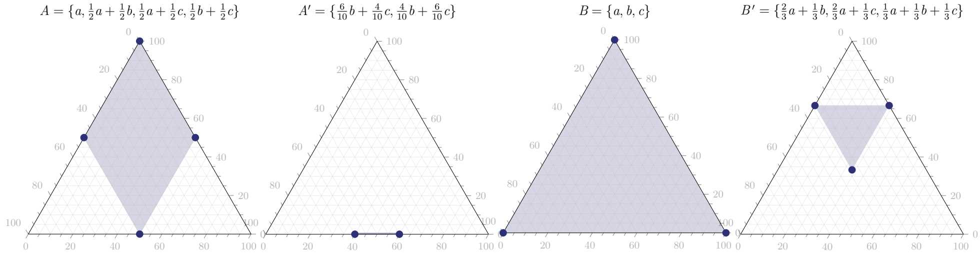
If $\{\, W_{A,P} | P \in \mathscr{P} \}$ is μ-equivalent to $\{\, W_{B,Q} | Q \in \mathscr{Q} \}$ then $(A, \mathscr{P}) \sim (B, \mathscr{Q})$.

- Structural properties of experiments are irrelevant
- Also, 0-probability events are irrelevant

Consider our EU maximizing subject choosing lotteries over $\{a,\, b,\, c\}$.

$$\begin{aligned}\text{EXP A}: A \ &= \{a,\ \ \tfrac{1}{2}a + \tfrac{1}{2}b,\ \ \tfrac{1}{2}a + \tfrac{1}{2}c,\ \ \tfrac{1}{2}b + \tfrac{1}{2}c\} \\ A^{'} &= \{\tfrac{6}{10}b + \tfrac{4}{10}c,\ \ \tfrac{4}{10}b + \tfrac{6}{10}c\}.\end{aligned}$$

$$\begin{aligned}\text{EXP B}: B \ &= \{a,\ \ b,\ \ c\} \\ B^{'} &= \{\tfrac{2}{3}a + \tfrac{1}{3}b,\ \ \tfrac{2}{3}a + \tfrac{1}{3}c,\ \ \tfrac{1}{3}a + \tfrac{1}{3}b + \tfrac{1}{3}c\}.\end{aligned}$$



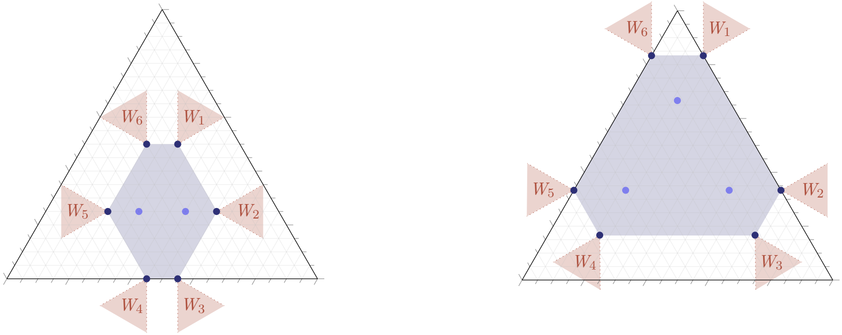
$A = \{a, \frac{1}{2}a + \frac{1}{2}b, \frac{1}{2}a + \frac{1}{2}c, \frac{1}{2}b + \frac{1}{2}c\}$
$A' = \{\frac{6}{10}b + \frac{4}{10}c, \frac{4}{10}b + \frac{6}{10}c\}$
$B = \{a, b, c\}$
$B' = \{\frac{2}{3}a + \frac{1}{3}b, \frac{2}{3}a + \frac{1}{3}c, \frac{1}{3}a + \frac{1}{3}b + \frac{1}{3}c\}$

- Linearity states that

$$\left.\begin{array}{l} x \in \arg\max_A u(\cdot) \\ y \in \arg\max_B u(\cdot) \end{array}\right\} \quad \text{if and only if} \quad \alpha x + (1-\alpha) y \in \arg \max_{\alpha A + (1-\alpha) B} u(\cdot)$$

- Therefore, observing A followed by A' is equivalent to observing $\frac{1}{2}A + \frac{1}{2}A'$

$\frac{1}{2}A + \frac{1}{2}A'$ $\qquad\qquad$ $\frac{1}{2}B + \frac{1}{2}B'$



- Structural invariance reflects the symmetries of the given domain
- With linear utility, the symmetry is ***translation invariance***:

Structural Invariance for Expected Utility

$$(A, \{P_1, \ldots P_n\}) \sim (A + B, \{P_1 + B, \ldots P_n + B\})$$

- This isn't exactly correct, since $\{P_1 + B, \ldots P_n + B\}$ might have overlaps....

"Experiments that induce sharper identification are (weakly) better"

(P2) - Information Monotonicity

If $\mathscr{P}$ refines $\mathscr{Q}$ then $(A, \mathscr{P}) \succcurlyeq (A, \mathscr{Q})$.

- Preference respects Blackwell order

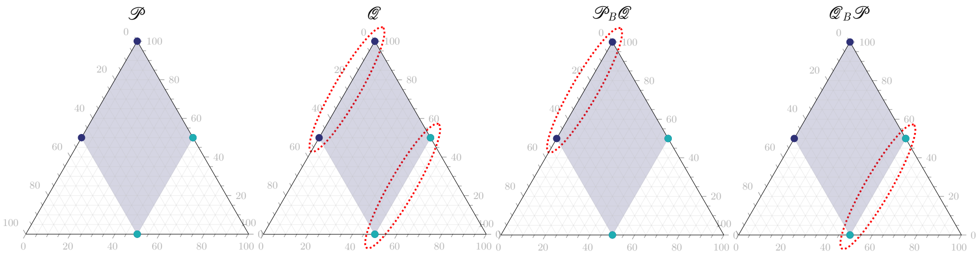
"The value of identification not depend on counterfactuals"

(P3) - Identification Separability

$$\tfrac{1}{2}(A,\mathscr{P})+\tfrac{1}{2}(A,\mathscr{Q})\sim\tfrac{1}{2}(A,\mathscr{P}_B\mathscr{Q})+\tfrac{1}{2}(A,\mathscr{Q}_B\mathscr{P}).$$

- $\mathscr{P}$ and $\mathscr{Q}$ partitions of A and $B\subseteq A$, then $\mathscr{P}_B\mathscr{Q}$ denotes the partition that coincides with $\mathscr{P}$ over B and with $\mathscr{Q}$ over $A\setminus B$

Consider decision problem A (from before) with the following partitions



- The set B is the two south-east lotteries (in teal)

Theorem

Let $\succcurlyeq$ be an expected utility preference, represented by index $F : \mathbb{E} \to \mathbb{R}$.

Then $\succcurlyeq$ satisfies **P1-3** if and only if there exists a $\tau : \Omega \to \mathbb{R}$ such that:

$$F(A, \mathscr{P}) = \sum_{P \in \mathscr{P}} \tau(W_{A,P})\mu(W_{A,P})$$

with $W \subseteq V$ implies

- $\tau(W)\mu(W|V) + \tau(V \setminus W)(1 - \mu(W|V)) \geq \tau(V)$
- $\mu(V \setminus W) = 0$ implies $\tau(W) = \tau(V)$

Representation reflects our normative principles:

$$F(A,\mathscr{P}) = \sum_{P\in\mathscr{P}} \tau(W_{A,P})\mu(W_{A,P})$$

- Only dependents on $W_{A,P} \rightarrow$ Structural Invariance
- Additive $\rightarrow$ Identification Separability
- $\tau(W)\mu(W|V) + \tau(V\setminus W)(1\text{-}\mu(W|V)) \geq \tau(V) \rightarrow$ Monotonicity

Entropy

- Entropy is a common measure of information
- Entropy of a probability measure is

$$-\sum_{x \in \mathrm{supp}(\mu)} \log(\mu(x))\mu(x)$$

- The experimenter's value for an experiment is the (expected) entropy of the induced identification

$$F(A, \mathscr{P}) = -\sum_{P \in \mathscr{P}} \log(\mu(W_{A,P}))\mu(W_{A,P})$$

We can specialize each of the normative principals to this context

Structural Invariance for Entropy: Symmetry

Fix $(A, \{P_1, \ldots P_n\})$ and $(B, \{Q_1, \ldots Q_n\})$. Then if for all $i \leq n$,

$$|\mu(W_{B,Q_i}) - \tfrac{1}{n}| \geq |\mu(W_{A,P_i}) - \tfrac{1}{n}|$$

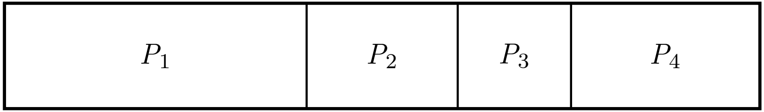
it follows that

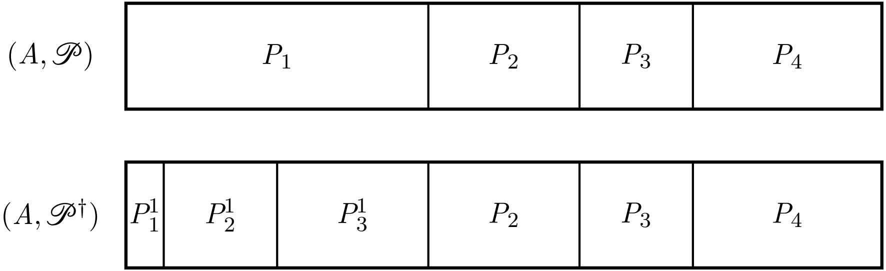
$$(A, \{P_1 \ldots P_n\}) \succcurlyeq (B, \{Q_1 \ldots Q_n\}).$$

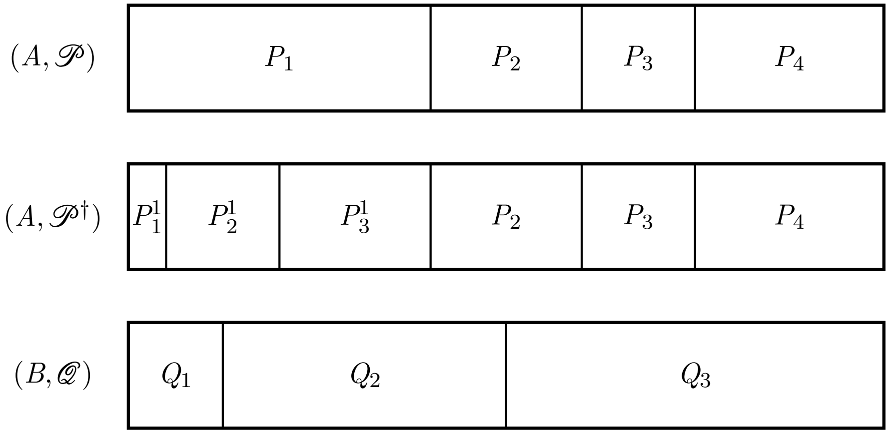
- implies structural invariance:

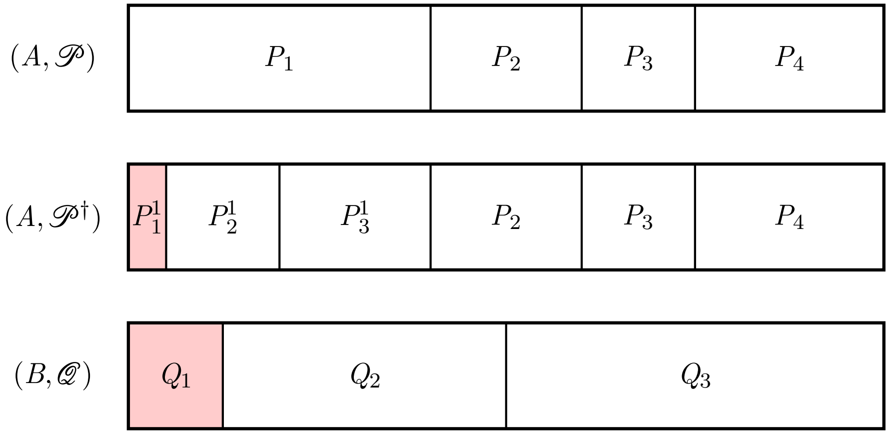
- Fix $(A, \mathscr{P} = \{P_1, \ldots P_n\})$ and let $\mathscr{P}^1 = \{P_1^1, \ldots P_k^1\}$ partition P_1.
- Then $\mathscr{P}^\dagger = \{P_1^1, \ldots P_k^1, P_2, \ldots P_n\}$ is also partition of A.
  - As if observing $\mathscr{P}$ and then if P_1 is realized, further observing $\mathscr{P}^1$
  - $\mathscr{P}$ observed with prob 1, $\mathscr{P}^1$ observed with probability $\mu(W_{A,\mathscr{P}_1})$
- Let $(B, \mathscr{Q} = \{Q_1, \ldots Q_k\})$ with $\mu(W_{B,Q_i}) = \mu(W_{A,P_i^1} \mid W_{A,P_1})$
  - Observing $\mathscr{Q}$ has same 'informational content' as observing $\mathscr{P}^1$ conditional on realization of P_1

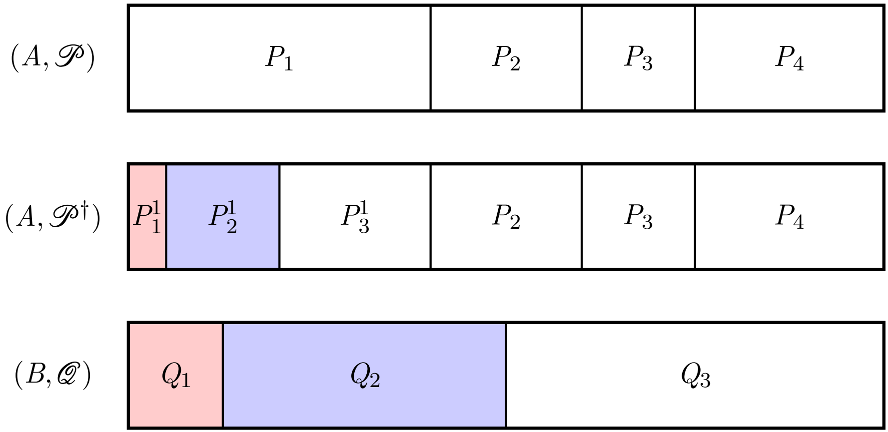
$(A, \mathscr{P})$

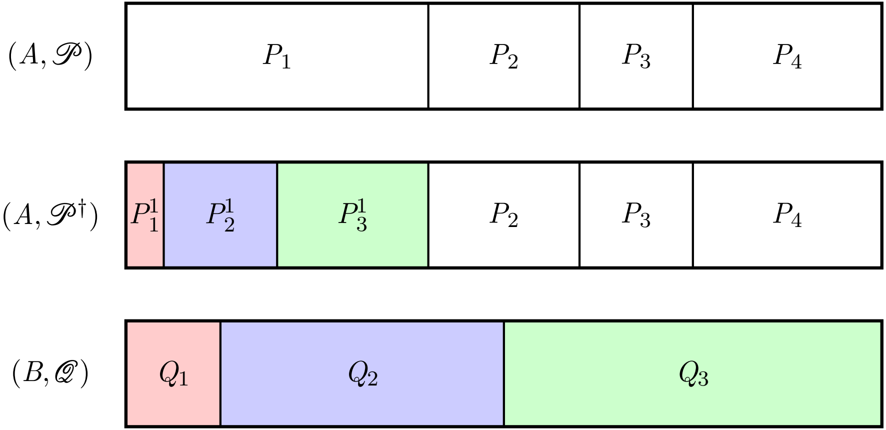



$(A, \mathscr{P})$
P_1
P_2
P_3
P_4
$(A, \mathscr{P}^\dagger)$
P_1^1
P_2^1
P_3^1
P_2
P_3
P_4


$(A, \mathscr{P})$
P_1 P_2 P_3 P_4
$(A, \mathscr{P}^\dagger)$
P_1^1 P_2^1 P_3^1 P_2 P_3 P_4
$(B, \mathscr{Q})$
Q_1 Q_2 Q_3


$(A, \mathscr{P})$
P_1
P_2
P_3
P_4
$(A, \mathscr{P}^\dagger)$
P_1^1
P_2^1
P_3^1
P_2
P_3
P_4
$(B, \mathscr{Q})$
Q_1
Q_2
Q_3


$(A, \mathscr{P})$
P_1
P_2
P_3
P_4
$(A, \mathscr{P}^\dagger)$
P_1^1
P_2^1
P_3^1
P_2
P_3
P_4
$(B, \mathscr{Q})$
Q_1
Q_2
Q_3



$(A, \mathscr{P})$
P_1
P_2
P_3
P_4
$(A, \mathscr{P}^\dagger)$
P_1^1
P_2^1
P_3^1
P_2
P_3
P_4
$(B, \mathscr{Q})$
Q_1
Q_2
Q_3

Option 1	Option 2
◇ Observe $\mathscr{P}^\dagger$ with prob α	◇ Observe $\mathscr{P}$ with prob α
◇ Observe nothing with prob $(1-\alpha)$	◇ Observe $\mathscr{Q}$ with prob $(1-\alpha)$

- To be indifferent α must capture the likelihood of reviving extra information
- This is $\alpha = \frac{1}{1+\mu(W_{A,P_1})}$

Identification Separability for Entropy

Fix $(A, \mathscr{P} = \{P_1, \dots P_n\})$ and $\mathscr{P}^\dagger = \{P_1^1, \dots P_k^1, P_2, \dots P_n\}$.

Set $\alpha = \frac{1}{1+\mu(W_{A,P_1})}$.

Then if $(B, \{Q_1, \dots Q_k\})$ is such that $\mu(W_{B,Q_i}) = \mu(W_{A,P_i^1} \mid W_{A,P_1})$, it follows that

$$\alpha(A, \mathscr{P}^\dagger) + (1-\alpha)(A, \{A\}) \sim \alpha(A, \mathscr{P}) + (1-\alpha)(B, \mathscr{Q})$$

- implies identification separability

Theorem

Let $\succcurlyeq$ be an expected utility preference.

Then $\succcurlyeq$ satisfies the entropic versions of `P1-3` if and only if

$$F(A, \mathscr{P}) = -\sum_{P \in \mathscr{P}} \log(\mu(W_{A,P}))\mu(W_{A,P})$$

is an utility index for $\succcurlyeq$

Thank You!