Modeling the Modeler:
A Normative Theory of Experimental Design

Evan Piermont
Royal Holloway, University of London

&

Fernando Payró
Universitat Autònoma de Barcelona

University of Surrey

The Twentieth Day of November in the Year of Our Lord Twenty Twenty Five

An experimenter seeks to learn about a subject:

- has a theory about how subjects makes choices, conditional on their type
  - type $=_{def}$ parameters of preference / beliefs / whatever drives choices
- provides experiments to the subject, observes outcome
  - experiment $=_{def}$ a choice problem, from which the subject's behavior is observed

Our paper: how should the experimenter's goals (and her theory of behavior) determine her valuation of experiments

There are different goals for different experimenters:

Information Maximizing: wishes to learn as much as possible about the subject, independently of what she learns

There are different goals for different experimenters:

Information Maximizing: wishes to learn as much as possible about the subject, independently of what she learns

Theory Testing: Gains more value from validating a hypothesis than from anything else

There are different goals for different experimenters:

Information Maximizing: wishes to learn as much as possible about the subject, independently of what she learns

Theory Testing: Gains more value from validating a hypothesis than from anything else

Profit Seeking: Will use her information about the subject to sell him something

There are different goals for different experimenters:

Information Maximizing: wishes to learn as much as possible about the subject, independently of what she learns

Theory Testing: Gains more value from validating a hypothesis than from anything else

Profit Seeking: Will use her information about the subject to sell him something

Deceitful: Uninterested in truth, only wishes to validate hypothesis

- We offer a framework unifying the many goals of experimentation
- We propose three normative principles for how to rank experiments
  - minimal rationality properties, independent of specific motivations
- We show that they imply a particular representation
  - Relates a experiment to the expected value of identification
  - Provides a recipe for choosing experiments
  - Distinguish between various ‘types’ of experimenters
  - Test to ensure experimenter does not have an ‘agenda’

- A space of choice parameters Θ
- A (costly) experiment A has
  - possible observable outcomes $\mathscr{P} = \{P_1, \ldots, P_n\}$
  - cost on implementation $c \geq 0$
- Observing $P \in \mathscr{P}$ *partially identifies* a set of parameters:
  - $W_{A,P} \subseteq \Theta$ consistent with observation

Normative Principles

Structural Invariance: Prefer the least costly implementation of a given identification of parameters

Information Monotonicity: (Weakly) prefer experiments that induce sharper identification

Identification Separability: Consider only the elements of experiments that can be controlled

Expected Identification Value

These principles characterize *expected identification value* maximization

- Exists some τ: for $W \subseteq \Theta$, $\tau(W)$ is the value of identifying W
- Experiment $(A, \mathscr{P}, c)$ is valued according to:

$$\sum_{P \in \mathscr{P}} \tau(W_{A,P})\mu(W_{A,P}) - c$$

- where μ is the (exogenous) prior probability

Special Case: Entropy

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

- Value of experiment is expected reduction in entropy
- Exp is valued proportional to information gain
- ‘Information Maximizing’ experimenter

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 parameter lies in W^*
- Value of exp is the probability the hypothesis can be accepted or rejected
- ‘Theory Testing’ experimenter

Special Case: Actions

$$\tau(W) = \max_{\alpha \in \mathbb{A}} \int_{\theta \in W} \xi(\alpha, \theta)\, d\mu.$$

- The analyst will take action $\alpha \in \mathbb{A}$
- Utility of outcome depends on the parameter: $\xi(a, \theta)$
- Value of exp is expected value of conditionally optimal action
- ‘Profit Seeking’ experimenter

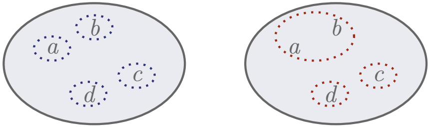
Experimental Paradigms

An **Experimental Paradigm** is (Z, E, T)

- Z is a set of **choice patterns**
- E is a set of **experiments**.
- T is the experimenters **theory of behavior**.

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

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



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

A **theory of behavior** $T = \langle \Theta, \Omega, \mu \rangle$:

- Θ is a set of types (choice parameters), each $\theta \in \Theta$ associated with a choice function

$$c_\theta : 2^Z \to Z \qquad \text{such that} \qquad c_\theta(A) \in A.$$

- Ω algebra of measurable sets of Θ
- μ prior over (Θ, Ω)

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

$$W_{A,P} = \{\theta \in \Theta \mid c_\theta(A) \in P\}$$

Observing $P \in \mathscr{P}$ identifies that the subject's type is in $W_{A,P}$

Concordant Exp. Paradigms

We call an experimental paradigm (Z, E, T) **concordant** if the following hold:

(1) Any observable outcome can be measured

(2) Any measurable hypothesis can be tested

Concordant Exp. Paradigms

We call an experimental paradigm (Z, E, T) **concordant** if the following hold:

(1) Any observable outcome can be measured

- For all $(A, \mathscr{P}) \in E$ and $P \in \mathscr{P}$, $W_{A,P} \in \Omega$

(2) Any measurable hypothesis can be tested

Concordant Exp. Paradigms

We call an experimental paradigm (Z, E, T) **concordant** if the following hold:

(1) Any observable outcome can be measured

- For all $(A, \mathscr{P}) \in E$ and $P \in \mathscr{P}$, $W_{A,P} \in \Omega$

(2) Any measurable hypothesis can be tested

- For all Ω-meas. partitions $\{W_1, \ldots, W_n\}$ of Θ, there exists $(A, \mathscr{P}) \in E$:

$$\{W_{A,P} | P \in \mathscr{P}\} =_\mu \{W_1, \ldots, W_n\}$$

($=_\mu$ means up to μ-measure 0 differences)

Example: Choice Functions

Let X be a finite set of alternatives, subject chooses element out of $A \subseteq X$

- $Z = X \times \ldots \times X$
- $E = \{(A_1 \times \ldots \times A_n, \mathscr{P}) \mid A_i \subseteq X, \mathscr{P}$ partitions $A_1 \times \ldots \times A_n\}$
- Θ is all (strict) preference orders over X:

$$c_\theta(A_1 \times \ldots \times A_n) = \prod_{i \leq n} \{x \in A_i \mid x \succ_\theta y, \text{ for all } y \in A_i\}$$

- Then for discrete Ω, any μ forms a concordant paradigm

Example: Choice Functions

Let X be a finite set of alternatives, subject chooses element out of $A \subseteq X$

- $Z = X \times \ldots \times X$
- $E = \{(A_1 \times \ldots \times A_n, \mathscr{P}) \mid A_i \subseteq X, \mathscr{P}$ partitions $A_1 \times \ldots \times A_n\}$
- Θ is all (strict) preference orders over X:

$$c_\theta(A_1 \times \ldots \times A_n) = \prod_{i \leq n} \{x \in A_i \mid x \succ_\theta y, \text{ for all } y \in A_i\}$$

- Then for discrete Ω, any μ forms a concordant paradigm
  - this is **not** true for weak preferences; no way to test for tie-breaking

Example: Choice Correspondences

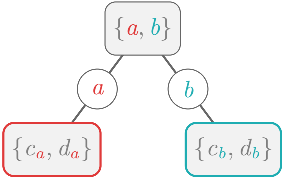
Let X be a finite set of alternatives, subject chooses subset out of $A \subseteq X$

- $Z = 2^X \times \ldots \times 2^X$
- $E = \big\{(2^{Y_1} \times \ldots \times 2^{Y_n}, \mathscr{P}) \mid Y_i \subseteq X, \mathscr{P}$ partitions $2^{Y_1} \times \ldots \times 2^{Y_n}\big\}$
- Θ is all preference orders over Z:

$$c_\theta(2^{Y_1} \times \ldots \times 2^{Y_n}) = \prod_{i \leq n} \{x \in Y_i \mid x \succcurlyeq_\theta y, \text{ for all } y \in Y_i\}$$

- Then for discrete Ω, any μ forms a concordant paradigm

Example: Adaptive Experiments



Let $X = \{a, b, c_a, c_b, d_a, d_b\}$

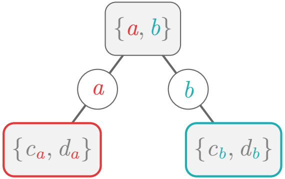
- $Z = \{(a, c_a), (a, d_a), (b, c_b), (b, d_b)\}$
- $E = \{(Z, \mathscr{P}) \mid \mathscr{P}$ partitions $Z\}$
- Θ collect utility functions over X: could represent naive subjects

$$c_\theta(Z) = (x, y) \text{ such that } x \in \arg\max_{z \in \{a,b\}} \theta(z) \text{ and } y \in \arg\max_{w \in \{c_x, d_x\}} \theta(w)$$

or sophisticated subjects

$$c_\theta(Z) = (x, y) \text{ such that } x \in \arg\max_{(z,w) \in Z} \theta(z) + \theta(w)$$

Example: Adaptive Experiments



Let $X = \{a, b, c_a, c_b, d_a, d_b\}$

- $Z = \{(a, c_a), (a, d_a), (b, c_b), (b, d_b)\}$
- $E = \{(Z, \mathscr{P}) \mid \mathscr{P} \text{ partitions } Z\}$
- In either case: Ω must reflect the inability to simultaneously measure preferences over $\{c_a, d_a\}$ and also $\{c_b, d_b\}$
- Ω must be generated by the four choice patterns in Z

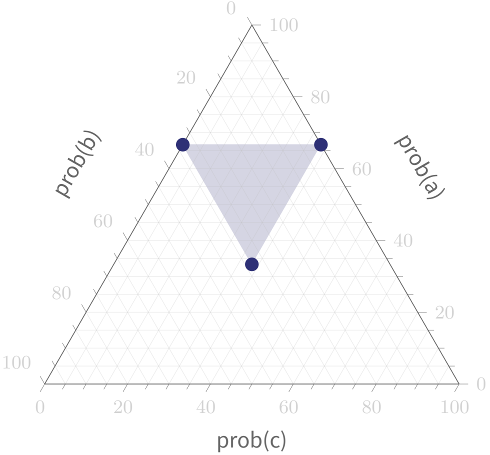
Example: Expected Utility

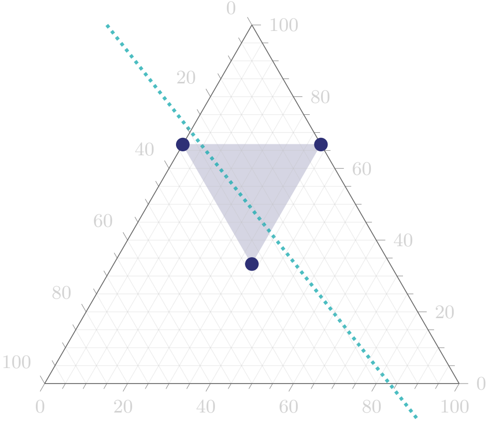
- Z is lotteries over $\{a, b, c\}$
- E is all finite subsets of Z (and all partitions thereof)
- Θ is all affine functions over Z:

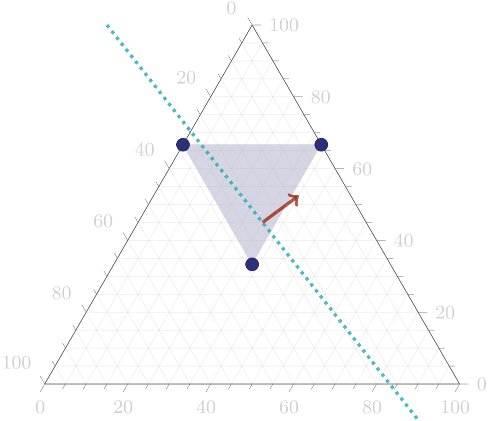
$$\theta(A) \in \arg\max_{x \in A} \theta(x)$$

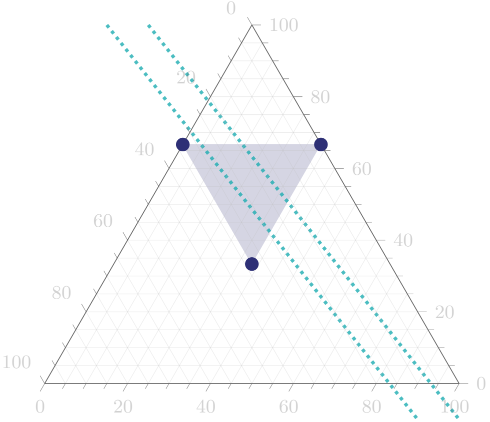
(so each parameter also specifies how ties are broken)

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

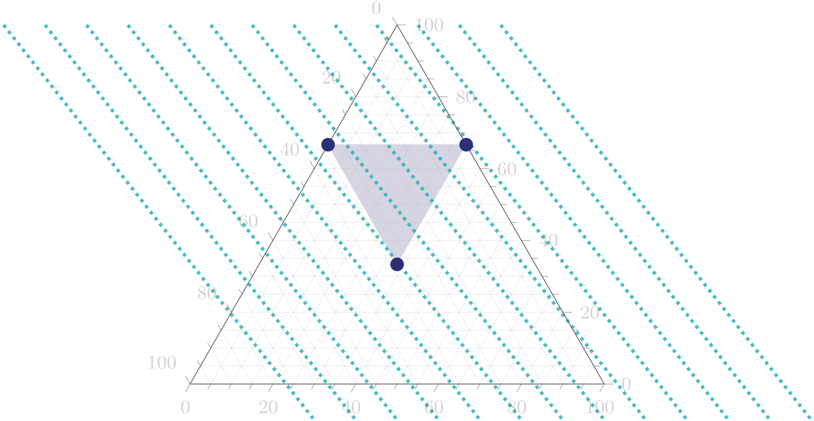



0
100
20
80
40
60
60
40
80
20
100
0
0
20
40
60
80
100

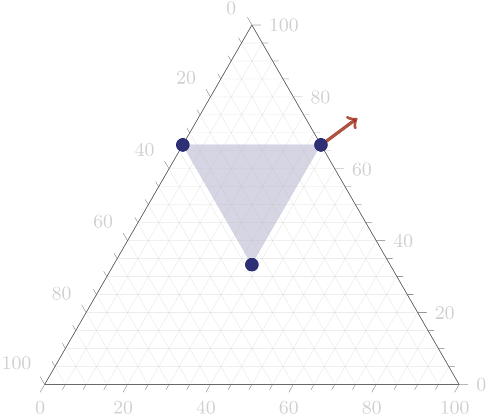

0
100
20
80
40
60
60
40
80
20
100
0
0
20
40
60
80
100

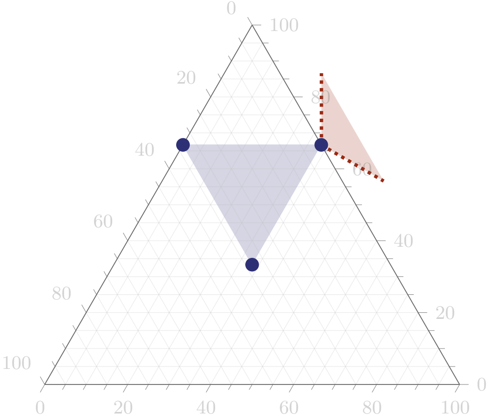


0
100
20
80
40
60
60
40
80
20
100
0
0
20
40
60
80
100

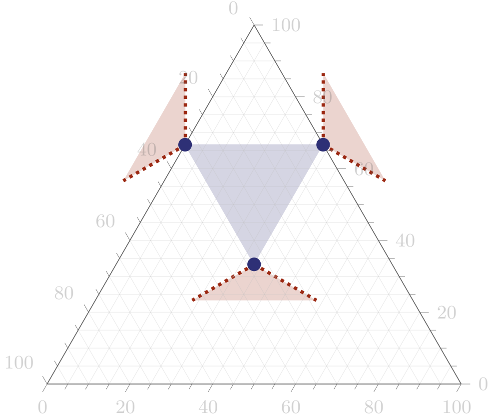


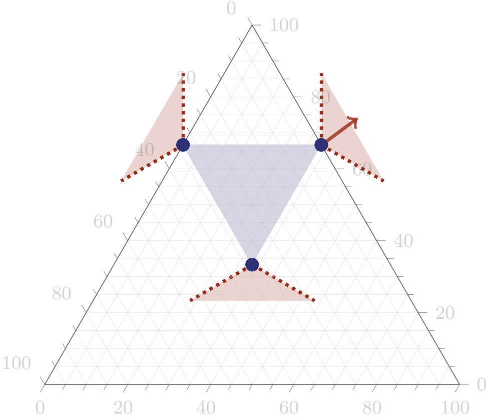
0
20
40
60
80
100
0
20
40
60
80
100
0
20
40
60
80
100


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

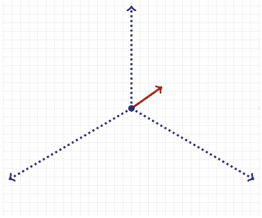


0
20
40
60
80
100
100
80
60
40
20
0
0
20
40
60
80
100


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



$\Theta \cong \mathbb{R}^2$



Example: Expected Utility

- Z is lotteries over $\{a, b, c\}$
- E is all finite subsets of Z (and all partitions thereof)
- Θ is all affine functions over Z:

$$\theta(A) \in \arg\max_{x \in A} \theta(x)$$

Theorem: Random Expected Utility

$(Z, (\Theta, \Omega, \mu), E)$ is concordant iff μ is regular à la Gul & Pesendorfer (2006):

$$\text{for all } p, q \in Z, \qquad \mu(\{\theta \mid \theta(p) = \theta(q)\}) = 0$$

Normative Principles

Primitive

- Fix some concordant paradigm: (Z, E, T):
- Our primitive is a ranking $\succcurlyeq$ over the set of all *costly* experiments:

$$(A, \mathscr{P}, c) \qquad \text{where } (A, \mathscr{P}) \in E \quad \text{and} \quad c \geq 0$$

- We entertain 4 axioms:
  - 3 for our normative principles
  - 1 for quasi-linearity

“ Prefer the least costly implementation of a given identification of parameters ”

(P1) - Structural Invariance

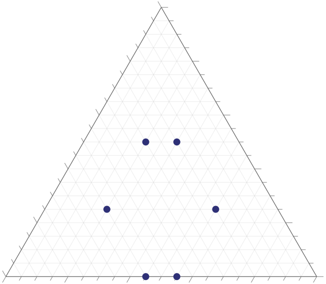
Let $\{W_{A,P} | P \in \mathscr{P}\} =_\mu \{W_{B,Q} | Q \in \mathscr{Q}\}$:

$$(A, \mathscr{P}, c) \succcurlyeq (B, \mathscr{Q}, c') \qquad \text{if and only if} \qquad c' \geq c.$$

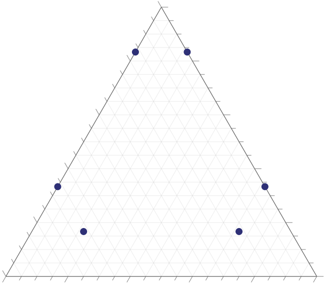
- Structural properties of experiments are irrelevant
- Also, 0-probability events are irrelevant ($=_\mu$)

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

A

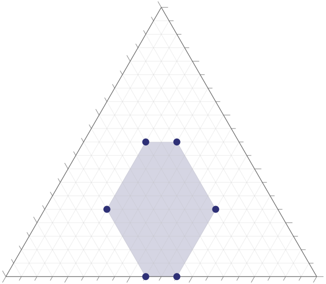


B

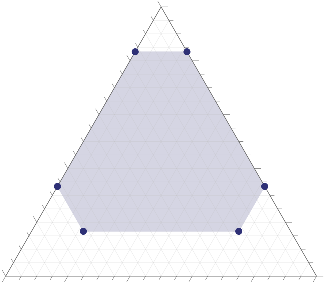


Consider our EU maximizing subject choosing lotteries over $\{a, b, c\}$.

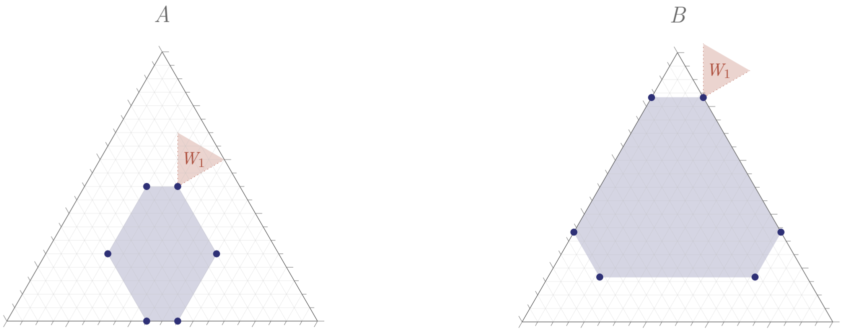
A



B

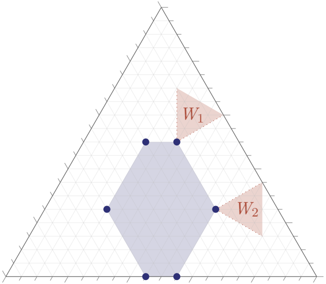


Consider our EU maximizing subject choosing lotteries over $\{a, b, c\}$.

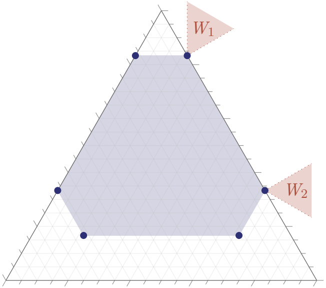


Consider our EU maximizing subject choosing lotteries over $\{a, b, c\}$.

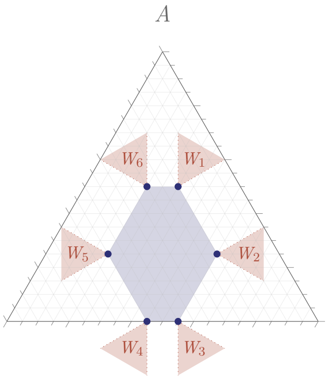
A

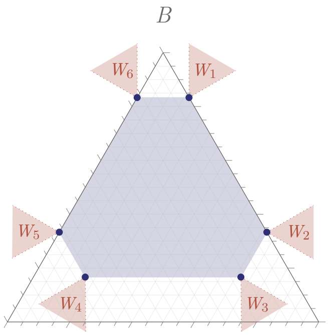


B



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





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

Structural Invariance for Expected Utility

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

“ (Weakly) prefer experiments that induce sharper identification ”

(P2) - Information Monotonicity

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

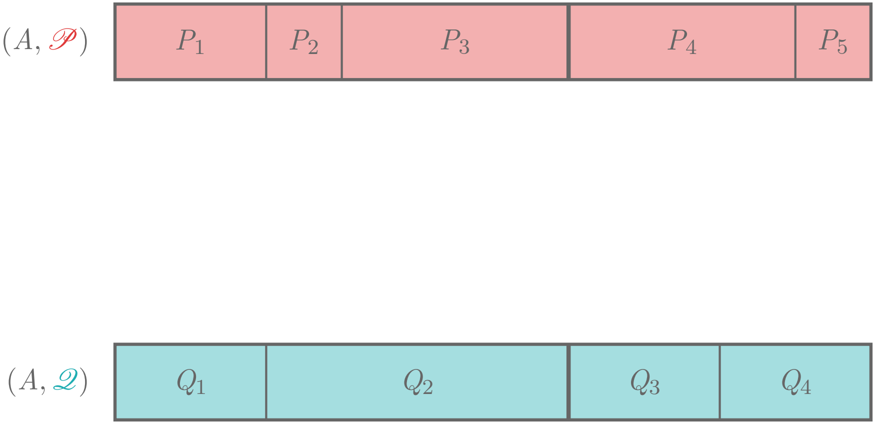
- Preference respects Blackwell order

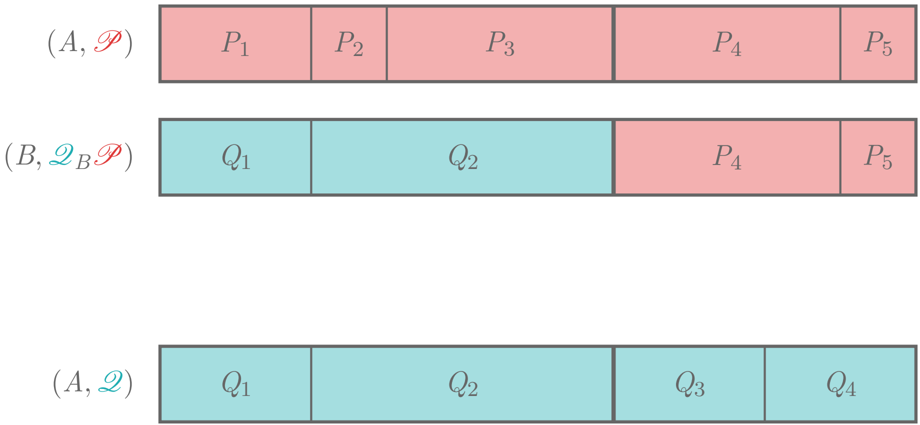
“ Consider only the elements of experiments that can be controlled ”

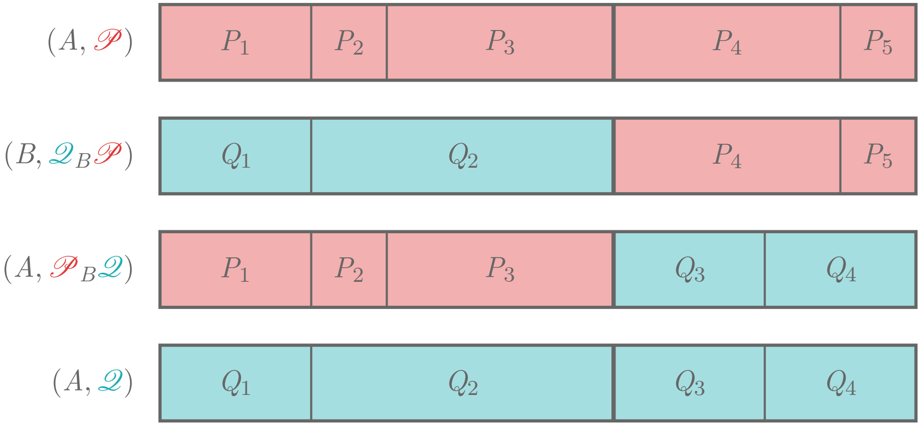
(P3) - Identification Separability

$$(A, \mathscr{P}, c) \sim (A, \mathscr{Q}_B \mathscr{P}, 0) \quad \text{if and only if} \quad (A, \mathscr{P}_B \mathscr{Q}, c) \sim (A, \mathscr{Q}, 0).$$

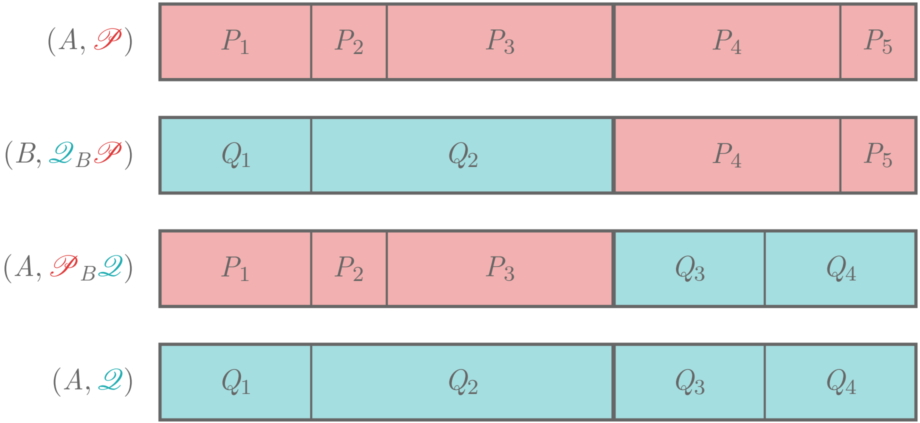
- If $\mathscr{P}$ and $\mathscr{Q}$ partition of A and $B \subseteq A$, then $\mathscr{P}_B \mathscr{Q}$ denotes the partition that coincides with $\mathscr{P}$ over B and with $\mathscr{Q}$ over $A \setminus B$


$(A, \mathscr{P})$
P_1
P_2
P_3
P_4
P_5
$(A, \mathscr{Q})$
Q_1
Q_2
Q_3
Q_4


$(A, \mathscr{P})$
P_1 P_2 P_3 P_4 P_5
$(B, \mathscr{Q}_B\mathscr{P})$
Q_1 Q_2 P_4 P_5
$(A, \mathscr{Q})$
Q_1 Q_2 Q_3 Q_4


$(A, \mathscr{P})$
P_1 P_2 P_3 P_4 P_5
$(B, \mathscr{Q}_B\mathscr{P})$
Q_1 Q_2 P_4 P_5
$(A, \mathscr{P}_B\mathscr{Q})$
P_1 P_2 P_3 Q_3 Q_4
$(A, \mathscr{Q})$
Q_1 Q_2 Q_3 Q_4

$$(A, \mathscr{P}, c) \sim (A, \mathscr{Q}\,{}_B\,\mathscr{P}, 0) \quad \text{if and only if} \quad (A, \mathscr{P}\,{}_B\,\mathscr{Q}, c) \sim (A, \mathscr{Q}, 0).$$



Let $\succcurlyeq$ be an quasi-linear preference, represented by index $U : E \times \mathbb{R}_+ \to \mathbb{R}$.

Theorem: Expected Identification Value Maximization

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

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

with $W \subseteq V$ implies

- $\tau(V) \leq \tau(W)\mu(W|V) + \tau(V \setminus W)(1\text{-}\mu(W|V))$
- if also $\mu(V \setminus W) = 0$ then $\tau(W) = \tau(V)$

Representation reflects our normative principles:

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

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

Information Based Models

Information Based Models

- Often, we seek only to reduce uncertainty
  - Experimenter does not have a preference for *what* is learned
- Value of identification $\tau(W)$, depends only of $\mu(W)$.
  - We can specialize structural invariance to capture this.
- E.g., value for an experiment is the (expected) reduction in entropy

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

Probability Vectors

For each $e = (A, \{P_1, \ldots, P_n\}) \in E$ let

$$\mathbf{p}_e = \{\mu(W_{A,P_1}), \ldots, \mu(W_{A,P_n})\} \in \Delta^n$$

denote the probability vector induced by the experiment.

(P2*) - Information Based Structural Invariance

Let $\mathbf{p}_e = \mathbf{p}_{e'}$ and $c' \geq c$: then

$$(e, c) \succcurlyeq (e', c'),$$

with a strict preference whenever $c' > c$.

(P2) - Schur Concavity**

Let $\mathbf{p}_e = \mathbf{D}\,\mathbf{p}_{e'}$ for some doubly stochastic matrix $\mathbf{D}$, and $c' \geq c$: then

$$(e, c) \succcurlyeq (e', c'),$$

with a strict preference whenever $c' > c$.

Let $\succcurlyeq$ be an Expected Identification Value Maximization preference:

Theorem: Belief Based Models

Let μ be non-atomic. $\succcurlyeq$ satisfies P2* if and only

$$U(A, \mathscr{P}, c) = \sum_{P \in \mathscr{P}} h(\mu(W_{A,P}))\mu(W_{A,P}) - c.$$

for some $h : [0, 1] \to \mathbb{R}$ with $h(0) = 0$ is such that

- For all $p, \alpha \in [0, 1]$,

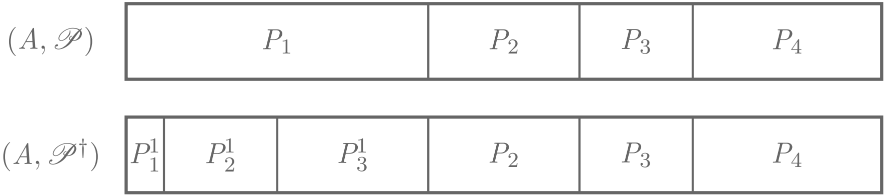
$$h(p) \leq \alpha h(\alpha p) + (1 - \alpha)h((1 - \alpha)p)$$

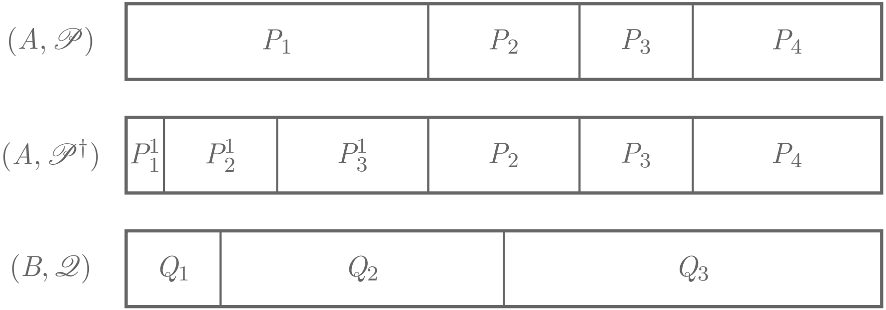
- $p \mapsto h(p)p$ is concave iff P2* is strengthened to P2**.

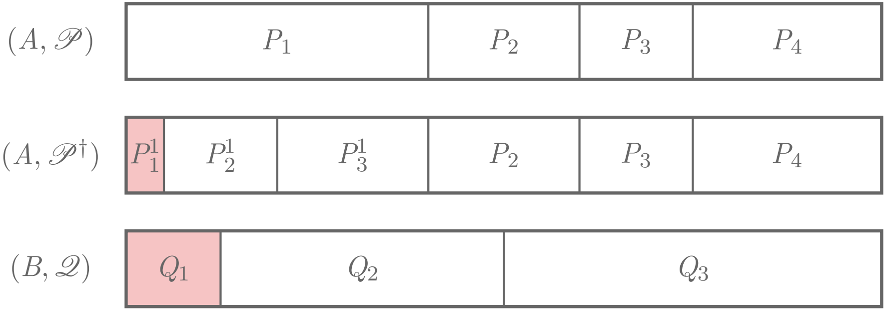
Entropic Partitions

- Fix $(A, \mathscr{P} = \{P_1, \ldots P_n\})$ and let $\mathscr{P}^1 = \{P^1_1, \ldots P^1_k\}$ partition P_1.
- Then $\mathscr{P}^\dagger = \{P^1_1, \ldots P^1_k, P_2, \ldots P_n\}$ 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 $\mu(W_{B,Q_i}) = \mu(W_{A,P^1_i} \mid W_{A,P_1})$
  - Observing $\mathscr{Q}$ has same 'informational content' as observing $\mathscr{P}^1$ conditional on realization of P_1

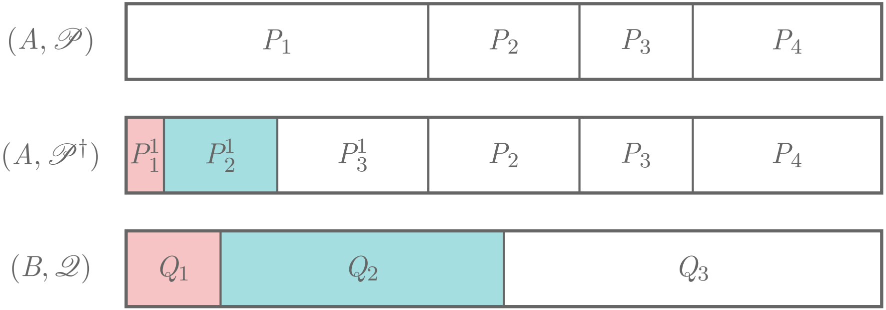

$(A, \mathscr{P})$
P_1
P_2
P_3
P_4


$(A, \mathscr{P})$
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

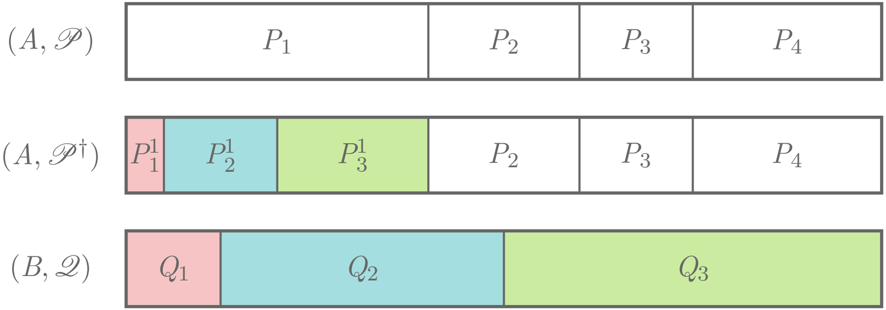

$(A, \mathscr{P})$
P_1
P_2
P_3
P_4
$(A, \mathscr{P}^\dagger)$
P_1^1
P_2^1
P_3^1
P_2
P_3
P_4
$(B, \mathscr{Q})$
Q_1
Q_2
Q_3



$(A, \mathscr{P})$
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

(P3*) - Identification Separability for Entropy

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

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

$$(B, \mathscr{Q}, c) \sim (A, \{A\}, 0) \quad \text{if and only if} \quad (A, \mathscr{P}^\dagger, \mu(W_{A,P_1}) \cdot c) \sim (A, P, 0)$$

- Conditional on P_1 being realized, observing $\mathscr{P}^1$ has the same information content as (unconditionally) observing $\mathscr{Q}$
- P_1 is realized with probability $\mu(W_{A,P_1})$.

Let $\succcurlyeq$ be an Expected Identification Value Maximization preference:

Theorem: Entropy Minimization

Let μ be non-atomic. $\succcurlyeq$ satisfies P2** and P3* if and only

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

Thank You!