Making Information More Valuable

Mark Whitmeyer

Doctor's visit

A patient (with a hurt hand) visits a doctor.

Three possibilities–three states of the world:

1. State 0: sprain.
2. State 1: broken bone, but not displaced.
3. State 2: displaced fracture.

Doctor's visit

Insurance dictates what doctor can do.

Suppose just one treatment available, so two options:

1. Cast on the hand: optimal in states 1 and 2.
2. Do nothing: optimal in state 0.

Representation on the 2-simplex

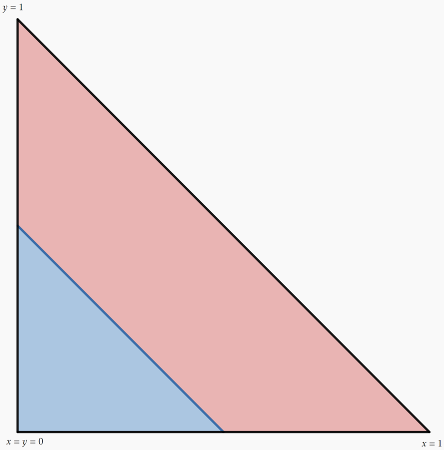


Figure: Whether to cast.

Giving her another option

Insurance makes another treatment available–*an additional action*.

One possible new treatment: conservative treatment. Not optimal in any state but not that bad in any state.

Another possible new treatment: surgery. Optimal if and only if fracture is displaced (state 2).

New representation on the simplex I

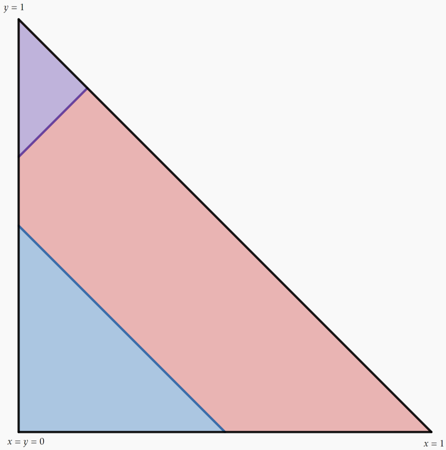


Figure: Cast, surgery, or nothing.

New representation on the simplex II

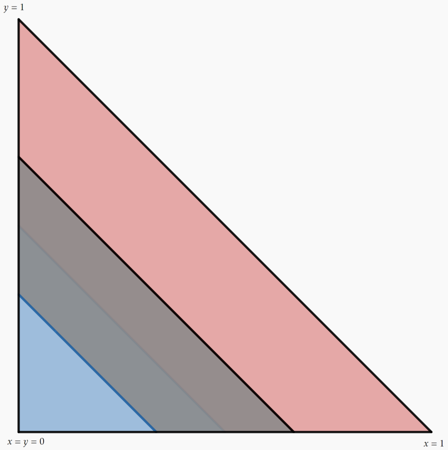


Figure: Cast, conservative, or nothing.

The value of information: question

Which of these new treatments (if any) guarantees that the value of information for the doctor has increased (or hasn't decreased)?

The value of information: answer

Surgery, not the conservative treatment.

Why? The new action is **refining**.

Only one action's region of optimality (cast) is changed–if nothing is optimal at some belief beforehand, it is still optimal after.

Information is valuable

to a Bayesian decision-maker (DM). What about *comparative* values for information?

Research question: What modifications of a DM's decision problem make information more valuable?

Two cases:

1. Buying some info.
2. Acquiring some info.

Vague answer

More convexity...

Hint: correct but **not in the sense of Pratt (1964)**.

Why relevant?

Regulators (principals in general): enact policies that modify incentives of firms/agents.

Often, **add** or **subtract** actions.

Contracts insurers may not offer.

Limits to amount vessels can fish.

Or, **scale** payoffs.

Insurance reduces risk–scale down.

Bonuses scale state-contingent payoffs up.

The Formal Setting

Setup I

Compact grand set of actions $\mathcal{A}$.

Initially, DM has access to a compact subset of actions $A \subseteq \mathcal{A}$.

Unknown state of the world, $\theta \in \Theta$ where $|\Theta| = n \in \mathbb{N}$.

Continuous utility function $u : \Theta \times \mathcal{A} \to \mathbb{R}$ & no action in A is weakly dominated.

Modifying the decision problem

Now let this DM have access to $\hat{A} \subseteq \mathcal{A}$ and utility $\hat{u} : \Theta \times \mathcal{A} \to \mathbb{R}$ instead.

Leading scenario: A finite & keep $u = \hat{u}$ but enlarge agent's action set by one, from A to $\hat{A} := A \cup \{\hat{a}\}$, where $\hat{a} \in \mathcal{A} \setminus A$. Agent **becomes more flexible** when A enlarged to $\hat{A}$.

Could also add multiple actions from $\mathcal{A} \setminus A$ (preserving $u = \hat{u}$). Agent **becomes much more flexible**.

Or, Agent **becomes much less flexible:** from A to $\emptyset \neq \hat{A} \subset A$.

Another scenario: $A = \hat{A}$ but $\hat{u} = \phi \circ u$ for some monotone ϕ (Agent's **utility is transformed**).

One more preliminary

When the set of actions is A, define the value function

$$V(x) := \max_{a \in A} \mathbb{E}_x u(a, \theta) \ ,$$

where $x \in \Delta(\Theta)$ is the agent's belief.

V is continuous and convex.

Finitely many actions $\Rightarrow$ V is piecewise affine.

$\hat{V}$ is the value function after the transformation (set of actions is $\hat{A}$).

Obtaining some information

Equivalence between Bayesian learning from a signal $\pi : \Theta \rightarrow \Delta(S)$ (for compact set of signal realizations S) and (Bayes-plausible) distribution over posteriors $\Phi \in \mathcal{F}(\mu) \subset \Delta\Delta(\Theta)$.

Say that the transformation **generates a greater value for information** if

$$\mathbb{E}_{\Phi} \hat{V}(x) - \hat{V}(\mu) \geq \mathbb{E}_{\Phi} V(x) - V(\mu),$$

for all $\Phi \in \mathcal{F}(\mu)$ and $\mu \in \operatorname{int} \Delta(\Theta)$.

Research question 1

What are necessary and sufficient conditions on the decision problems for the transformation to generate a greater value for information?

Acquiring some information

Agent's flexible information acquisition problem

$$\max_{\Phi \in \mathcal{F}(\mu)} \int_{\Delta(\Theta)} V(x)\, d\Phi(x) - D(\Phi),$$

where D is a uniformly posterior-separable cost functional, i.e.,

$$D(\Phi) = \int_{\Delta(\Theta)} c(x)\, d\Phi(x) - c(\mu)$$

for some strictly convex, continuous, function $c : \Delta(\Theta) \to \mathbb{R}$.

Acquiring some information

Say that the transformation **does not generate less information acquisition** if for any prior $\mu \in \operatorname{int}\Delta(\Theta)$, UPS cost functional D, and optimal solution to the agent's information acquisition problem when her value function is V, Φ_V^*, there exists an optimal solution to the agent's information acquisition problem when her value function is $\hat{V}$, $\Phi_{\hat{V}}^*$, that is not a strict mean-preserving contraction of Φ_V^*.

DM doesn't want to acquire less information.

Research question 2

What are necessary and sufficient conditions on the decision problems for the transformation to not generate less information acquisition?

Preliminaries

Value functions and polyhedral subdivisions

Recall value function representation

$$V(x) := \max_{a \in A} \mathbb{E}_x u(a, \theta).$$

(Finite A): Project V onto $\Delta(\Theta)$ yielding a finite collection C of polytopes C_i:

$$C_i := \{x \in \Delta \mid \mathbb{E}_x u(a_i, \theta) = V(x)\}.$$

Action a_i is optimal for any belief $x \in C_i$ and uniquely optimal for any $x \in \operatorname{int} C_i$.

C is a **regular polyhedral subdivision** of $\Delta(\Theta)$. To save space, just **subdivision**.

Subdivision illustration

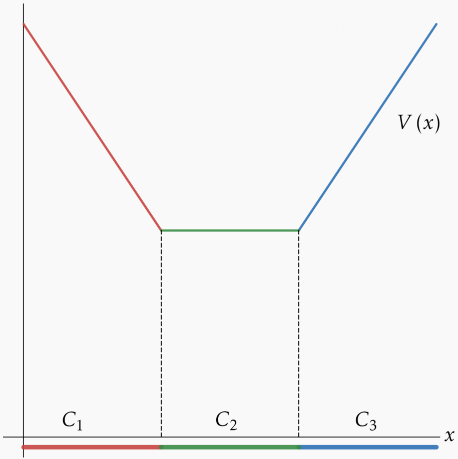


Figure: Two States, Three Actions

Subdivision illustration II

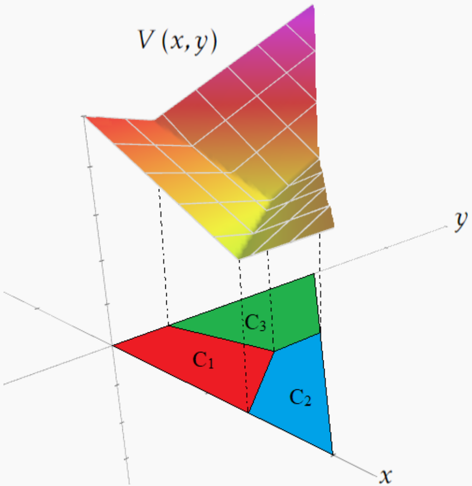


Figure: Three States, Three Actions

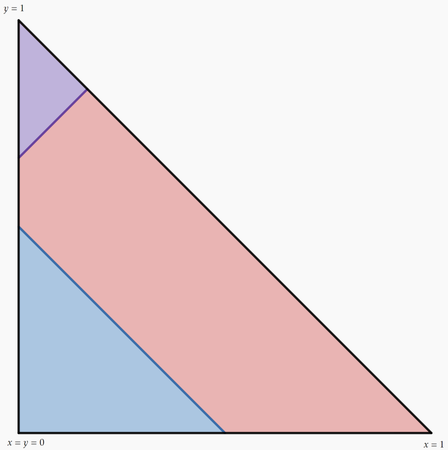
Adding a new action

Adding $\hat{a}$ leads to a new $\hat{V}$ and a new $\hat{C}$.

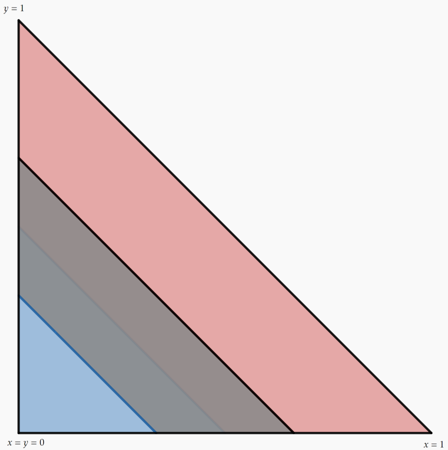
Subdivision $P = \{P_1, \ldots, P_l\}$ is finer than (or refines) a subdivision $Q = \{Q_1, \ldots, Q_m\}$ if for each $j \in \{1, \ldots, l\}$, there exists $i \in \{1, \ldots, m\}$ such that $P_j \subseteq Q_i$.

Write this $P \succeq Q$ (and $\succ$ when the relation is strict).

Finer



Not finer



Potential incomparibility of subdivisions

Remark. Any of the following can occur:

1. $C \succ \hat{C}$;
2. $C \prec \hat{C}$;
3. $C = \hat{C}$;
4. C and $\hat{C}$ incomparable.

Subdivision comparisons I

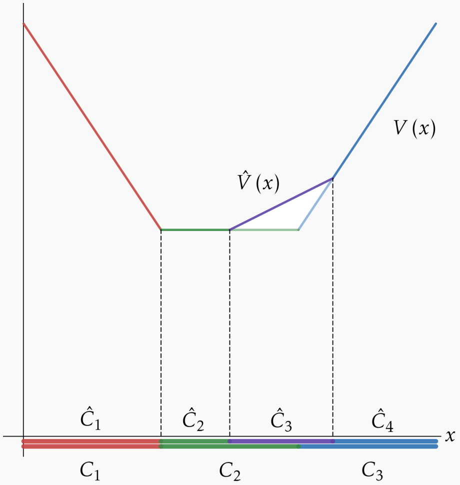


Figure: Incomparable C and $\hat{C}$

Subdivision comparisons II

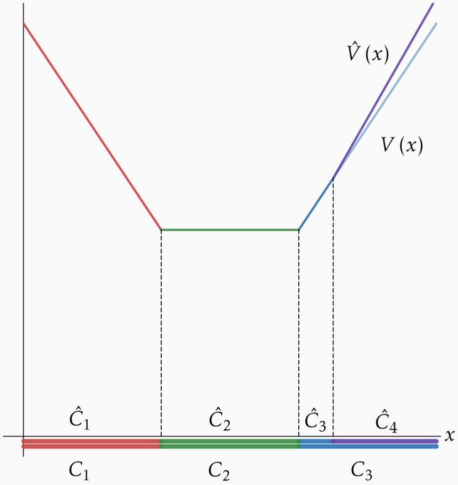


Figure: $\hat{C} \succ C$

Subdivision comparisons III

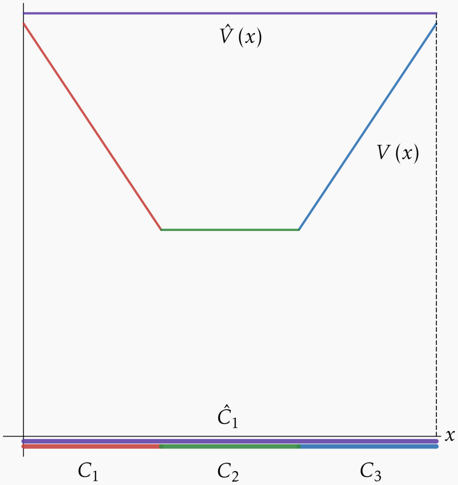


Figure: $C \succ \hat{C}$

Results

Answering the research questions

New action $\hat{a}$ is **refining** if $\hat{C} \geq C$.

$\hat{a}$ can be a (partial) replacement for most one action in A.

Main result

Theorem. The following are equivalent:

1. The transformation generates a greater value for information.
2. The transformation does not generate less information acquisition.
3. $\hat{V} - V$ is convex.
4. ★ $\hat{C} \geq C$ ★

Proof step 1

Lemma. $\hat{V} - V$ is convex $\Leftrightarrow \hat{C} \geq C$.

Proof. Easy.

Proof step 2

Lemma. $\hat{V} - V$ is convex $\Rightarrow$ For any $\mu \in \operatorname{int}\Delta(\Theta)$ and $\Phi, \Upsilon \in \mathcal{F}_\mu$ with $\Phi \in MPS(\Upsilon)$,

$$\mathbb{E}_\Phi \hat{V}(x) - \mathbb{E}_\Upsilon \hat{V}(x) \geq \mathbb{E}_\Phi V(x) - \mathbb{E}_\Upsilon V(x).$$

Proof. Rearrange definition of MPS:

$$\mathbb{E}_\Phi\left[\hat{V}(x) - V(x)\right] \geq \mathbb{E}_\Upsilon\left[\hat{V}(x) - V(x)\right].$$

Proof step 3

Lemma. $\hat{V} - V$ is convex ⇒ The transformation does not generate less information acquisition.

Proof. Similar to previous lemma. Fix an optimizer Φ^* in $\mathcal{D}$ and suppose FSOC every optimizer in $\hat{\mathcal{D}}$, $\hat{\Phi}^*$, is a strict MPC of Φ^*.

$$\mathbb{E}_{\hat{\Phi}^*} \hat{V} - D\left(\hat{\Phi}^*\right) > \mathbb{E}_{\Phi^*} \hat{V} - D\left(\Phi^*\right).$$

Analogously,

$$\mathbb{E}_{\Phi^*} V - D\left(\Phi^*\right) \geq \mathbb{E}_{\hat{\Phi}^*} V - D\left(\hat{\Phi}^*\right).$$

Combining these produces

$$\mathbb{E}_{\hat{\Phi}^*}\left[\hat{V} - V\right] > \mathbb{E}_{\Phi^*}\left[\hat{V} - V\right].$$

Something stronger

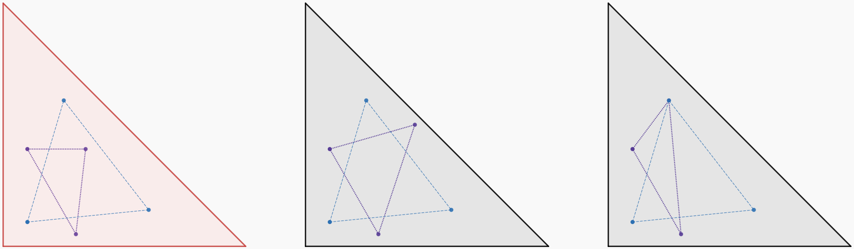
Yoder (2022): $\left(\hat{V} - V\right)$'s convexity $\Rightarrow$ the intersection of the support of any $\hat{\Phi}^*$ with the convex hull of the support of any Φ^* is a (possibly empty) subset of the extreme points of the convex hull of the support of Φ^*.

Implies the lemma.

Denti (2022): an (almost) stronger result as well.

Both: vaguely $\hat{\Phi}^*$ more extreme than Φ^*. Rely on posterior-separability.

One wrong and two rights



Proof step 4

Lemma. The transformation does not generate less information acquisition $\Rightarrow \hat{V} - V$ is convex.

Proof. Contraposition. Suppose $\hat{V} - V$ isn't convex.

Let $\rho(x)$ be some strictly convex continuous function on $\Delta(\Theta)$ and for arbitrary $\varepsilon > 0$, define

$$c_\varepsilon(x) := \varepsilon\rho(x) + \hat{V}(x).$$

For all sufficient small ε, can find $\mu \in \operatorname{int} \Delta\Theta$ s.t. DM with V acquires strictly more info than with $\hat{V}$ (latter acquires nothing).

Proof step 5 (FINAL STEP)

Lemma. The transformation generates a greater value for information $\Rightarrow$ $\hat{V} - V$ is convex.

Proof. Again by contraposition. Take an optimal $F \neq \delta_\mu$ for V from previous slide/Lemma and uniquely optimal δ_μ for $\hat{V}$. By construction,

$$\mathbb{E}_F V - V(\mu) > C(F) > \mathbb{E}_F \hat{V} - \hat{V}(\mu).$$

Acquiring some information redux

Say that the transformation **generates more information acquisition** if for any prior $\mu \in \operatorname{int}\Delta(\Theta)$, UPS cost functional D, and optimal solution to the agent's information acquisition problem when her value function is V, Φ_V^*, there exists an optimal solution to the agent's information acquisition problem when her value function is $\hat{V}$, $\Phi_{\hat{V}}^*$, that is a mean-preserving spread of Φ_V^*.

DM wants to acquire more information.

Two states and more (info acquisition)

Theorem. If $|\Theta| = 2$, the following are equivalent:

1. The transformation generates a greater value for information.
2. The transformation generates more information acquisition.
3. $\hat{V} - V$ is convex.
4. ⋆ $\hat{C} \geq C$ ⋆

Follows from Yoder (2022) and Curello & Sinander (2022).

Why two states?

More than two states and more (info acquisition)

Proposition. If $|\Theta| \geq 3$, a transformation generates more information acquisition if and only if $\hat{V} - V$ and/or V is affine.

Sufficiency is immediate.

Necessity: can always find a cost function to make binary learning optimal, *but on different line segments.*

Beyond Adding an Action

Adding multiple actions

Agent **becomes much more flexible** when A enlarged to $\hat{A} := A \cup B$, where B is a finite subset of $\mathcal{A} \setminus A$, and $\hat{u} = u$.

Adding multiple actions

Remark. $\hat{V} - V$ convex $\Rightarrow \hat{C} \geq C$.

$\hat{C} \geq C \Rightarrow \hat{V} - V$ convex?

No

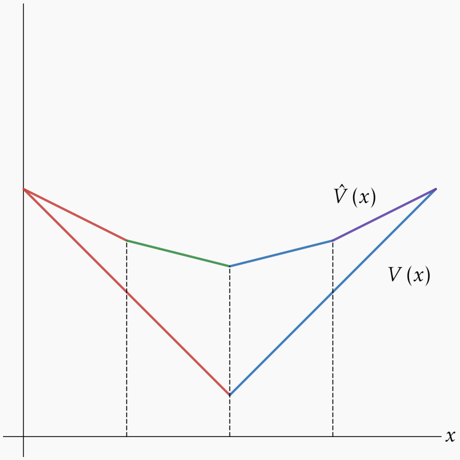


Figure: $\hat{C} \succ C$ But $\hat{V} - V$ Not Convex

A sufficient condition via subdivisions

Set of actions being added, B, is **totally refining** if each $b \in B$ is refining.

B is totally refining $\Rightarrow \hat{V} - V$ is convex. Converse is false

Proposition. Much more flexibility generates a greater value for information and does not generate less information acquisition if the set of additional actions is totally refining.

Non-necessity of total refinement

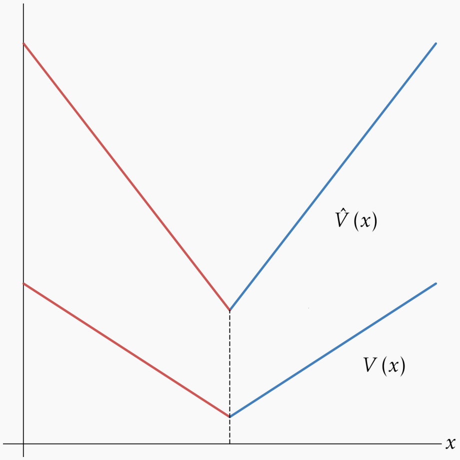


Figure: $\hat{V} - V$ Convex But B Not Totally Refining

Amost-necessity of total refinement

For B, understand $u \in \mathbb{R}^{B \times \Theta}$

Denote $\hat{V}_u$ new value function, given u.

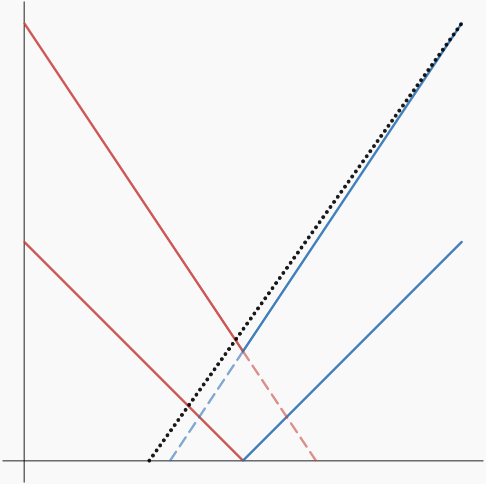
Making the agent much more flexible generically generates a greater value for information and does not generate less information if $\hat{V}_{\tilde{u}} - V$ is convex for all $\tilde{u}$ in an open ball around u.

Fragility

Proposition. Making the agent much more flexible generically generates a greater value for information and does not generate less information acquisition only if the set of additional actions is totally refining.

Proof. $\hat{C} \succeq C$ is necessary for $\hat{V} - V$'s convexity. But if B not totally refining, value function is partially "lifted up." Replicates subdivision but this is fine-tuned system of equations. Can always perturb and shatter $\hat{C} \succeq C$.

Fragility illustration



Removing an action

Agent **becomes much less flexible** if set of actions reduced from A to $\hat{A} \neq \emptyset$.

What makes information more valuable?

Nothing

Proposition. Making the agent much less flexible does not generate a greater value for information and may generate less information acquisition.

Proof. Take $x \in \text{int}\, C_i$ for some removed a_i (so $\hat{V}(x) < V(x)$), and $\mu \neq x'$ for which $V(\mu) = \hat{V}(\mu)$, $V(x') = \hat{V}(x')$ and $\mu \in \ell(x, x')$.

Then,

$$\lambda \underbrace{\left(\hat{V}(x) - V(x)\right)}_{<0} + (1-\lambda)\underbrace{\left(\hat{V}(x') - V(x')\right)}_{=0} - \underbrace{\left(\hat{V}(\mu) - V(\mu)\right)}_{=0} < 0,$$

so $\hat{V} - V$ is not convex.

Transforming the agent's utility

Agent's utility is transformed if set of actions stays the same but $\hat{u} = \phi \circ u$ for strictly increasing ϕ.

What sorts of transformations make $\hat{V} - V$ convex?

Affine transformations of u

$u \mapsto \alpha u + \beta =: \hat{u}$, where $\alpha \in \mathbb{R}_{++}$ and $\beta \in \mathbb{R}$.

Obviously subdivision is preserved, but what about $\hat{V} - V$?

Proposition. A positive affine transformation of the agent's utility function, u, generates a greater value for information and does not generate less information acquisition if and only if $\alpha \geq 1$.

Need utilities "scaled up."

What scales up utilities?

1. Direct manipulation (by, e.g., a principal). Of course.

2. Repetition: repeat a decision problem > 1 times.

3. Aggregate risk with CARA utility.

Aggregate risk + CARA

$\Theta \subset \mathbb{R}_+$.

Agent's endowed wealth is (finite-mean) random variable $Y \sim H$, uncorrelated with θ.

Utility function over terminal wealth, w, is (CARA): $v(w) = -\exp(-\gamma w)$, w/ $\gamma \in \mathbb{R}_{++}$.

In the language of this paper,

$$u(a,\theta) = -\int \exp(-\gamma(f_a(\theta) + y))\, dH(y) = -\exp(-\gamma f_a(\theta)) \int \exp(-\gamma y)\, dH(y).$$

Aggregate risk + CARA

Well known: CARA utility means wealth/aggregate risk does not affect decision-making.

A change to Y's distribution just scales u *linearly*.

Aggregate risk increases if H transformed to $\hat{H} \in MPS(H)$.

$$u \mapsto \underbrace{\frac{\int \exp(-\alpha y)\, d\hat{H}(y)}{\int \exp(-\alpha y)\, dH(y)}}_{\alpha} u =: \hat{u}.$$

CARA + Aggregate risk

Corollary. For an agent with CARA utility, increased aggregate risk generates a greater value for information and does not generate less information acquisition.

Decision-making (in decision problem) unchanged, **but value of information changes.**

Application 1. Delegation

Extreme actions in delegation

Szalay (2005) "The Economics of Clear Advice and Extreme Options:"

Delegation problem with interval state space and action, say $[0, 1]$.

Principal and agent with common quadratic loss–*ex post* agreement on optimal action.

Agent pays private cost to acquire information.

Delegation problem: WLOG for principal to allow agent to choose action from closed subset of $[0, 1]$.

Optimal delegation: prohibit actions within a certain distance around mean, i.e., delegation set is $[0, \alpha] \cup [\beta, 1]$, $0 < \alpha < \beta < 1$.

Our delegation problem

Similar problem: same utility function for principal and agent.

Agent can acquire information by paying some cost $\gamma > 0$ to see the realization of some signal.

Agent initially has (finite) action set A (in which no action is weakly dominated).

Principal wants to know whether to give agent access to an additional finite set of actions, B, before the agent acquires information.

Remark. The principal prefers to give the agent access to an additional set of actions, B, if it is totally refining.

But is there any subset of actions that the principal would like to remove?

Application 2. Monopolistic Screening

Selling Information

Principal (monopolist) and agent.

Two States.

Agent's type ω_i corresponds to her set of available actions $A_i \subseteq \mathcal{A}$.

Just two types, $\omega_1 > \omega_2$.

$V_1 - V_2$ is convex.

Principal and agent share a common prior $\mu \in \text{int}\,\Delta(\Theta)$

Principal can "produce" any distribution over posteriors Φ subject to a UPS cost $D(\Phi)$.

By the revelation principle, she offers a contract $((t_1, \Phi_1), (t_2, \Phi_2))$.

Selling Information: First Best

Principal solves

$$\max_{\Phi_1 \in \mathcal{F}(\mu)} \left\{ \int_0^1 V_1(x)\, d\Phi_1(x) - \kappa D(\Phi_1) \right\}, \quad \text{and} \quad \max_{\Phi_2 \in \mathcal{F}(\mu)} \left\{ \int_0^1 V_2(x)\, d\Phi_2(x) - \kappa D(\Phi_2) \right\},$$

and charges each type a price produced by that type's binding participation constraint.

$V_1 - V_2$ convex $\Rightarrow$ ω_1 is provided with "higher quality" than type ω_2: $\Phi_{1,FB}$ is an MPS of $\Phi_{2,FB}$.

$t_1 \geq t_2$.

Selling info: second best

IR_2 and IC_1 bind (as usual). Principal's objective reduces to

$$(1-\rho)\left(\frac{1}{1-\rho}\int_0^1 \left(V_2(x)-\rho V_1(x)\right) d\Phi_2(x) - \kappa D(\Phi_2)\right) + \rho\left(\int_0^1 V_1(x)\, d\Phi_1(x) - \kappa D(\Phi_1)\right),$$

where $\rho := \mathbb{P}(\omega_1)$.

$V_2 - \frac{V_2-\rho V_1}{1-\rho}$ is convex $\Rightarrow \Phi_{2,SB}$ an MPC of $\Phi_{2,FB}$.

Downward distortion for the "low" type relative to the first-best optimum.

$\Phi_{1,SB} = \Phi_{1,FB}$. *No output (quality of information) distortion at the top.*

Related Work & Conclusion

Related work

Value of information: Blackwell (1951, 1953), Athey & Levin (2018), De Lara & Gossner (2020), Radner & Stiglitz (1984), De Lara & Gilotte (2007), and Chade & Schlee (2002).

Rational inattention: Especially Caplin & Martin (2021):

- (Binary) relation between joint distributions over actions and states.
- One such joint distribution dominates another if for every utility function, every experiment consistent with the former is more valuable than every experiment consistent with the latter.
- Here, a partial order over (equivalence classes of) value functions: one dominates another if information must be more valuable for the former.

Related work

Comparative Statics: Especially Yoder (2022) and Curello & Sinander (2022): what changes to a persuader's indirect payoff lead to greater (or no less) information provision?

Regular Polyhedral Subdivisions: Kleiner, Moldovanu, Strack, & yt (2023*).

Risk Aversion: Pease, & yt (2023): binary relation between actions in a decision problem. What actions have beliefs comparatively robust to increased risk aversion?

All in all,

"Right" notion of convexity for comparing utility functions: $u = \phi \circ \hat{u}$ for monotone concave ϕ.

"Right" notion of convexity for comparing decision problems: $\hat{V} - V$ is convex.

Thanks for coming!