

Comparing Experiments in Discounted Problems

Ludovic Renou[♣]

Xavier Venel[◇]

Cole Williams[♣]

Mark Whitmeyer[♣]

♣ Arizona State

◇ LUISS

♣ Nebraska

USC
April 24, 2026

Dynamic decision problems

States: Θ , w/ generic element θ .

Time: $t = 1, \dots, T$ ($T \leq \infty$).

Dynamic information structures:

$$f: \Theta \rightarrow \Delta(X_1 \times \dots \times X_T), \quad \text{vs.} \quad g: \Theta \rightarrow \Delta(Y_1 \times \dots \times Y_T).$$

Dynamic decision problem: @ each t DM chooses $a_t \in A_t$ after observing past and current signals... *but not future ones.*

Utility: $u(a_1, \dots, a_T, \theta)$.

All sets finite, except (possibly) # of pds.

Discounted problems

Additively-separable (A-S) problem:

$$u(a_1, \dots, a_T, \theta) = \sum_{t=1}^T u_t(a_t, \theta),$$

for some cllxn. $(u_t: A_t \times \Theta \rightarrow \mathbb{R})_t$

Discounted problem: Common $A = A_t \forall t$, common $u: A \times \Theta \rightarrow \mathbb{R}$, &

$$u(a_1, \dots, a_T, \theta) = \sum_{t=1}^T \delta_t u(a_t, \theta), \quad \text{for some } (\delta_t)_t \in \Delta(\{1, \dots, T\}).$$

Misc. subclasses: all, decreasing, decreasing & convex, exponential, fixed.

Comparisons of info structures

Static problems (dynamic w/ $T = 1$).

Blackwell ('51, '53).

Dynamic problems.

Greenshtein ('96).

Discounted problems.

This talk.

Questions

1. Comparisons of info structures for discounted problems?
2. De Jarnette, Dillenberger, Gottlieb, Ortoleva ('20): attitudes toward timing risk. Here: timing risk over info arrival?

Roadmap

1. Brief review of Blackwell ('51, '53) & Greenshtein ('96).
2. Some examples.
3. Discounted problems: fixed discount factor.
4. Discounted problems: sets of discount factors + AS problems.
5. Risk preferences over info arrival.
6. Preview: more questions.

Review

Review of Blackwell ('51, '53)

Single pd. dynamic problem: A finite set Θ , w/ generic element θ .

Information: Two statistical experiments $f: \Theta \rightarrow \Delta(X)$ and $g: \Theta \rightarrow \Delta(Y)$ to be compared.

Decision problem: (A, u) , where A is the set of decisions and $u: A \times \Theta \rightarrow \mathbb{R}$ the utility/loss function.

Strategies: $\sigma: X \rightarrow \Delta(A)$ and $\tau: Y \rightarrow \Delta(A)$.

Distributions: $\mathbb{P}_{\theta, f, \sigma}$ and $\mathbb{P}_{\theta, g, \tau}$ are the induced distributions over signals and actions (given state θ).

Review of Blackwell ('51, '53)

Definition

The experiment f is *more valuable* than the experiment g if for all decision problems (A, u) , for any strategy τ , there exists a strategy σ such that

$$\mathbb{E}_{\theta, g, \tau}[u(\mathbf{a}, \theta)] \leq \mathbb{E}_{\theta, f, \sigma}[u(\mathbf{a}, \theta)] \quad \forall \theta$$

Key: The information is fixed, but *all* decision problems are considered.

Remark: No prior is considered, hence “for all θ .”

Given a prior $\mu_0 \in \Delta(\Theta)$, f induces (ex-ante) distribution over posteriors $F \in \Delta(\Delta(\Theta))$ w/ $\mathbb{E}_F(\mu) = \mu_0$.

Review of Blackwell ('51, '53)

Marginal dist. over actions: $(\text{marg}_A \mathbb{P}_{\theta, g, \tau})(a) := \sum_{y \in Y} \mathbb{P}_{\theta, g, \tau}(y, a)$.

Theorem

T.f.a.e:

- 1. The experiment f is more valuable than the experiment g .*
- 2. For all decision problems (A, u) , for all τ , there exists σ such that*

$$\text{marg}_A \mathbb{P}_{\theta, g, \tau} = \text{marg}_A \mathbb{P}_{\theta, f, \sigma}, \quad \forall \theta.$$

- 3. Sufficiency: There exists a garbling $\gamma: X \rightarrow \Delta(Y)$ such that*

$$g(y|\theta) = \sum_x \gamma(y|x) f(x|\theta), \quad \forall y, \forall \theta$$

- 4. For all $\mu_0 \in \Delta(\Theta)$, $F \succeq_{cx} G$.*

Back to dynamics

Recall $f: \Theta \rightarrow \Delta(X_1 \times \dots \times X_T)$ & utility $u(a_1, \dots, a_T, \theta)$.

Write

$$X^t = X_1 \times \dots \times X_t, \quad \& \quad A^t = A_1 \times \dots \times A_t.$$

Strategy: a stochastic map

$$\alpha: X^T \rightarrow \Delta(A^T)$$

that is **Adapted:** for each t , the marginal

$$\alpha^t(a_1, \dots, a_t \mid x_1, \dots, x_T) = \sum_{a_{t+1}, \dots, a_T} \alpha(a_1, \dots, a_T \mid x_1, \dots, x_T)$$

depends only on (x_1, \dots, x_t) .

$$\Theta \xrightarrow{f} X^T \xrightarrow{\alpha} A^T.$$

Key observation: strategies and garblings are the same type of object.

Review of Greenshtein ('96)

Definition

The dynamic information structure f is *more informative* than g if for all dynamic decision problems and every adapted strategy τ for g , there exists an adapted strategy σ for f such that

$$\mathbb{E}_{\theta,g,\tau}[u(\mathbf{a}_1, \dots, \mathbf{a}_T, \theta)] \leq \mathbb{E}_{\theta,f,\sigma}[u(\mathbf{a}_1, \dots, \mathbf{a}_T, \theta)] \quad \forall \theta.$$

Key: mirrors Blackwell, only now the actions and signals arrive over time.

Feasible distributions over action histories:

$$\Lambda_f(A_1, \dots, A_T) = \{\alpha \circ f: \alpha \text{ adapted}\}.$$

Expected utility depends only on the element of $\Lambda_f(A_1, \dots, A_T)$ that is induced.

Review of Greenshtein ('96). Also De Oliveira ('17)

Theorem

T.f.a.e:

1. f is more informative than g in every dynamic decision problem.
2. For every action space $(A_t)_t$,

$$\Lambda_f(A_1, \dots, A_T) \supseteq \Lambda_g(A_1, \dots, A_T).$$

3. There exists an *adapted garbling*

$$\gamma: X^T \rightarrow \Delta(Y^T)$$

s.t. $g = \gamma \circ f$.

4. Sequences of distributions over posteriors?

Blackwell's siren call

For each t , view dynamic experiment f as static experiment $f^t \rightarrow \Delta(X^t)$.

Why not $f^t \succeq_B g^t$ for all t ?

Equivalently, $F^t \succeq_{cX} G^t$ for all $\mu_0 \in \Delta(\Theta)$, for all t .

Examples

Example 1: a stopping problem

Two states and two actions:

$$\theta \in \{0, 1\}, \quad a \in \{0, 1\}, \quad \& \quad \mu_0 = \mathbb{P}(\theta = 1).$$

Two conditionally-independent binary signals: *Weak* and *Strong*:

	$\mathbb{P}(W = 1 \mid \theta)$	$\mathbb{P}(S = 1 \mid \theta)$
$\theta = 0$	1/3	1/4
$\theta = 1$	2/3	3/4

Two sequential experiments are

$$f = (W, S), \quad \& \quad g = (S, W).$$

The first signal is “free.” Observing the second signal costs $c = 1/100$.

Example 1: utilities

Payoffs:

	$\theta = 0$	$\theta = 1$
$a = 0$	1	0.9
$a = 1$	0	1

Plus: if the second signal is bought, subtract $c = 1/100$.

$$v(q) = \max\left\{1 - \frac{q}{10}, q\right\}.$$

Static cutoff:

$$a = 1 \iff q \geq \frac{10}{11}.$$

Example 1: decisions

Set $\mu_0 = \mathbb{P}(\theta = 1) = \frac{4}{5}$.

B/c second signal is costly, the agent may “stop” after the first signal.

Optimal continuation behavior:

	lower first posterior	higher first posterior
$f = (W, S)$	$2/3$: stop	$8/9$: buy S
$g = (S, W)$	$4/7$: stop	$12/13$: buy W

Supports are

$$f: \frac{4}{5} \rightarrow \left\{ \frac{2}{3}, \frac{8}{9} \right\} \rightarrow \left\{ \frac{2}{5}, \frac{8}{11}, \frac{6}{7}, \frac{24}{25} \right\},$$

$$g: \frac{4}{5} \rightarrow \left\{ \frac{4}{7}, \frac{12}{13} \right\} \rightarrow \left\{ \frac{2}{5}, \frac{8}{11}, \frac{6}{7}, \frac{24}{25} \right\}.$$

Example 1: discussion

Continuation decision makes the “effective” info structures incomparable.

S better than W , but not much better as a screening device.

Comparative value higher as a decision tool for high beliefs.

Example 1.5: discounting

Previous decision problem: not a discounted problem.

Not A-S.

Take arbitrary A-S problem; for $t = 1, 2$ and $p \in [0, 1]$ write
 $V_t(p) := \max_{a_t} \mathbb{E}_p u_t(a_t, \theta)$.

$$\mathbb{E}_{G_1} V_1 \geq \mathbb{E}_{F_1} V_1 \ \& \ \mathbb{E}_{G_2} V_2 \geq \mathbb{E}_{F_2} V_2 \quad \implies \quad g \text{ better for AS problems!}$$

$\implies g$ superior to f for discounted problems (all varieties).

Example 2: info arrival

Fix an *informative* Blackwell experiment $\pi: \Theta \rightarrow \Delta(S)$.

R.V. \mathcal{T} : realized t is when the DM gets to see draw from π .

Discounted problems:

$$u(a_1, \dots, a_T, \theta) = \sum_{t=1}^T \delta_t u(a_t, \theta), \quad \text{for some } (\delta_t)_t \in \Delta(\{1, \dots, T\}).$$

Risk attitudes toward arrival? E.g., $T = 3$: prefer @ $t = 2$ for sure vs. 50/50 1 or 3.

Example 2: consumption streams benchmark

De Jarnette, Dillenberger, Gottlieb, Ortoleva ('20): prize x @ random time w/
 $u(x) > u(0) = 0$.

Compare $\delta_t u(x)$ w/ $\mathbb{E}\delta_T u(x)$.

Shape of δ_t v. important: exponential $\delta_t = \beta^t$ is convex \implies risk-seeking over time lotteries.

What about over information?

Example 2: 3 periods

Suppose $T = 3$ and $\delta_1 = 1/2$ $\delta_2 = 2/5$ & $\delta_3 = 1/10$.

Re. De Jarnette et. al.: set $u(x) = 1$.

$$\frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{10} = \frac{3}{10} < \frac{4}{10} = 1 \times \frac{2}{5} \implies \text{Risk averse!}$$

Now decision problem: π informative, VOI positive for some us strictly so.
Normalize no info value to 0, value to $\alpha \geq 0$.

$$\left[\frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{10} \right] \alpha \quad \underbrace{+ \frac{1}{2} \left[\frac{2}{5} + \frac{1}{10} \right] \alpha}_{\text{WHY?}} \quad \text{vs.} \quad \frac{2}{5} \alpha \quad \underbrace{+ \frac{1}{10} \alpha}_{\text{WHY?}}$$

$$\frac{11}{20} \alpha > \frac{10}{20} \alpha \implies \text{Risk seeking!}$$

The Main Result

The discounted model – reminder

States: A set Θ , with generic element θ .

Horizon: $T \leq +\infty$ periods, w/ generic element t .

Information: Two statistical experiments $f := (f_t)_t$ and $g := (g_t)_t$ w/

$$f_t: \Theta \times X^{t-1} \rightarrow \Delta(X_t), \quad \& \quad g_t: \Theta \times Y^{t-1} \rightarrow \Delta(Y_t)$$

A-S Decision problems: $(A_t, u_t)_t$, w/ A_t is the set of decisions @ t , $u_t: A_t \times \Theta \rightarrow \mathbb{R}$ the utility @ t .

Notation. $X^t := X_1 \times \dots \times X_t$, etc.

The discounted model – the baseline

Discounting: The discount factor δ is a probability distribution over $\{1, \dots, T\}$.

E.g.,

1. Finite horizon.
2. Geometric discounting. ($T = \infty$ & $(1 - \beta)\beta^{t-1}$ for some $\beta \in [0, 1)$.)
3. Quasi-hyperbolic discounting.

The discounted model – reminder

Discounted problem: $(A_t, u_t) = (A, u)$ for all t , & DM evaluates $(a_t)_t$ as

$$\sum_t \delta_t u(a_t, \theta).$$

A-S problem: DM evaluates $(a_t)_t$ as

$$\sum_t u_t(a_t, \theta).$$

Strategies:

$$\sigma_t: X^t \times A^{t-1} \rightarrow \Delta(A_t) \quad \& \quad \tau_t: Y^t \times A^{t-1} \rightarrow \Delta(A_t).$$

Distributions: $\mathbb{P}_{\theta, f, \sigma}$ and $\mathbb{P}_{\theta, g, \tau}$ are the induced distributions over signals and actions (given θ).

Discounted problems – comparing experiments

Definition

Let δ be the discount factor. The statistical experiment f is *more valuable* than g in discounted problems, if for **all decision problems** (A, u) , for any strategy τ , there exists strategy σ such

$$\mathbb{E}_{\theta, g, \tau} \left[\sum_t \delta_t u(\mathbf{a}_t, \theta) \right] \leq \mathbb{E}_{\theta, f, \sigma} \left[\sum_t \delta_t u(\mathbf{a}_t, \theta) \right], \quad \forall \theta$$

Key: Discount factor is fixed in the basic model.

Aim: Derive the analog of Blackwell's theorem.

Discounted problems – a lemma

Time-weighted marginal dist. over actions:

$$(\sum_t \delta_t \text{marg}_{A_t} \mathbb{P}_{\theta, g, \tau})(a) := \sum_t \sum_{y^t \in Y^t} \delta_t \mathbb{P}_{\theta, g, \tau}(y^t, a)$$

Note:

$$\mathbb{E}_{\theta, f, \sigma} \left[\sum_t \delta_t u(\mathbf{a}_t, \theta) \right] = \sum_a u(a, \theta) \underbrace{\left[\sum_t \delta_t \text{marg}_{A_t} \mathbb{P}_{\theta, f, \sigma}(a) \right]}_{\text{discounted prob. of "a"}}.$$

Hence, as in static problems (equivalence between (1) and (2) in Blackwell's theorem), we have:

Lemma. f is more valuable than g in all discounted problems if, and only if, for all τ , there exists σ , such that

$$\sum_t \delta_t \text{marg}_{A_t} \mathbb{P}_{\theta, g, \tau} = \sum_t \delta_t \text{marg}_{A_t} \mathbb{P}_{\theta, f, \sigma}, \quad \forall \theta.$$

Discounted problems – the baseline

Now want to prove an equivalence similar to the one between (2) and (3) in Blackwell's theorem, *viz.*, via garblings.

From the lemma, we can restrict attention to ensuring that:

$$\sum_t \delta_t \text{marg}_{A_t} \mathbb{P}_{\theta, g, \tau} = \sum_t \delta_t \text{marg}_{A_t} \mathbb{P}_{\theta, f, \sigma},$$

i.e., matching the discounted probabilities of decisions.

Discounted problems – the baseline

We have:

$$\sum_t \delta_t \text{marg}_{A_t} \mathbb{P}_{\theta, f, \sigma}(a) = \sum_t \sum_{x^t \in X^t} \delta_t \mathbb{P}_{\theta, f, \sigma}(x^t) \mathbb{P}_{\theta, f, \sigma}(a|x^t).$$

Moreover, $\mathbb{P}_{\theta, f, \sigma}(x^t)$ does not depend on σ and is equal to:

$$f^t(x^t|\theta) := f_1(x_1|\theta) \times \dots \times f_t(x_t|x^{t-1}, \theta).$$

So, we can rewrite the above in the form:

$$\sum_t \sum_{x^t \in X^t} \underbrace{\delta_t f^t(x^t|\theta)}_{\text{info}} \underbrace{\bar{\sigma}(a|t, x^t)}_{\text{strategy}}.$$

Punchline: Back to a static problem, with the reinterpretation of a signal as being “ (t, x^t) ” with prob. $\delta_t f^t(x^t|\theta)$, when state is θ .

Similarly,

$$\sum_t \delta_t \text{marg}_{A_t} \mathbb{P}_{\theta, g, \tau}(a) = \sum_t \sum_{y^t \in Y^t} \delta_t g^t(y^t | \theta) \bar{\tau}(a | t, y^t).$$

From the equivalence between (2) and (3) in Blackwell's theorem, we therefore need the existence of a garbling γ satisfying

$$\delta_t g^t(y^t | \theta) = \sum_{t', x^{t'}} \delta_{t'} f^{t'}(x^{t'} | \theta) \gamma(\underbrace{t, y^t}_{\text{"simulations"}} | t', x^{t'}),$$

for all θ , for all (t, y^t) .

We call this condition **δ -sufficiency**.

Intuition. It does not matter when "a" is chosen, as long as the averages match.

Discounted problems – the baseline

Example: Info today vs. info tomorrow.

Two periods, equally discounted $(\delta_1, \delta_2) = (1/2, 1/2)$.

- ▶ f : no info in period 1, but full info in period 2.
- ▶ g : some info in period 1, but no further info in period 2.

For a numerical example,

$\theta \backslash x$	x_0	x_1
θ_0	1	0
θ_1	0	1

Table: f

$\theta \backslash y$	y_0	y_1
θ_0	7/12	5/12
θ_1	5/12	7/12

Table: g

Clearly, f is more informative than g in static problems, but is it “sufficiently more informative”?

Discounted problems – the baseline

δ -sufficiency amounts to comparing a modified version of f with g , where

$\theta \backslash x$	x_0	x_1	\emptyset
θ_0	1/2	0	1/2
θ_1	0	1/2	1/2

Table: the modified f

Modified f : with prob.1/2, no info; with prob.1/2, full info.

Answer: The “modified f ” is sufficient for g , hence f is δ -sufficient for g .

Alternative approach

Given $\mu_0 \in \Delta(\Theta)$, sequences of distributions over posteriors:

$$(F_t)_t \quad \text{vs.} \quad (G_t)_t.$$

Define

$$F^\delta := \sum_{t=1}^T \delta_t F_t \quad \& \quad G^\delta := \sum_{t=1}^T \delta_t G_t.$$

Proposition

f m.v. than g in discounted problems w/ factor δ if and only if $F^\delta \succeq_{cx} G^\delta \forall \mu_0$.

Alternative approach

Proposition

f m.v. than g in discounted problems w/ factor δ if and only if $F^\delta \succeq_{cx} G^\delta \forall \mu_0$.

proof.

$$\begin{aligned} \sum_{t=1}^T \delta_t \mathbb{E}_{F_t} V(\mu) \geq \sum_{t=1}^T \delta_t \mathbb{E}_{G_t} V(\mu) \forall \text{ convex } V &\iff \mathbb{E}_{\sum_{t=1}^T \delta_t F_t} V(\mu) \geq \mathbb{E}_{\sum_{t=1}^T \delta_t G_t} V(\mu) \\ &\iff \mathbb{E}_{F^\delta} V(\mu) \geq \mathbb{E}_{G^\delta} V(\mu). \end{aligned}$$

Discounted problems – the basics

Theorem

T.f.a.e:

1. f m.v. than g in discounted problems w/ factor δ .
2. For all decision problems (A, u) , for all strategies τ , there exists a strategy σ s.t.

$$\sum_t \delta_t \text{marg}_{A_t} \mathbb{P}_{\theta, g, \tau} = \sum_t \delta_t \text{marg}_{A_t} \mathbb{P}_{\theta, f, \sigma}, \quad \forall \theta.$$

3. δ -sufficiency: There exist garbling γ s.t.

$$\delta_t g^t(y^t | \theta) = \sum_{t'} \sum_{x^{t'} \in X^{t'}} \delta_{t'} f^{t'}(x^{t'} | \theta) \gamma(t, y^t | t', x^{t'}), \quad \forall (t, y^t), \forall \theta.$$

4. For all $\mu_0 \in \Delta(\Theta)$, $F^\delta \succeq_{cX} G^\delta$.

Sets of discount factors + AS problems

Sets of discount factors

Use extreme points

1. (Weakly) decreasing discount factors:

$$\Delta_{\downarrow} := \{\delta \in \Delta(\{1, \dots, T\}) : \delta_1 \geq \delta_2 \geq \dots \geq \delta_T\}$$

$(1, 0, \dots)$, $(1/2, 1/2, 0, \dots)$, $(1/3, 1/3, 1/3, 0, \dots)$, etc.

2. All discount factors:

$$\Delta_{\updownarrow} := \Delta(\{1, \dots, T\})$$

$(1, 0, \dots)$, $(0, 1, 0, \dots)$, etc.

3. Exponential discount factors in $(0, 1)$.

$$\Delta_{\beta} := \left\{ \delta^{\beta} : \delta_t^{\beta} = (1 - \beta)\beta^{t-1}, \beta \in (0, 1) \right\}.$$

Not convex!

Sets of discount factors

Proposition

f m.v. than g in discounted problems over $\Delta_{\downarrow} \iff \sum_{i=1}^t \frac{1}{t} f^i \succeq_B \sum_{i=1}^t \frac{1}{t} g^i \forall t.$

Proposition

f m.v. than g in discounted problems over $\Delta_{\uparrow} \iff f^t \succeq_B g^t \forall t \iff f$ m.v. than g in AS problems.

Time Lotteries over Info

Time lotteries over info

Fix static experiments $\xi: \Theta \rightarrow \Delta(Z)$ and $\eta: \Theta \rightarrow \Delta(W)$. Let H and P be cdfs on $\{1, \dots, T\}$, with arrival times Y_H and Y_P .

Dynamic experiment $\mu^{H,\xi}$: draw $Y_H \sim H$ and $z \sim \xi(\cdot | \theta)$ independently, and set

$$x_t = \begin{cases} z, & t = Y_H, \\ \emptyset, & t \neq Y_H. \end{cases}$$

Define $\mu^{P,\eta}$ analogously. If $\xi = \eta$, write μ^H and μ^P .

Fix prior $\mu_0 \in \Delta(\Theta)$.

Notation. $\mu^{H,\xi} \geq_{\Delta} \mu^{P,\eta}$ means: for every $\delta \in \Delta$ and every δ -discounted problem, $\mu^{H,\xi}$ gives weakly larger value.

Time lotteries over info: the key reduction

For a static experiment ζ , define its static value

$$W(\zeta; u) = \mathbb{E} \left[\max_{a \in A} \mathbb{E}[u(a, \theta) \mid s] \right] - \max_{a \in A} \mathbb{E}[u(a, \theta)].$$

Set $w_\delta(y) := \sum_{t \geq y} \delta_t$.

If the signal arrives at time y , the gain is $w_\delta(y)W(\zeta; u)$: once the signal arrives, it can be used in every remaining period.

Proposition

Value of $\mu^{H, \xi} = \mathbb{E}_H [w_\delta(Y_H)] W(\xi; u)$, and value of $\mu^{P, \eta} = \mathbb{E}_P [w_\delta(Y_P)] W(\eta; u)$

So the timing question is about the shape of w_δ , not the shape of δ_t !

Back to example 2

For $T = 3$ and $(\delta_1, \delta_2, \delta_3) = \left(\frac{1}{2}, \frac{2}{5}, \frac{1}{10}\right)$,

$$w_\delta(1) = 1, \quad w_\delta(2) = \frac{1}{2}, \quad w_\delta(3) = \frac{1}{10}.$$

So if π is informative and $W(\pi; u) = \alpha > 0$,

$$\frac{1}{2}w_\delta(1)\alpha + \frac{1}{2}w_\delta(3)\alpha = \frac{11}{20}\alpha > \frac{1}{2}\alpha = w_\delta(2)\alpha.$$

Thus: 50/50 arrival at 1 or 3 beats sure arrival at 2.

By contrast, for a prize received at a random time,

$$\frac{1}{2}\delta_1 + \frac{1}{2}\delta_3 = \frac{3}{10} < \frac{2}{5} = \delta_2.$$

Same discount factor, opposite timing attitude.

Time lotteries over info: characterizations

Write $H(t) = \mathbb{P}(Y_H \leq t)$ and $P(t) = \mathbb{P}(Y_P \leq t)$.

Proposition

$\mu^{H,\xi} \succeq_{\Delta} \mu^{P,\eta}$ if and only if the corresp. ineq. below holds $\forall (A, u)$.

1. All discount factors Δ_{\uparrow} :

$$H(t)W(\xi; u) \geq P(t)W(\eta; u) \quad \forall t.$$

2. Decreasing discount factors Δ_{\downarrow} :

$$\sum_{t=1}^m H(t)W(\xi; u) \geq \sum_{t=1}^m P(t)W(\eta; u) \quad \forall m.$$

3. Exponential discounting $\Delta_{\beta}, T = \infty$:

$$\mathbb{E}[\beta^{Y_H}]W(\xi; u) \geq \mathbb{E}[\beta^{Y_P}]W(\eta; u) \quad \forall \beta \in (0, 1).$$

Same experiment, only timing differs

Assume $\xi = \eta$ and the common static experiment is nontrivial. Then the comparison depends only on the order of arrival times.

Corollary

$$\mu^H \succeq_{\Delta_{\downarrow}} \mu^P \iff Y_P \succeq_1 Y_H \quad (\text{FOSD})$$

$$\mu^H \succeq_{\Delta_{\downarrow}} \mu^P \iff Y_P \succeq_2 Y_H \quad (\text{increasing-concave order})$$

$$\mu^H \succeq_{\Delta_{\beta}} \mu^P \iff Y_P \succeq_{\text{pgf}} Y_H.$$

Interpretation. Impatience \implies Time-risk love. Shape (beyond decreasing) does not matter!

probability-generating-function order

If the underlying experiments also differ

Define

w/ conventions $0/0 := 0$ and $a/0 := +\infty$ for $a > 0$

$$r_{\uparrow} := \max_{1 \leq t \leq T} \frac{P(t)}{H(t)}, \quad r_{\downarrow} := \max_{1 \leq m \leq T} \frac{\sum_{t=1}^m P(t)}{\sum_{t=1}^m H(t)},$$

and, for $T = \infty$,

$$r_{pgf} := \sup_{\beta \in (0,1)} \frac{\mathbb{E}[\beta^{Y_P}]}{\mathbb{E}[\beta^{Y_H}]}.$$

Write $\xi \geq_r \eta$ if $W(\xi; u) \geq rW(\eta; u)$ for every decision problem (A, u) . Then

$$\mu^{H,\xi} \geq_{\Delta_{\uparrow}} \mu^{P,\eta} \iff \xi \geq_{r_{\uparrow}} \eta,$$

and analogously for Δ_{\downarrow} and Δ_{β} .

If the relevant $r \leq 1$, equivalent to ξ Blackwell-dominating the r -dilution of η .

Timing and informativeness separate multiplicatively.

Conclusion

In progress, r as a measure of comparative-VOI.

For discounted problems, aggregate then Blackwell.

Impatient DMs in discounted problems love risk.

Generalizations in the paper:

1. DM can (permanently) switch experiments after any history.
2. Actions affect info.
3. Actions affect discounting.
4. Time (In-)consistency?

Robust-to-outside-info analog is easy. AS and full class coincide.

Thank you!

Generalizations

Discounted problems – sequential comparisons

So far: assumed that the DM observes signals from either f or g (but not both).

As if the DM chooses e-a whether to observe the X -signals or the Y -signals and *cannot switch*.

Strengthen our comparison: give the DM the opportunity to permanently switch from observing the X -signals to observing the Y -signals after any history x^t of X -signals.

Viz., want the DM to prefer observing signals from f not only at the e-a stage, but at all (on-path) histories.

Discounted problems – sequential comparisons

Consider the following example: two periods, two states L (ow) and H (igh). Signals are ℓ (ow) and h (igh). For $\theta = L, H$,

$$f_1(\ell|L) = f_1(h|H) = \frac{3}{4}, f_2(h|h, \theta) = f_2(\ell|\ell, \theta) = 1,$$

$$g_1(\ell|L) = g_1(h|H) = \frac{3}{4}, g_2(h|h, \theta) = g_2(\ell|\ell, \theta) = 1.$$

The two experiments seem identical and, indeed, have the same law.

As random processes, they may differ, however.

E.g., if they are cond. independent, after observing the first-period signal from either of them, the DM always has an incentive to switch.

If perfectly correlated, the DM never has an incentive to switch. \implies Need to model the interdependence.

Discounted problems – control of information

So far, the DM cannot influence/control the information he receives.

But...selling a product with unknown demand, experimentation/bandit problems, POMDP, etc.

Natural idea: To introduce actions in the experiments:

$$f_t: \Theta \times X^{t-1} \times A^{t-1} \rightarrow \Delta(X_t),$$
$$g_t: \Theta \times Y^{t-1} \times A^{t-1} \rightarrow \Delta(Y_t).$$

Problem: As the decision problem varies, so do the experiments!

Discounted problems – control of information

Our proposal: To have experiments depend on fixed covariates/controls K_t , which themselves are influenced/controlled by actions in A_t .

Conceptually, the same as having preferences over consequences, and actions influencing the consequences, as in mechanism design.

Controlled information: Two statistical experiments $f := (f_t)_t$ and $g := (g_t)_t$ to be compared, where

$$f_t: \Theta \times X^{t-1} \times K^{t-1} \rightarrow \Delta(X_t),$$

$$g_t: \Theta \times Y^{t-1} \times K^{t-1} \rightarrow \Delta(Y_t),$$

and K_t is the set of covariates at period t .

Decision problems: $(A_t, u_t, \kappa_t)_t$, with $\kappa_t: K^{t-1} \times A_t \rightarrow \Delta(K_t)$ the “control map.”