## The Value of Information (for an expected-utility maximizer)

- Finite state space $S$; prior $p \in \Delta(S)$

- Finite set of actions $A$, utility function $u : S \times A$
  (or outcome function $\rho : S \times A \to Z$ and utility function $u : Z \to R$ .)

- Ex-ante optimal action is $a^0 := \arg\max_a \sum_{s \in S} p(s)u(s,a)$,
  value $V(p) := \max_a \sum_{s \in S} p(s)u(s,a)$ .

- $V$ is a convex function of $p$, and linear when the optimal action is constant:
  If $p = \lambda p' + (1-\lambda) p''$ then
  $V(p) = \max_a \sum_{s \in S} p(s)u(s,a) = \max_a \left[ \sum_{s \in S} \left( \lambda p'(s) + (1-\lambda) p''(s) \right) u(s,a) \right]$

  $\leq \lambda \max_{a'} \sum_{s \in S} p'(s)u(s,a') + (1-) \sum_{s \in S} p''(s)u(s,a'') = \lambda V(P') + (1-\lambda)V(p'')$

- Suppose agent can observe signals $y$ with Markov kernel $K$ before making a choice, so joint probability of signal $s$ and state $y$ is $P^K(s,y) = p(s)k(y|s)$, the probability of observing $y$ is $P^K(y) = \sum_{s\in S} p(s)k(y|s)$ , and the conditional probability of $s$ given $y$ is $P^K(s|y) = P^K(s,y)/P^K(y)$

- Agent will choose $a^*(y) := \arg\max_a \sum_{s\in S} p^K(s|y)u(s,a)$ after seeing $y$.

- Expected payoff of $K$ with prior $p$ is

$$V(K) = \sum_{s\in S}\sum_{y\in Y} P^K(s,y)u(s,a^*(y))$$

$$= \sum_{y\in Y} P^K(y)\left[\max_a \sum_{s\in S} P^K(s|y)u(s,a)\right].$$

- Value of signal structure $K$ is $I(K) := V(K) - V(\varnothing)$ . (*note that both V and I depend on p, u, and A; hold those fixed for now and so don't include them as arguments.)*

- $I(K) \geq 0$ as could always ignore the information, only use it if it's helpful. (*not true for biased learners, or for the equilibria of games…*)

- $K_1$ is called more informative than $K_2$ if $K_2$ is a garbling of $K_1$.

- Intuitively this means we can get from $K_1$ to $K_2$ by "adding noise."

- The signals generated by $K_1$ and $K_2$ could have different names, and different cardinalities.

- With finitely many states and signals can represent the signal structures as matrices whose rows are the probability distribution over signals in each state.

- Then $K_2$ is a garbling of $K_1$ if there is a stochastic matrix $Q$ s.t. $K_2 = QK_1$, so that $k_2(s, y_2) = \sum_{y_1} q(y_2 \mid y_1) k_1(y_1 \mid s)$: First observe $y_1$, then use the resulting signal to generate $y_2$.

- Garbling is a weak, partial order on signal structures.

- Agent can't gain from garbling; Blackwell shows that this condition is also a characterization: $K_2$ is a garbling of $K_1$ iff $V(K_1) \geq V(K_2)$ for all *u*,*p*, and *A*.

- Instead of explicitly dealing with the signals, can alternatively describe information structures by the induced distribution $\mu \in \Delta(\Delta(S))$ of posterior distributions on *S*, where $p = E_\mu p'$.

- $K_2$ is a garbling of $K_1$ iff the posteriors induced by $K_1$ are "more informative" in the sense that averaging over the posteriors according to $\mu_1$ yields the distribution of posteriors under $\mu_2$.

- For example with two states $s_1, s_2$, and $p(s_1) = p(s_2) = .5$ $\mu_2$ might assign probability 1/3 to posterior $(3/4, 1/4)$ and 2/3 to (3/8, 5/8) while $\mu_1$ assigns probability 1/3 to posterior $(3/4, 1/4)$, ¼ to (1,0) and 5/12 to (0,1).

- Here $\mu_1$ corresponds to getting an "extra signal" when the posterior is to (3/8, 5/8); a lot like a mean-preserving spread (but not on the reals).

------------------------------------

Next topic: apply the dual self and quasi-hyperbolic models to some simple dynamic decision problems.

Doing a 1-shot task (O'Donoghue and Rabin *QJE* [2001]

- Each period can act or not.
- Acting ends the game, costs $c$, pays $v$ next period.
- Don't act: same decision next period.

Exponential discounting: If $\delta v > c$ act immediately, $\delta v < c$ never act.

Naïve quasi-hyperbolic: If $\beta\delta v > c$ act immediately; if $c > \beta\delta c$ never act.

Sophisticated quasi-hyperbolic: In general sophisticated quasi-hyperbolic models are like many-player games, so they often have multiple equilibria in infinite horizon settings, and implausible "terminal period dependence" with a fixed and known horizon.

This can be the case here.

Specifically:

- If $c > \beta\delta v$ the unique equilibrium is "never act."

- If $\beta\delta v > c$ (so the current self gets a benefit from acting)

  and $\beta\delta v - c > \beta\delta^2 v - \beta\delta c$
  (the current self would rather act now then let the next self do it)

  then the unique equilibrium is "act immediately."

- Otherwise, let

$$d^* = \max\left\{d \in \{0,1,...\} \mid \beta\delta v - c < \beta\delta^d (\delta v - c)\right\}:$$

each self strictly prefers to act if future self won't act for $d^*$ or more periods, but prefers to wait if future self will act sooner than that.

- Family of pure-strategy subgame-perfect equilibria: pick a $T \le d^*$ , if the game hasn't stopped yet then act if $t = T, T + d^*, T + 2d^*, ..$ else wait.

  *Prove it by testing for 1-shot deviations: If supposed to stop at t then* $t = T + kd^*$ *for some integer k.* If stop get $\beta\delta v - c$, waiting gives $\beta\delta^{d^*}(\delta v - c)$ *which is less from the definition of* $d^*$*. And if supposed to wait at t then waiting gives* $\beta\delta^k(\delta v - c)$ *for some* $k < d^*$ *which is better …*

- With a known finite horizon *H* act at *H*, $H - d^*, H - 2d^*, ..$ else wait.

- For example if $d^* = 2$ then stop in odd periods if H is odd and even periods if H is even, no matter how long H is.

- And because the game is continuous at infinity, every accumulation point of finite horizon equilibria as $H \to \infty$ is an equilibrium of the infinite-horizon game.

<u>Dual-self model:</u> Solution more like a control problem than a multiplayer game.

- The temptation is to not act, save *c*.

- So with linear control cost $\gamma$ , the solution is act at once if $\delta v > (1+\gamma)c$ , else never act.

- Simple and clean, but "now or never" is a very stark conclusion; it comes from the assumption that the environment is deterministic.

FL *AER* [2006] study a stochastic version where

- Each period can act or wait.
- Waiting in period $t$ yields utility $x_t$, known at the start of the period; these are iid with known continuous cdf $P$ on $[\underline{x}, \overline{x}]$ .
- Acting ends the game, and yields flow of $v$ starting next period; present value $\delta v/(1-\delta) := \delta V$ .
- Here waiting has no self-control cost, acting costs $\gamma x_t$, the foregone utility of doing something fun.
- Assume $-\gamma \underline{x} + \delta V > \underline{x} + \delta Ex/(1-\delta)$ so when the opportunity cost is at its lowest value it is better to act than to wait forever.
- Solve by dynamic programming.

- Unique optimum has a cutoff $x^*$ : act when $x_t \leq x^*$, $\underline{x} \leq x^* < \overline{x}$ .

- $x^*$ is the level that makes the agent indifferent between acting and continuing:

$$-\gamma x^* + \delta V = x^* + \delta W^*(x^*)$$

where $W^*$ is the value of waiting and using the cutoff rule from tomorrow on.

- $dx^* / d\gamma < 0$ so the expected waiting time is increasing in the cost of self control.

- The sophisticated quasi-hyperbolic model again has multiple equilibria as the equilibrium cut-off can depend on calendar time.

- And now the non-stationary aspects of these equilibria would be observable on the path of play.

## What is a time period? What is the "temptation horizon"?

- In both the quasi-hyperbolic model and the basic dual self model, only current rewards are "tempting."

- This one-period "temptation horizon" makes unrealistic predictions about the timing of decision and payoff.

- "Period" is artificial and not observed. Choices should depend on the real time between decision and information nodes, independent of how this time is cut into periods.

- Not clear how to adapt the quasi-hyperbolic model for this, more straightforward with the dual self model.

## Fudenberg and Levine *Ema* [2012]: Non-myopic short run selves

- Discrete time periods $n = 1, 2, \ldots$ with period length $\tau$ which we think of as "short" (to proxy for continuous time).

- So want temptations "a few periods from now" to matter more and more as the periods get shorter.

- Normalize so that per-period utility *u* flow is independent of $\tau$: a constant utility of 1 corresponds to flow payoff of 1.

- SR selves: discount factor $\delta = \exp(-\rho\tau)$, survival probability $\mu = \exp(-\eta\tau)$. SR's objective function from any period on is normalized discounted value with discount factor $\delta\mu$.

*Long Run Self:* Discounts future utility by $\delta$ $(> \delta\mu)$, incurs a "control cost" to change behavior of SR self. This cost depends on how much value the SR is asked to forego.

Encode influence of past actions on continuation payoffs and choice sets in a state variable *y* (*e.g. wealth or contents of liquor cabinet.*)

*Temptation value* $\bar{U}(y_n)$:
maximum feasible average present value for SR at state $y_n$. (*assume finite)*

*Foregone value:*

$$\Delta(y_n, a_n) = \bar{U}(y_n) - \Big((1-\delta\mu)u(y_n, a_n) + \delta\mu E_{y_n, a_n} \bar{U}(y_{n+1})\Big)$$

*One interpretation:* future temptation value is SR's prediction of future utility, and SR doesn't think about the way current actions might change future self-control. (other interpretations can also fit.)
This reduces to the original model if SR only care about the current period.

Assume linear cost: LR's objective at any period *n* is

$$\sum_{\ell=0}^{\infty} \delta^{\ell} \left( (1-\delta) u_{n+\ell} - \Gamma \Delta_{n+\ell} \right).$$

Note that this functional form scales *u* but not $\Gamma$ by $(1-\delta)$ .

- w/o self control a constant flow of *u* is worth *u* for any $\tau$.
- Action that foregoes 1 util in *every* period costs $\Gamma$ for any $\tau$.
- Foregoing 1 util in 1 period costs $\Gamma(1-\delta\mu)$: $O(\tau)$.

*Important:* An infinite sequence of one-period-at a time 1-util sacrifices costs

$$\Gamma(1-\delta\mu) \,/\, (1-\delta).$$

This is more than the cost $\Gamma$ of acting now to resist the entire sequence of future temptations.

Apply to "simple versus persistent temptations."

***<u>Simple temptation</u>*****<u>:</u>** Choose accept/reject

- Reject: no change in utility flow.
- Accept: get $u_g > 0$ for $N = T / \tau$ periods, $-u_b < 0$ thereafter.
    - (so $T$ is in units of real time.)

- LR value *P*, SR value *S*.

- Assume $S > 0 > P$ : the quintessential LR/SR conflict.)

- *Solution:* Reject if $\left|P\right| > \Gamma S$ (regardless of $N$ and $\tau$).

*"The only way to get rid of a temptation is to yield to it."*
Wilde, *Portrait of Dorian Grey*

***Persistent*** temptation: If resist, face same choice next period.
(*same as regular temptation if periods very long…)*

With non-myopic SR selves, the persistent temptation more costlyt o resist than the simple one because declining a simple temptation commits to avoid <u>future</u> temptations.

- Foregone value if resist the persistent temptation is $S - \delta\mu S = (1-\delta\mu)S$.

- So resist if $\left|P\right|(1-\delta) > \Gamma(1-\delta\mu)S$.

- As $\tau \to 0$ resist simple temptation and accept persistent one if

$$(\rho + \eta)\Gamma S > \rho \left| P \right| > \rho \Gamma S$$

Here

$\rho \left| P \right|$: value of postponing P for an interval $dt$

$(\rho + \eta)\Gamma S$: flow cost of resisting the persistent temptation

$\rho \Gamma S$: flow equivalent of one-time payment of $\Gamma S$.

Same logic implies agent can decline a "bundle" of two simple temptations, one occurring tomorrow and one next week, even though he would take each of them if presented in isolation, as in Ainslie *Breakdown of Will* [2001], Kirby and Guassello *J. Exp. Psych.* [2001].

- Linear costs make things simple, but we believe costs are typically convex.
- But a conceptual problem with convex costs and short time periods.
- So the second half of the paper introduces as “stock of willpower” which can be depleted in the short run, in the spirit of Buameister et al *J. Pers. And Soc. Psych.* [1998], Baumeister and Muraven *Psych. Bulletin* [2000].
- Conceptually straightforward but leads to more complicated optimal control problems.

- Money isn't consumption, so need an auxiliary assumption (like "mental accounts") to explain why people treat at least small amounts of money like consumption.

- One attempt: FL propose that people set mental accounts when not tempted: "pocket cash."

- But many papers use choices over money payment to look at present bias and self control.

T**haler [1981], Myerson and Green** [1995], etc.: How much (hypothetical) money now to make indifferent to hypothetical $1,000 after delay $t$?

For each $t$ estimate marginal interest rate

- Geometric discounting*:* marginal interest rates time invariant.

- Standard dual-self and quasi-hyperbolic models Initial marginal interest rate high, subsequent ones equal and low.

- Data shows smooth decline, both overall and within subject. (Andersen et al [2008] find smooth but smaller decline using lotteries with real stakes.)

- Fit with non-myopic SR players.

**Keren and Roelsofsma** [1995]:

- 100 Florins now (82%) or 110 in 4 weeks.

- 100 in 26 weeks (37%) or 110 in 30 weeks.

- Many more choose 100 when "now" than when all choices pushed off half a year, which isn't consistent with exponential discounting.

- When money payments replaced with .5 chances of same payment, fractions choosing 100 were .39 and .33: almost no present bias.

- Confirmed by Weber and Chapman [2005] with real stakes.

- Explained by convex control cost + SR cares about winnings in 30 weeks: costs less than ½ as much to resist a ½ probability of reward.

**Baucells and Heukamp** [2010]:common-ratio Allais paradox with immediate payoffs, then with all payoffs to be in 3 months.

The reversal rate is 36% without delay, 22% with delay.

Qualitatively consistent with convex control costs: Allais paradox comes from convex control costs, less temptation about future payoffs.

More motivation for non-myopic SR selves…

--

- No class next Thursday

- Suggested reading (*won't be on the exam*): MWG Ch. 21, 23 A-C