Wald Revisited: The Optimal Level of Experimentation

Giuseppe Moscarini
Yale

Lones Smith
M.I.T.

Fall, 1997

(2

Bayesian Reformulation of Sequential Paradigm

- actions: A and B (eg. accept H_0 and accept H_1) → H_1 true
- $p =$ decision maker's ($\mathcal{DM}$'s) prior on state H ($1-p$ on state L) H_0 true
- $\mathcal{DM}$ is risk neutral, with payoffs / utilities: $\pi_A^L, \pi_B^L, \pi_A^H, \pi_B^H$
- $\pi_a(p) \equiv p\pi_a^H + (1-p)\pi_a^L$ is $\mathcal{DM}$'s expected payoff to action a
- suppose one might never decide $\Rightarrow$ need null (zero payoff) action
- optimal static payoff $\bar{\pi}(p) \equiv \max\langle \pi_A(p), \pi_B(p), 0\rangle$
- *R&D model*

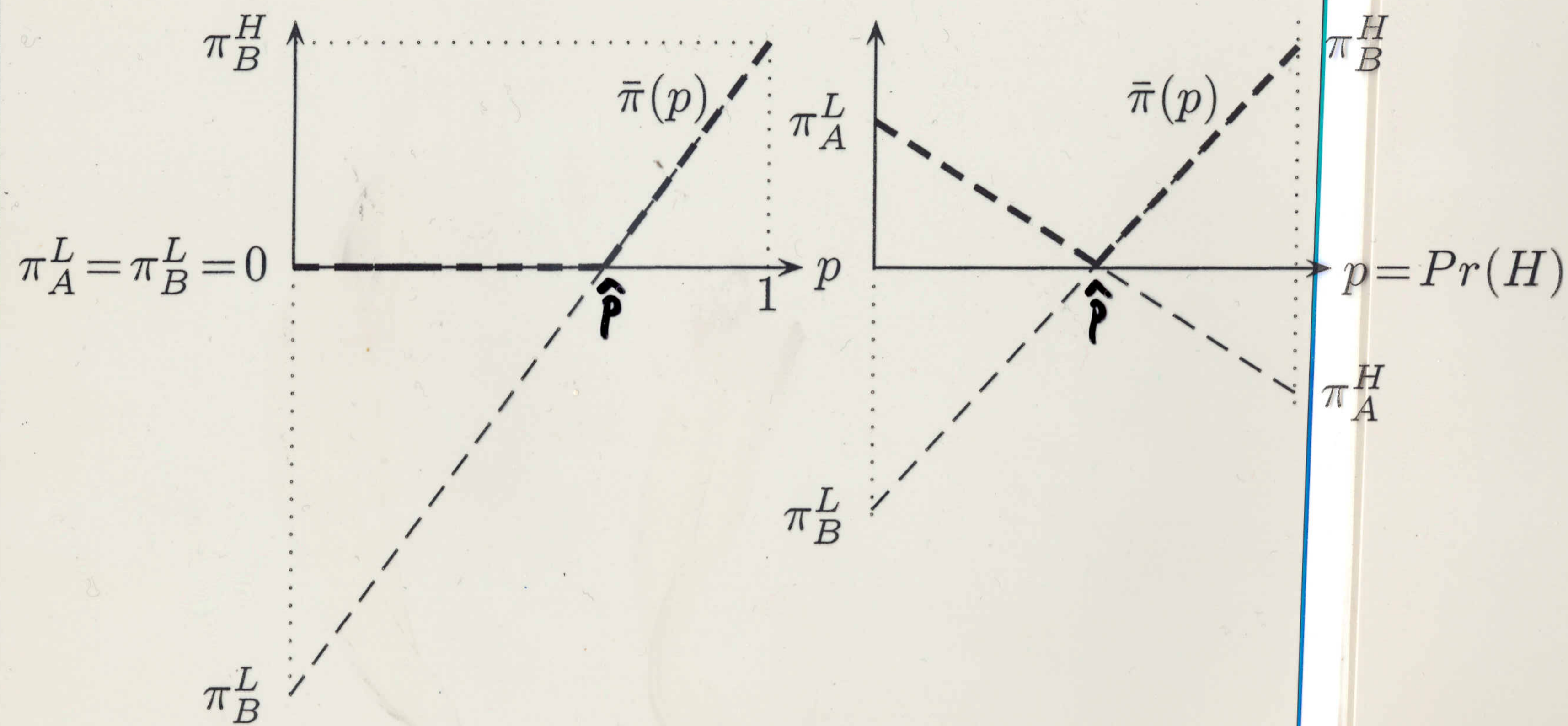
  - actions: $B =$ 'build' costly new prototype, and $A =$ 'abandon'

  - payoffs: $h = \pi_B(1) > 0$, $\ell = \pi_B(0) < 0$, $\pi_A(p) \equiv 0$

  $\Rightarrow \bar{\pi}(p) \equiv \max\langle 0, hp + \ell(1-p)\rangle$
- Wald-Wolfowitz (1949): If each signal $\cancel{x}$ costs a given amount, then the cost minimizing strategy is the sequential one:

  $\mathcal{DM}$ quits and chooses action A (B) with posterior $p \leq \underline{p}$ ($p \geq \bar{p}$)

Typical Static Payoff Frontiers



Our Economic Twist

- homo economicus
  - (a) is impatient $\Rightarrow$ encourages nonsequential parallel "stacking" of info purchases
  - (b) faces increasing, strictly convex cost of info $\Rightarrow$ encourages Wald's sequential behaviour

- Can we characterize level of experimentation?
  - eg. – binary signals $X = \begin{cases} \begin{array}{ccc} \theta=H & \theta=L & \\ 1 & -1 & 1/2+\varepsilon \\ -1 & 1 & 1/2-\varepsilon \end{array} \end{cases}$
  - buy N signals at cost $C(N)$, $C'>0$, $C''>0$
  - sufficient statistic for information = chance p of state H
  - payoff to $\langle N_0, N_1, N_2, \dots, N_T$, final belief $p_T \rangle$ is $E\left(\delta^T \pi(p_T) - \sum_{k=0}^{T} \delta^{k-1} C(N_k)\right)$
  - $V(p)$ = Bellman Value & $N(p)$ = optimal experimentation level

A Man, A Plan, A Canal

① - new model: if you can't solve the problem, change it
- introduce cont's time extrapolation of discrete time models ≡ control of variance of diffusion

② - characterization: $N(p)$ is an increasing f^n of $V(p)$ using micro IOI intuition
→ robust insight for countable state models and normal learning models

③ - testable implications → costs $\langle C(N(p_t)) \rangle$ = submartingale
→ sensitivity analysis
eg. nonstandard impatience result

④ - R&D interpretation when states are payoff-ordered
⇒ $N(p)$ increasing as a function of p

A. R&D → Kamien-Schwartz (1971-82), Grossman-Shapiro ('86)
Dutta ('97), Malueg & Tsutsui ('97) model 'D' not 'R'
→ Poisson model dominates learning-based models
⇒ beliefs don't move up and down

B. Sequential Analysis in Statistics
→ no characterizations exist except on whether sequential analysis or some form of variable-level experimentation is best OR numerical simulations
→ eg. Cressie & Morgan (1993)

Wald-Wolfowitz (1948)

Arrow-Blackwell-Girshick (1949)

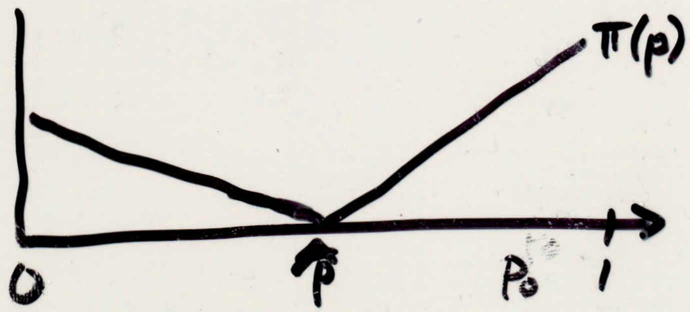
Optimal Experimentation

Dynamic Programming

C. Optimal Experimentation – mostly longrun learning studied
→ models confound information-gathering incentives with immediate payoff concerns (eg. bandits, monopoly pricing)

SCIENCE vs. ENGINEERING ⇒ DEMAND FOR INFO PAPER: PAYOFFS NOT RANDOM SIGNALS
(INFO = GOOD OF 21st CENTURY) (NONSTOCHASTIC PAYOFFS)

Static Intuitions



① → p_0 closer to cross-over point $\hat{p}$ ⇒ information more valuable at margin

(2) → p_0 closer to ½ ⇒ more "elastic" /variable posterior beliefs

So information demand is greatest in middle

Dynamic Intuition

→ time preference ⇒ information purchase should be backloaded ∴ info demand least in middle

Continuous Time Experimentation

- recall: $\tilde{X} = \pm 1$ with chances $0.5 \pm \varepsilon$ in state H & $0.5 \mp \varepsilon$ in state L
- sample average $\bar{X}_t = \sum X_i / N_t$ in period t has mean $\pm 2\varepsilon$ in states H, L and is a sufficient statistic for signals
- larger N proportionately raises its precision ($\equiv 1/\text{variance}$)

$\therefore$ choose observation process $\equiv$ diffusion $\langle S_t \rangle$:

- nature chooses drift μ_ω in state ω, $\mu_H = \mu > 0$, $\mu_L = -\mu < 0$
- DM controls flow variance $Var(dS_t) = \sigma^2 / n_t$ via the experimentation level n, AND chooses T = stopping time

- cost function $c(n)$ is increasing, strictly convex, $\underbrace{\& \; c(0) = 0}_{\text{optional}}$

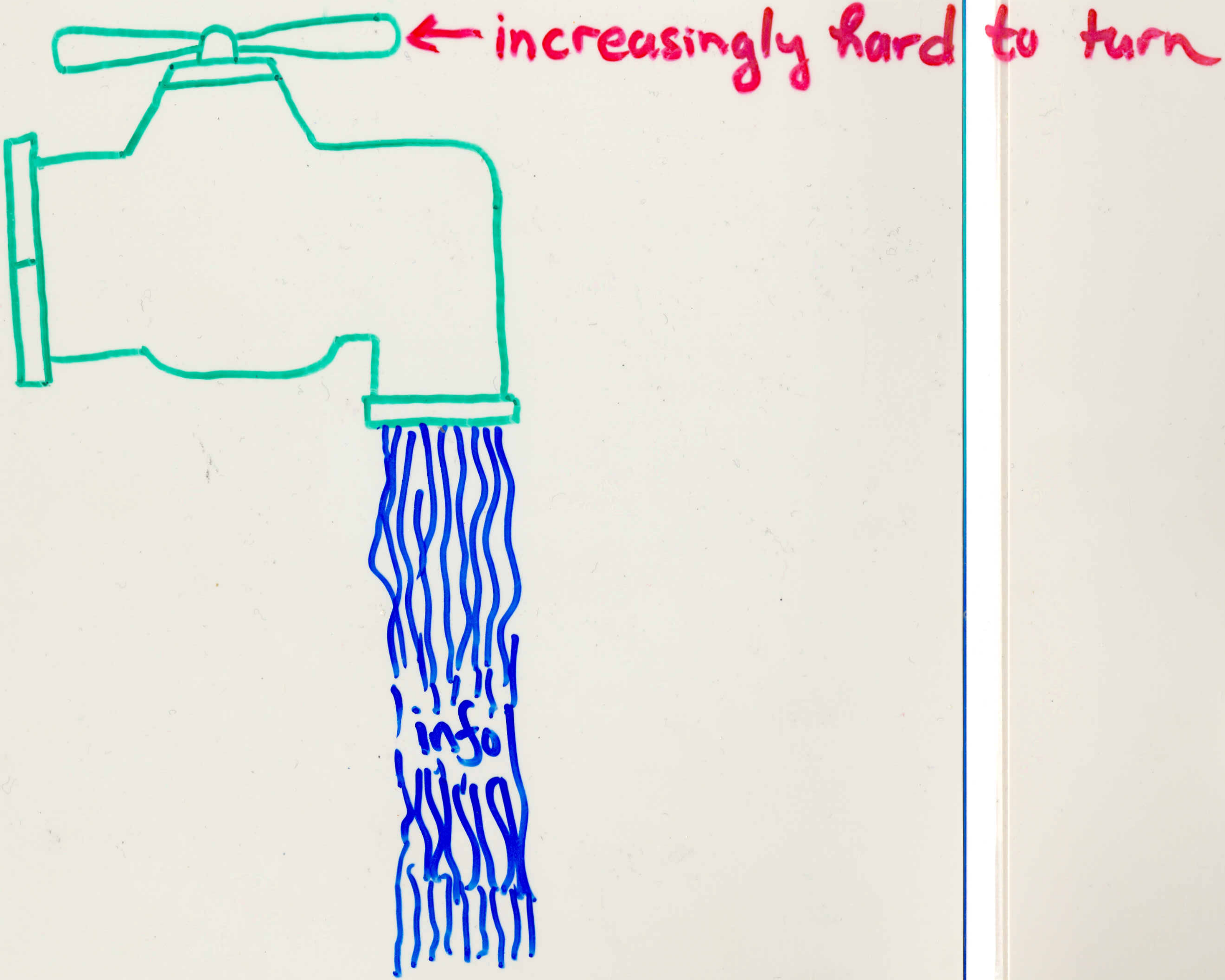
$$\boxed{dS_t = \mu_\omega dt + (\sigma / \sqrt{n_t}) dZ_t}$$

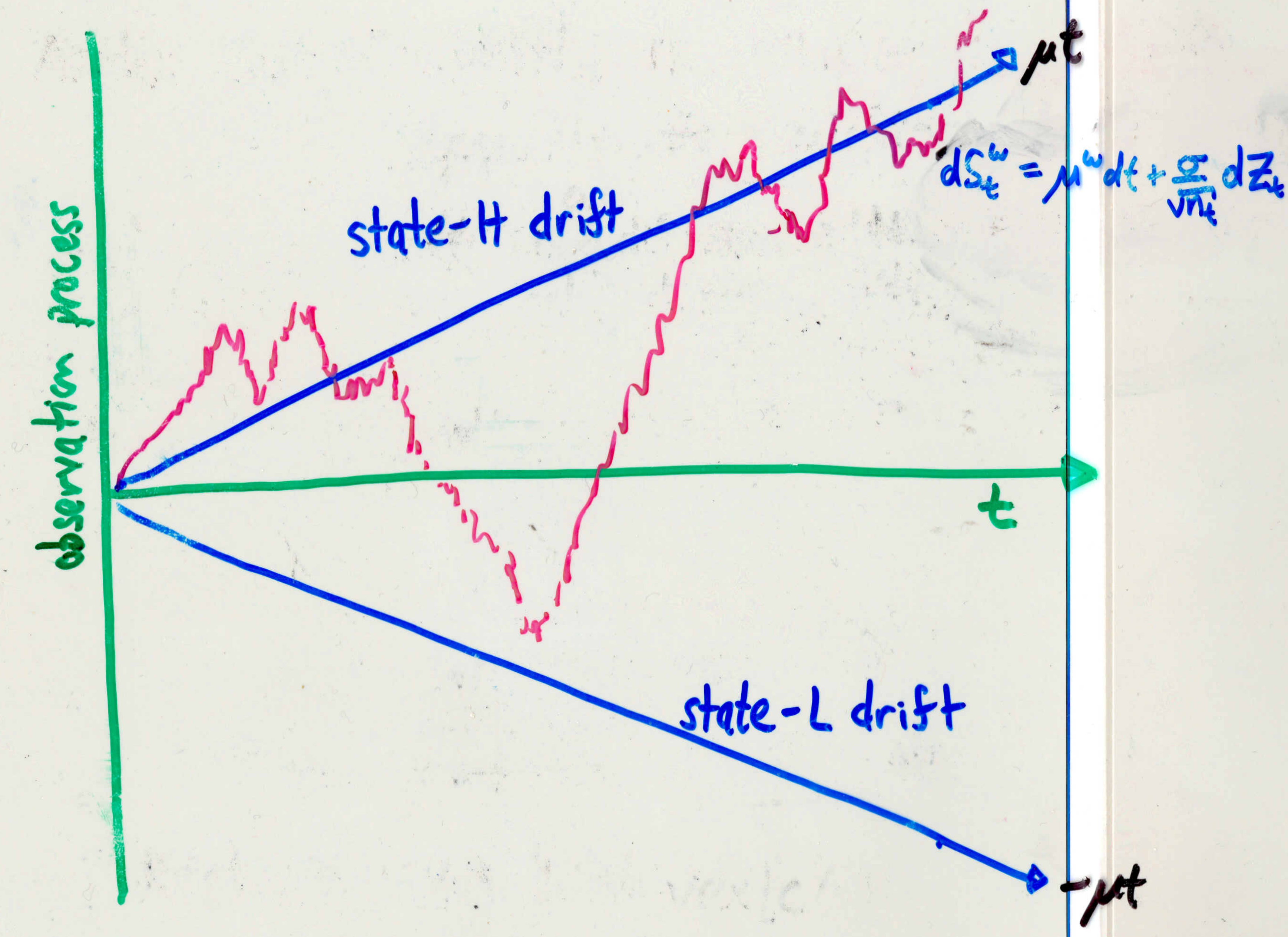
- interest rate $r > 0$ $\Rightarrow$ objective is the value function

$$v(p_0) \equiv \max_{T, \langle n_t \rangle} E\left[\int_0^T -c(n_t) e^{-rt} dt + e^{-rT} \bar{\pi}(p_T) \,\middle|\, p_0, \langle n_t \rangle\right]$$

$\rightarrow$ range of applicability of our insights: conditionally iid signal processes parameterized by first sample moments

WHAT IS THE QUESTION $\rightarrow$ ~~ARA~~ RESULT ONLY NEEDS CONDITIONALLY IID SIGNAL PROCESS CHARACTERIZED BY MEAN


increasingly hard to turn
info



observation process
state-H drift
state-L drift
μt
$-\mu t$
t
$dS_t^w = \mu^w dt + \frac{\sigma}{\sqrt{n_t}} dZ_t$

Cost Function Assumptions

WHAT IS NEEDED

→ $C(0) \geq 0$

→ C strictly increasing, weakly convex ($\Rightarrow C'$ exists a.e.)

→ surplus function $NC'(N) - C(N) > r \cdot$(max terminal payoff)

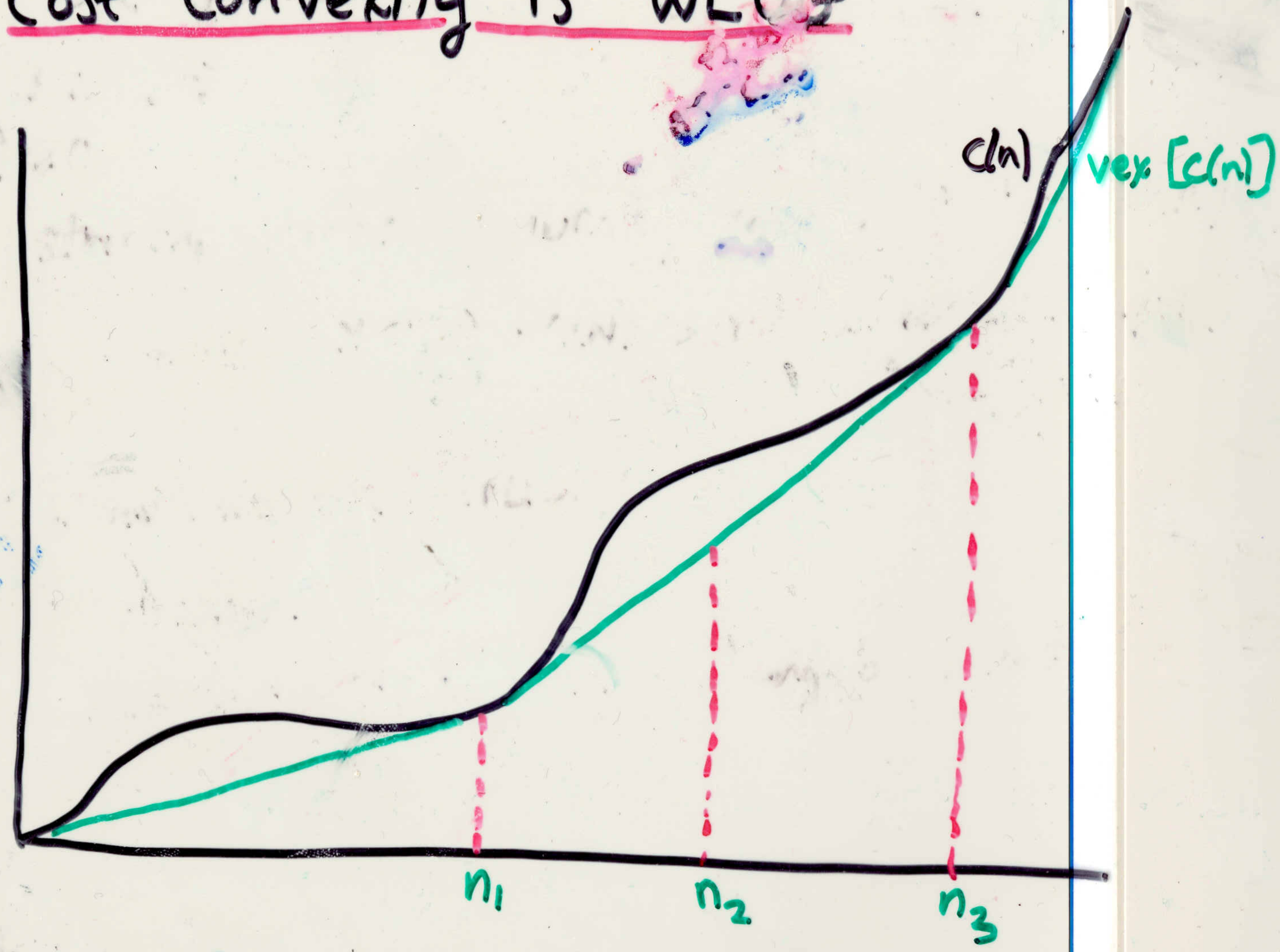
$\max\{\pi(0), \pi(1)\}$

WHAT IS ASSUMED FOR SIMPLICITY

→ C', C'' exist

→ strict convexity: $C'' > 0$ ($\Rightarrow$ no bang-bang control)

Aside: Cost Convexity is WLOG



Replace c(n) by vex[c(n)]

strict convexity & smoothness avoids corner sol'ns! ~~Smoothness allows us to~~ ISAAK NEWTON STORY JUSTIFY IT...

Recursive Formulation

- observation process $\langle S_t \rangle$ is a "diffusion" $\Rightarrow$ the belief process $\langle p_t \rangle$ is a driftless (martingale) "diffusion" of the form

$$dp_t = 0 \cdot dt + \#(p_t, n)\, dZ_t \quad \text{for some function } \#(\cdot) > 0$$

Cont's Time Bayes Rule: (Filtering Theory)

$$dp_t = p_t(1-p_t) \cdot \underbrace{\frac{[\mu - (-\mu)]}{\sigma / \sqrt{n_t}}}_{\text{signal-to-noise ratio of } \langle S_t \rangle} dZ_t = 2p_t(1-p_t)\frac{\sqrt{n_t}\,\mu}{\sigma}\, dZ_t$$

$$\therefore \text{Variance}(dp_t) \equiv \text{Var}(dp_t) = 4 n_t \frac{\mu^2}{\sigma^2} p_t^2 (1-p_t)^2 \equiv 2\Sigma(p)$$

- optimal value $v(p_0)$ equals

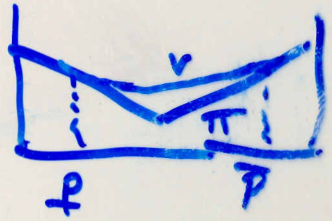
$$\max_{T, \langle n_t \rangle} E\left[\int_0^T -c(n_t)e^{-rt}dt + e^{-rT}\pi\left(\overbrace{p_0 + \int_0^T \underbrace{\sqrt{\text{Var}(dp_t)}\, dZ_t}_{dp_t}}^{p_t}\right)\right]$$

PROBLEM: MAX IS FORMULATED IN TERMS OF BELIEFS $\langle p_t \rangle$ BUT WE DON'T KNOW HOW THEY BEHAVE

$p(1-p)$ = variance of binomial coin flip

Optimality Equations

- optimal control (OC) exercise $\Rightarrow$ experimentation schedule $n(p)$
- optimal stopping (OS) exercise $\Rightarrow$ stopping boundaries $\underline{p}, \overline{p}$, and perhaps also $\underline{p}_0, \overline{p}_0$ if null action is taken
- ① for OS:



value matching: $v(\underline{p}) = \overline{\pi}(\underline{p})$ and $v(\overline{p}) = \overline{\pi}(\overline{p})$

smooth pasting: $v'(\underline{p}) = \overline{\pi}'(\underline{p})$ and $v'(\overline{p}) = \overline{\pi}'(\overline{p})$

and maybe also (if null action is optimally taken)

$$v(\underline{p}_0) = v(\overline{p}_0) = 0 \quad \text{and} \quad v'(\underline{p}_0) = v'(\overline{p}_0) = 0$$

- ② for OC, the Hamilton-Jacobi-Bellman equation for v in the interval $(\underline{p}, \overline{p})$, or $(\underline{p}, \underline{p}_0) \cup (\overline{p}_0, \overline{p})$ if null action is taken, is

FLOW COST — FLOW BENEFIT at time t of experimenting

$$rv(p) = \max_{n \geq 0} \{ \underbrace{-c(n)}_{\text{FLOW COST}} + \underbrace{\underbrace{0 \cdot v'(p)}_{E(dp_t)\, v'(p)} + \underbrace{n \Sigma(p) v''(p)}_{\frac{1}{2} Var(dp_t)\, v''(p)}}_{\text{FLOW BENEFIT}} \}$$

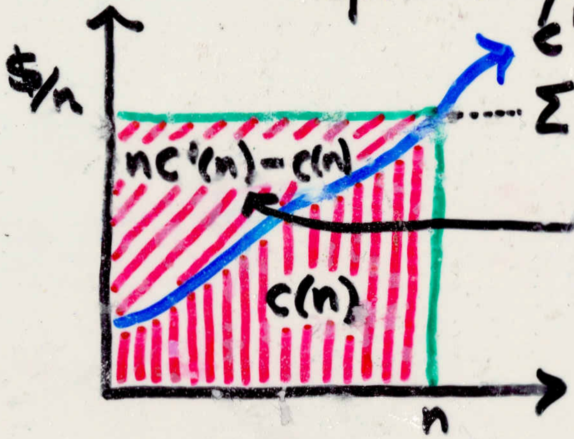
← SOC satisfied!

MORE MATH THAT I DIDN'T KNOW →
- FILTERING
- OPTIMAL STOPPING
- OPTIMAL CONTROL

Lemma (Experimentation Level Monotonicity)
If $c', c'' > 0$ then $v \mapsto n$ is increasing.

Proof: $\exists$ 2 decisions at t: stop / go & level (if go)

① level: → total benefit of experimentation $n\Sigma(p)v''(p)$ is linear
→ optimality $\Rightarrow c'(n) = \Sigma(p)v''(p)$



$\Sigma(p)v''(p)$ = marginal benefit of info

producer surplus $g(n) = nc'(n) - c(n)$

② stop / go: optimal stopping $\Rightarrow$ delay cost $rv(p) = g(n(p))$
− surplus rises in information quantity: $g'(n) = nc''(n) > 0$
$\Rightarrow rv(p) = g(n(p))$ has rising inverse $n(p) = f(rv(p))$
eg. $c(n) = n^2 \Rightarrow g(n) = n^2 \Rightarrow f(n) = \sqrt{n} \Rightarrow n = \sqrt{rV(p)}$

MAKING ECONOMIC SENSE || OF OPTIMALITY EQNS

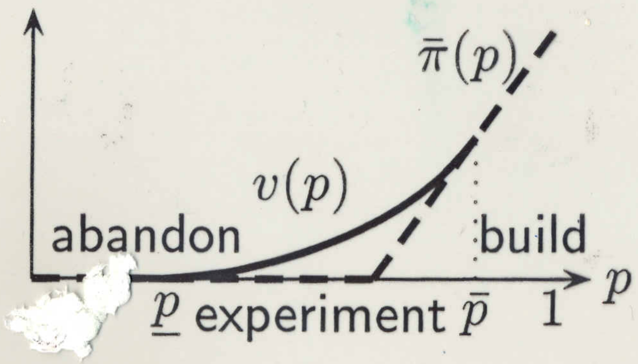
Properties of the Optimal Value Function

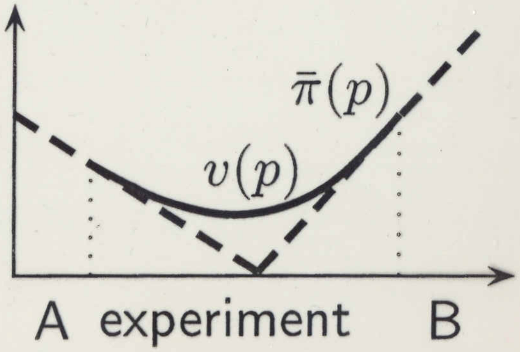
Lemma (a) *The value function v is convex.*

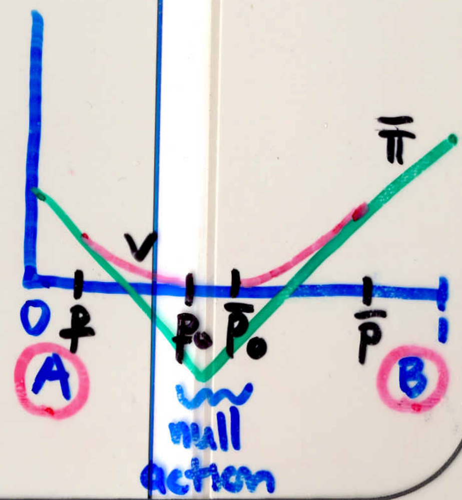
(b) *$v(p) = \bar{\pi}(p)$ for $p \leq \underline{p}$ and $p \geq \bar{p}$, for cut-offs $\underline{p}$ and $\bar{p}$.*

(c) *The static payoff $\bar{\pi}$ and value v are jointly monotone increasing (or decreasing) or U-shaped in p.*

(d) *If the null action is ever exercised, then $v(p) = \bar{\pi}(p)$ in $[\underline{p}_0, \bar{p}_0]$, where $\underline{p} < \underline{p}_0 < \bar{p}_0 < \bar{p}$.*



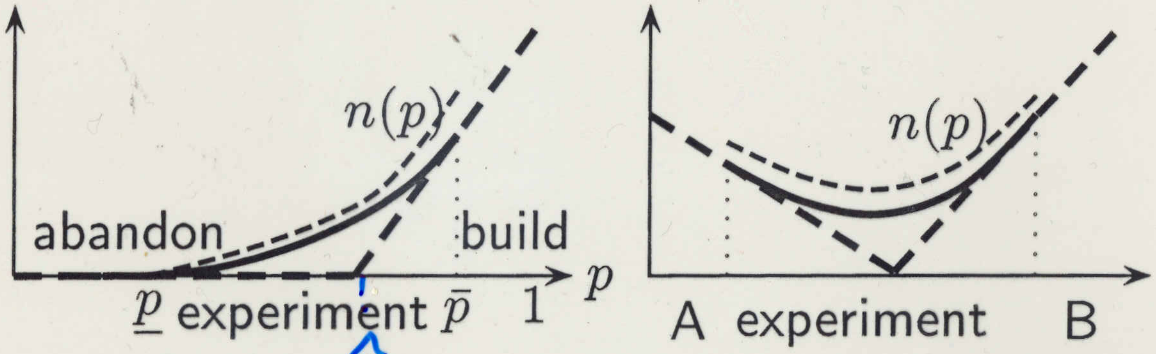




All we need do is figure out shape of value fn

The Optimal Level of Experimentation

Proposition *Assume the static payoff frontier $\bar{\pi}(p)$ is increasing (resp. decreasing, U-shaped) in p. ~~For any strictly increasing, strictly convex C^2 cost function $c(n)$,~~ T~~t~~he optimal experimentation level $n(p)$ is increasing (resp. decreasing, U-shaped) in p outside the stopping sets $[0, \underline{p}] \cup [\bar{p}, 1]$ for actions A and B.*



$\hat{p}$

- eg. research projects start small, grow larger with success
- discoveries that shift beliefs (eg. cold fusion experiment) can discretely kick up R&D levels

Convexity and Experimentation Drift

Since posterior beliefs $\langle p_t \rangle$ are a martingale & v is convex $v'' > 0$
$\Rightarrow \langle v(p_t) \rangle$ is a submartingale (drifts up)

Is n convex in p? Recall $c(n) = n^2 \Rightarrow n = \sqrt{rv}$

Information producer surplus $g(n)$ ↑ing and weakly concave $\Rightarrow$ inverse $f(n)$ is ↑ing & convex $\Rightarrow$ level $n(p) \equiv f(rv(p))$ is convex ($\because v$ convex)

Proposition If the producer surplus $g(n) = nc'(n) - c(n)$ is concave, then $\langle n(p_t) \rangle$ is a submartingale.

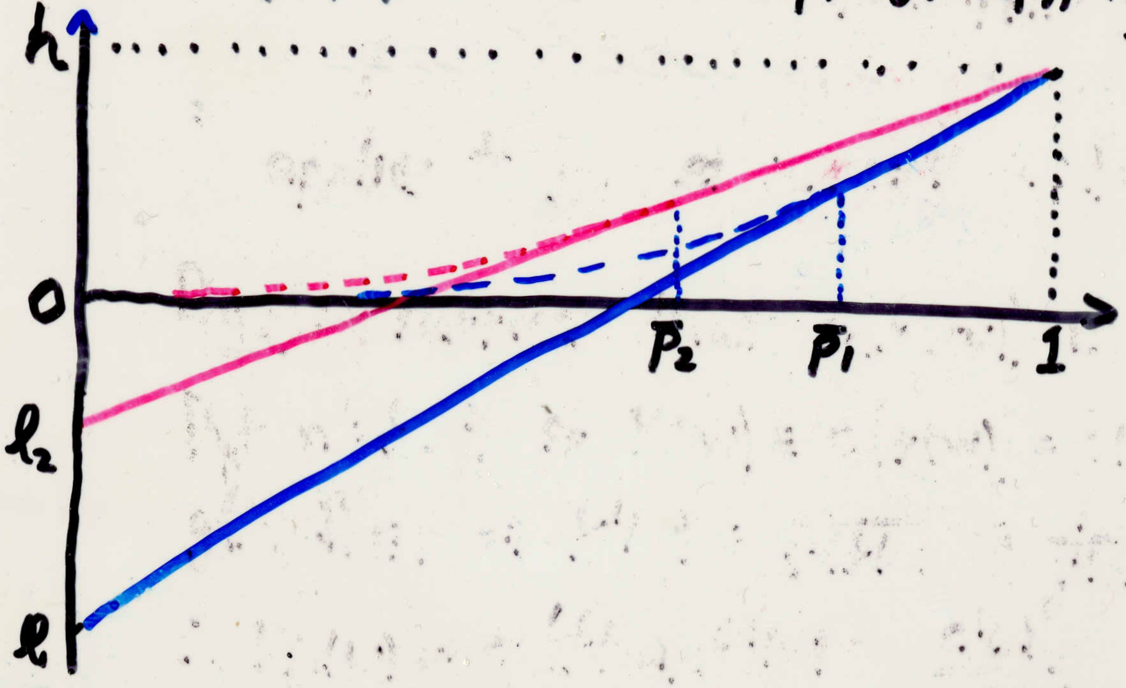
eg. $c(n) = n^k$, $k > 1 \Rightarrow g''(n) = [nc''(n)]' \geq 0$
but $c(n) = n \log n$ $(n>1)$ $\Rightarrow g''(n) = 0 \Rightarrow n''(p) > 0$

We set out to find the level. The drift is pure bonus.

Sensitivity Analysis ("Comparative Statics")

① PAYOFF SHIFTS

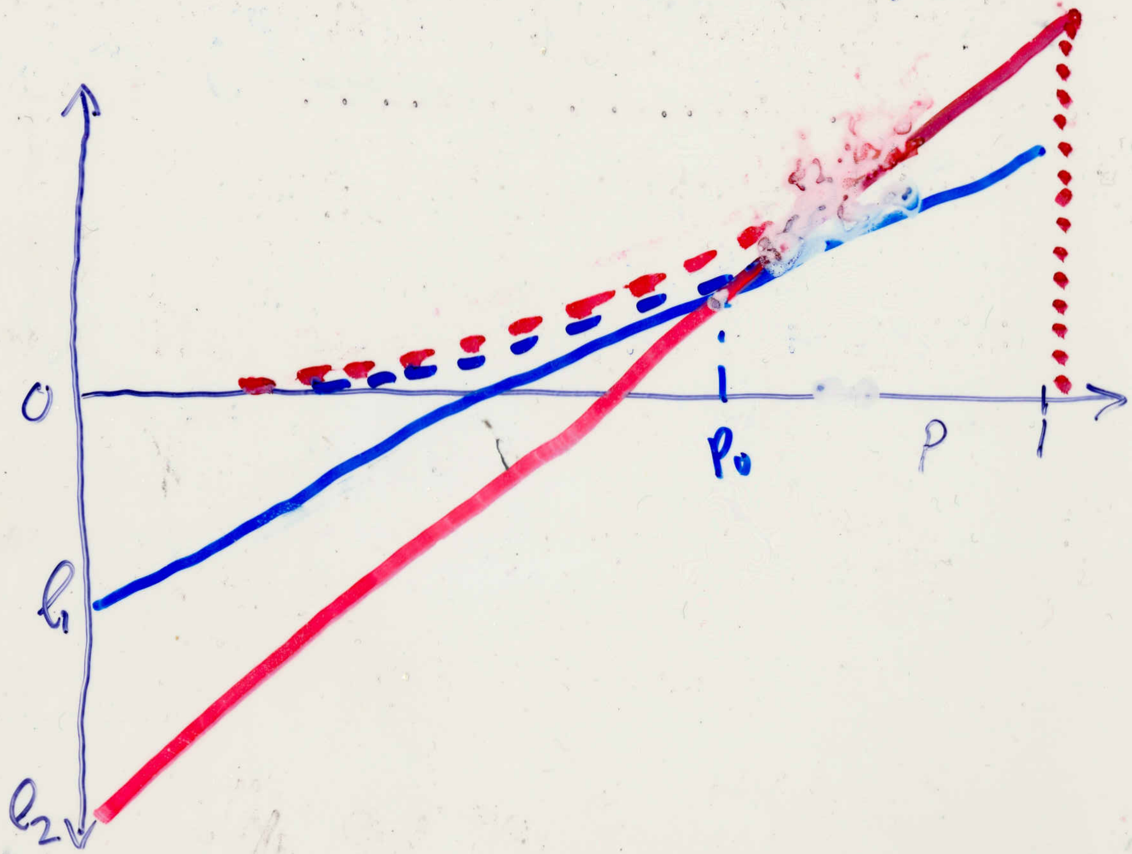
In R&D model, if ℓ rises, so does the value $v(p)$, clearly. But thresholds $\bar{p}$ & $\underline{p}$ fall! So $n(p) = f(rv(p))$ ↑ inside new thresholds.



More generally, any payoff $\pi_A^H, \pi_A^L, \pi_B^H, \pi_B^L$ ↑ leads to a higher value $v(p)$, thus a higher experimentation level $n(p)$, where $n>0$, while π_A^H ↑ ⇒ $\bar{p}, \underline{p}$ rise, while π_A^L ↑ ⇒ $\bar{p}, \underline{p}$ fall

② PAYOFF RISK

In R&D model, if $\ell\uparrow$, $h\downarrow$ so that $\pi(p)$ is unchanged, then the value $v(p)$ rises, so $n(p)$ rises, $\underline{p}\downarrow$, and $\overline{p}\uparrow$.



③ COST SHIFTS

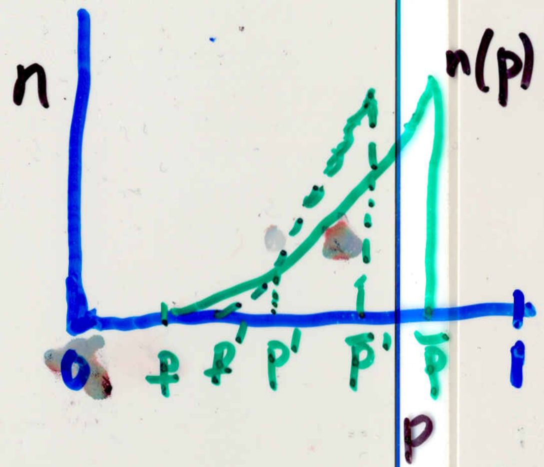
Similar comparative statics for cost of information (less convexity ⇒ higher experimentation level)

Just like search theory → tails matter // riskiness preferen

IMPATIENCE COMPARATIVE STATIC: "ANTI-FOLK" LOGIC

If the interest rate rises $r\uparrow$, ~~then~~ the value $v(p)$ falls [as usual] and thresholds shift in. While the experimentation level $n(p) = f(rv(p))$ rises near at least one threshold $\underline{p}$ or $\bar{p}$. In the R&D model, $n(p)$ declines $\forall p < p'$, rises for all $p > p'$, some $p' \in (\underline{p}, \bar{p})$

Proof: (ODE reasoning)



Wald's Limiting Sequential Paradigm

Proposition (Return to Wald's World)

(a) As the interest rate $r \to 0$, the experimentation level $n(p) \downarrow 0$ (and as $r \uparrow \infty$, $n(p) \uparrow \infty$).

(b) As the cost convexity vanishes ($\sup_n c''(n) \downarrow 0$), the experimentation level $n(p) \uparrow \infty$.

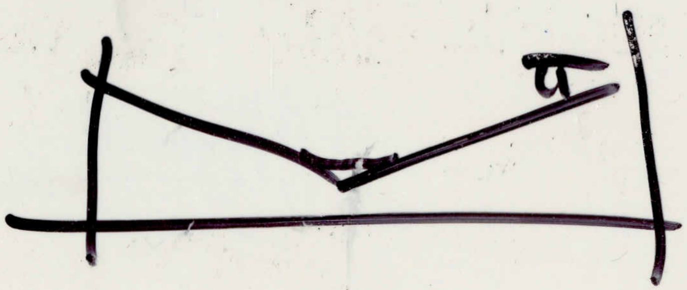
Proof of (a): $v(p) \leq \max\langle \pi_A^H, \pi_B^H, \pi_A^L, \pi_B^L \rangle < \infty$

$\Rightarrow n(p) = f(r v(p)) \downarrow 0$ as $r \downarrow 0$

$v \geq \pi > 0$ $\quad$ $rv \to \infty$ as $r \to \infty$

Proof of (b): On $[0, \bar{\pi}]$, $\max g'(n) = \max n c''(n) \downarrow 0$

So its inverse has derivative $f' \uparrow \infty \Rightarrow f \uparrow \infty$



RESTORING FOLK WISDOM ON INTEREST RATE SHIFTS

Let final payoff be annuity:

annuity final payoff

maximize $E\left(\int_0^T -c(n_t)e^{-rt}\,dt + e^{-rT}\boxed{\pi(p_T)/r} \,\middle|\, p\right) \equiv v(p)$

$\therefore$ return $w = rv$ obeys HJB equation

$$w(p) = \max_{n \geq 0}\left(-c(n) + n\,\Sigma(p)\,w''(p)/r\right)$$

$\therefore$ higher interest rate r has same effect on $w=rv$ as lower $\Sigma(p) = p^2(1-p)^2 \cdot \frac{2\mu^2}{\sigma^2}$ [i.e. lower signal-noise $\frac{2\mu}{\sigma}$] on value v

$\therefore$ <u>return</u> $w = rv$ falls

$\therefore$ $n = f(rv)$ falls with annuities

ROBUSTNESS OF MONOTONICITY & DRIFT RESULTS

K STATE MODEL OR NORMAL LEARNING MODEL

Bellman equation is still of form

$$rv(p) = \sup_{n \geq 0} \left(-c(n) + n \underbrace{\left[\frac{\text{signal}}{\text{noise}} \text{ term}\right] \cdot \left[\begin{matrix}\text{value}\\ \text{convexity}\\ \text{term}\end{matrix}\right]}_{\substack{\text{constant marginal}\\ \text{benefit of information}\\ MB}} \right)$$

belief vector or posterior mean

$$\Rightarrow c'(n) = MB$$

$$\Rightarrow rv = -c(n) + n\, c'(n)$$

$$\Rightarrow n = f(rv)$$

Likewise $\frac{n c''(n)}{c'(n)}$ non$\uparrow$ing $\Rightarrow \langle n(p_t) \rangle$ submartingale

"R&D" MODEL EXTENDS

Limiting Experimentation Models with Fine Discrete Time, Large Discrete Range

What are we approximating anyway?

Proposition Fix $\boxed{0 < \alpha < 1/2}$. Consider the ff. sequence of discrete time experimentation problems: Each period of length Δt, the DM may purchase N independent binary signals at total cost $C^{\Delta t}(N) \equiv c(N(\Delta t)^{1-2\alpha})(\Delta t)$. Each signal $X_i = \pm\sigma(\Delta t)^{\alpha}$ with chances $1/2 \pm \mu(\Delta t)^{1-\alpha}/2\sigma$ in states H, L. When DM stops with posterior p, his final payoff is $\pi(p)$. Then as $\Delta t \to 0$:

(a) The running sum $S_t^{\Delta t} \equiv \sum_{k=1}^{t} \bar{X}_k^{\Delta t}$ of per period average signals converges in distribution to the diffusion $\langle S_t \rangle$

(b) The Bellman value functions $V^{\Delta t}(p)$ and transformed optimal experimentation levels $n^{\Delta t}(p) \equiv N(\Delta t)^{1-2\alpha}$ pointwise converge to the limits $v(p)$ and $n(p)$.

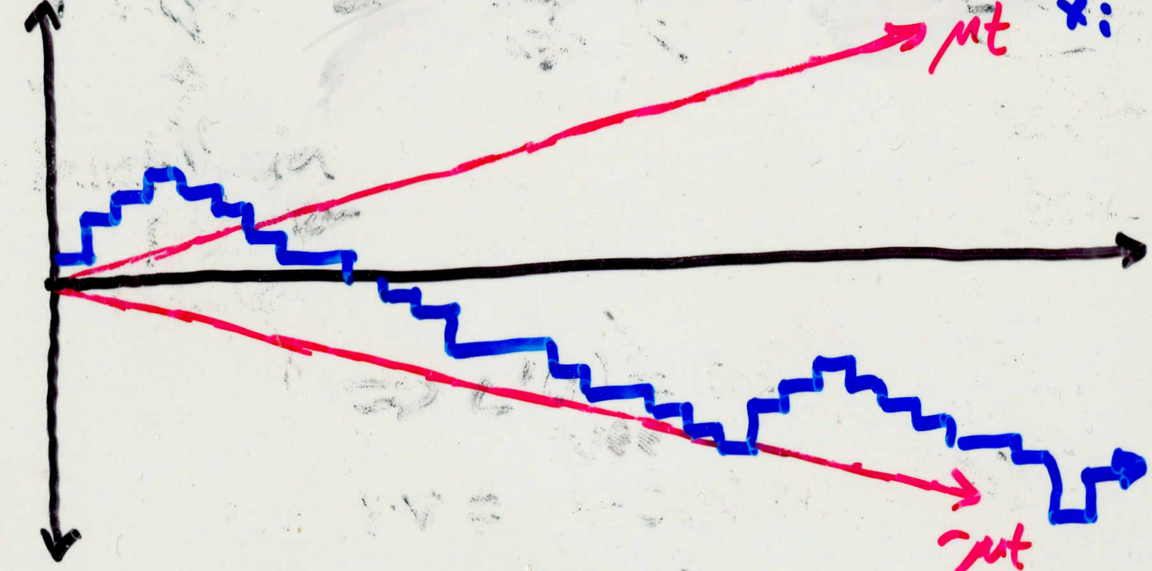
(c) The shape of $n(p)$ is inherited by $n^{\Delta t}(p)$.

CONT'S TIME ≡ RIGHT CHOICE

BUT REALLY ANY SEQUENCE OF DISCRETE TIME MODELS CAN CONVERGE TO OUR LIMIT IF 1ST TWO MOMENTS ARE OK

Quick Overview of Choice of Models

① observation process settles down in variance, and its drift in states H, L is $\pm \underbrace{\sigma(\Delta t)^{\alpha}}_{x_i} \cdot \underbrace{\mu(\Delta t)^{1-\alpha}/\sigma}_{\varepsilon} \equiv \pm\mu$



② PDV of costs converges when $C^{\Delta t}(N^{\Delta t}) = c(n^{\Delta t})\Delta t$

$$\sum_{i=1}^{T/\Delta t} e^{-i(r\Delta t)} C^{\Delta t}(N_i^{\Delta t}) \equiv \sum_{i=1}^{T/\Delta t} e^{-i(r\Delta t)} c(n_i^{\Delta t})\Delta t \to \int_0^T e^{-rt} c(n_t)\,dt$$