Controlled Learning in Continuous Time

Duarte Gonçalves

University College London

Topics in Economic Theory

Overview

1. Optimal Learning

2. Optimal Control: Necessity and Verification

3. Effort, Ability, and Decision Time

4. Final Remarks

How Much and How Quickly to Learn?

Moscarini & Smith (2001 Ecta).

DM chooses how much effort to put into learning.

Two choices:

When to stop: how much evidence is enough?

How intensively to experiment: how quickly to obtain it?

E.g., firm deploying more simultaneous tests/hiring more researchers delivers information sooner, but at higher marginal cost.

Discounting makes delay costly: there is return to reaching valuable decision sooner.

Should experimentation be most intense when DM is most uncertain?

Controlling the Speed of Learning

Keep known information quality $m > 0$ and binary state $\theta = \pm 1$.

DM now controls effort e_t at every instant:

$$dX_t = m\theta\sqrt{e_t}\,dt + dB_t, \qquad e_t \geq 0.$$

More effort increases information flow.

Effort e_t is chosen using observed information. Controlled filter:

$$dp_t = 2m\sqrt{e_t}\,p_t(1-p_t)\,d\widetilde{B}_t.$$

Filtering: $\widetilde{B}$ is innovation BM; choosing effort does not reveal unknown state.

With $\mathcal{L}_m\phi := 2m^2p^2(1-p)^2\phi''$, controlled generator is $\mathcal{L}_{\sqrt{e}m} = e\mathcal{L}_m$.

Double intensity gives double posterior variance:

same learning can occur in half the time.

Experimentation with Discounting

Scrap value $v(p) = \max_{\alpha} u(\alpha, p) \geq 0$; discount rate $r > 0$.

Flow cost of effort $c(e)$: $c, c', c'' > 0$ and $c''' \geq 0$.

Write η for inverse of $ec'(e) - c(e)$.

$$V(p) = \sup_{e,\tau} \mathbb{E}_p \left[e^{-r\tau} v(p_\tau) - \int_0^\tau e^{-rt} c(e_t)\, \mathrm{d}t \right].$$

Variational inequality: with new feature, optimise the continuation return over e.

$$\max \left\{ v - V,\ \sup_{e \geq 0} \{(e\mathcal{L}_m - r)V - c(e)\} \right\} = 0.$$

Convex value and piecewise linear scrap value again deliver two stopping thresholds.

DM chooses intensity (in continuation region), stopping (the continuation), and eventual decision (in stopping region).

Effort Rises with the Value of the Decision

Proposition (Moscarini & Smith (2001 Ecta), Proposition 1)

In continuation region, optimal intensity satisfies

$$e^*(p) = \eta(rV(p)).$$

It is strictly increasing in $V(p)$.

Proof Intuition

HJB and interior first-order condition give

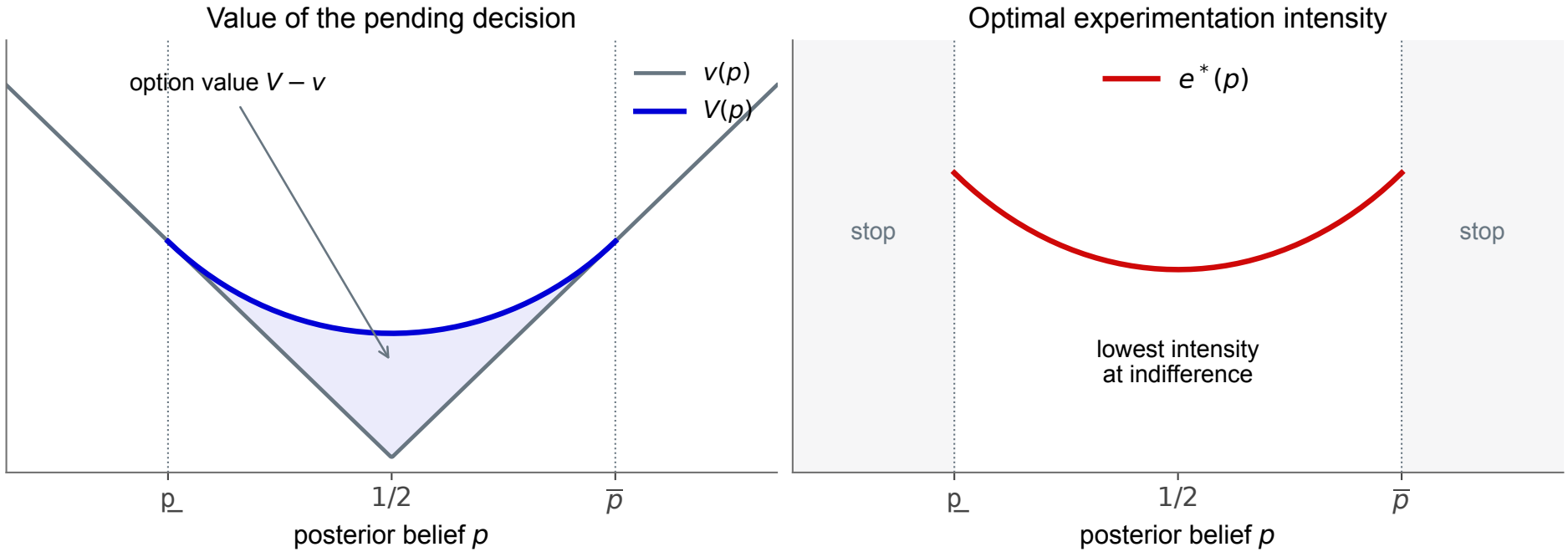
$$e^* \mathcal{L}_m V - c(e^*) = rV, \qquad \mathcal{L}_m V = c'(e^*).$$

Substitute 2nd equation into 1st. $e^*(p)c'(e^*(p)) - c(e^*(p)) = rV(p) \implies e^*(p) = \eta(rV(p))$.

LHS increases in effort: its derivative is $ec''(e) > 0$.

Information gains net of flow cost must compensate for discounting pending decision.

Uncertainty and Urgency Are Different



Effort not driven by option value but by cost in delaying.

Symmetric v: option value $V - v$ greatest at indifference, but effort is lowest there.

In contrast, near valuable decision, delay is more expensive: DM exerts more effort to get information faster.

In general, if v is increasing, then V is increasing and so is e.

Effort responds to realised beliefs; it need not increase along every sample path. Learning ends at either threshold and effort is (discontinuously) switched off.

What Changes without Discounting?

Set $r = 0$, retaining positive cost of time $c(0) > 0$.

Same arguments show

$$e^* c'(e^*) = c(e^*), \qquad e^* \in \arg\min_{e>0} \frac{c(e)}{e}.$$

Choose cheapest speed per unit of information.

Beliefs determine when to stop but don't change cheapest way to acquire information.

Next: why must an optimum satisfy the variational inequality (necessity),
and when does solution actually produce an optimum (sufficiency)?

Optimal Learning Intensity

Remark (Moscarini & Smith (2001 Ecta))

The following are equivalent:

(1) e is increasing/decreasing/U-shaped.

(2) V is increasing/decreasing/U-shaped.

(3) v is increasing/decreasing/U-shaped.

(Note: we're not even trying to solve V, just looking at its properties via HJB/BC.)

Important for U-shaped: minima are not the same!

Comparative Statics in Learning Intensity

Proposition (Moscarini & Smith (2001 Ecta), Proposition 5)

(a) Supppose $u(\alpha, \theta)$ increases. Then $\uparrow V$ and $\uparrow e$.

Moreover both $\underline{p}, \overline{p}$ decrease (increase) if $\alpha = 1$ $(= -1)$.

Interpretation: higher payoffs $\rightarrow$ cannot do any worse $\rightarrow$ higher value $\rightarrow$ higher effort

Increasing payoff to $\alpha \rightarrow$ need to be less convinced to stop and choose it (and more convinced to stop and choose alternative).

Comparative Statics in Learning Intensity

Proposition (Moscarini & Smith (2001 Ecta), Proposition 5)

(a) Supppose $u(\alpha, \theta)$ increases. Then $\uparrow V$ and $\uparrow e$.

Moreover both $\underline{p}, \overline{p}$ decrease (increase) if $\alpha = 1$ $(= -1)$.

(b) Let $\hat{u}$: $\hat{u}(\alpha, p_0) = u(\alpha, p_0)$ but $|\hat{u}(\alpha, a) - \hat{u}(\alpha, b)| > |u(\alpha, a) - u(\alpha, b)|$.

Then $\hat{C} \supset C$, $\hat{V} \geq V$ and $\hat{e} \geq e$ (strictly on $\hat{C}$).

Interpretation: $\hat{u}$ has greater payoff risk than u

Comparative Statics in Learning Intensity

Proposition (Moscarini & Smith (2001 Ecta), Proposition 5)

(a) Supppose $u(\alpha, \theta)$ increases. Then $\uparrow V$ and $\uparrow e$.

Moreover both $\underline{p}, \overline{p}$ decrease (increase) if $\alpha = 1$ $(= -1)$.

(b) Let $\hat{u} : \hat{u}(\alpha, p_0) = u(\alpha, p_0)$ but $|\hat{u}(\alpha, a) - \hat{u}(\alpha, b)| > |u(\alpha, a) - u(\alpha, b)|$.

Then $\hat{C} \supset C$, $\hat{V} \geq V$ and $\hat{e} \geq e$ (strictly on $\hat{C}$).

(c) If cost becomes less convex, lower, and flatter, then $\uparrow V$ and $\uparrow e$, and C expands.

(d) If signal-to-noise ratio increases, then $\uparrow V$ and $\uparrow e$, and C expands.

(e) If r rises, then $\downarrow V$, C shrinks; effort strictly increases near $\underline{p}$ or $\overline{p}$.

(att: typos in statement in main text — check proofs)

Extensions and Variations

No Discounting, Linear Costs

As $r \to 0$, effort becomes constant in $p \in C$

As cost becomes linear, effort explodes and $C \to [0, 1]$

Beyond Binary Actions/States

Optimal effort still $e(p) = \eta(rV(p))$

If $p >_{MLR} q \implies v(p) > v(q)$, then $e(p)$ monotone wrt $\geq_{MLR}$.

Next: How to solve for optimal control in general?

Overview

1. Optimal Learning

2. Optimal Control: Necessity and Verification

3. Effort, Ability, and Decision Time

4. Final Remarks

Controlling an Itô Process

Itô Process: state X_s, $\mathrm{d}X_s = \mu(s, X_s, \alpha_s)\,\mathrm{d}s + \sigma(s, X_s, \alpha_s)\,\mathrm{d}B_s, \quad X_t = x.$

Payoffs: Flow payoff $f(t, x, a)$; Scrap value $v(t, x)$; Discount rate $r \geq 0$.

Control α_t: $\alpha_t \in A$, progressively measurable, $\mathbb{E}\int_t^{t+h} a_s^2\,\mathrm{d}s < \infty\ \forall t + h < T$ (can be relaxed).

Controlled Generator: For smooth function f, define spatial generator

$$\mathcal{L}_t^a f(t, x) = \mu(t, x, a) \cdot \partial_x f(t, x) + \frac{1}{2}\sigma(t, x, a)^2 \partial_{xx} f(t, x).$$

Control changes drift and variance; changes the generator.

Time and discounting: Parabolic discounted and controlled generator: $\partial_t + \mathcal{L}_t^a - r$.

Control Problem: For horizon $T \in \mathbb{R}_{++} \cup \{\infty\}$, define

$$V(t, x) = \sup_{\alpha,\, t \leq \tau \leq T} \mathbb{E}_{t,x}\left[\int_t^\tau e^{-r(s-t)} f(s, X_s, \alpha_s)\,\mathrm{d}s + e^{-r(\tau - t)} v(\tau, X_\tau)\right].$$

Stopping immediately gives $V \geq v$; terminal condition gives $V(T, x) = v(T, x)$. Set $v(\infty, x) = 0$.

Continuation region: $C := \{(t, x) \mid V > v\}$.

Pure control: fix $\tau = T$ and optimise only a. Pure stopping: $|A| = 1$.

Dynamic Programming with a Control

Proposition (Dynamic Programming Principle)

Under sufficient regularity, for bounded stopping time $t \leq \eta \leq T$,

$$V(t,x) = \sup_{\alpha,\tau} \mathbb{E}_{t,x}\left[\int_t^{\tau\wedge\eta} e^{-r(s-t)} f(s, X_s, a_s)\, ds + \mathbf{1}_{\{\tau<\eta\}} e^{-r(\tau-t)} v(\tau, X_\tau) + \mathbf{1}_{\{\tau\geq\eta\}} e^{-r(\eta-t)} V(\eta, X_\eta)\right].$$

Proof Intuition

Any plan's continuation payoff is bounded by V at state reached at η.

Conversely, concatenate today's plan with measurably selected, nearly optimal continuation at that state.

Sufficient conditions for DPP: compact A, coefficients spacially Lipschitz uniformly in time, v bounded, f bounded or with finite expectation.

References: Øksendal (2003, Ch. 11), Fleming and Soner (2006, Ch. IV.7, IV.5, V); Pham (2009, Ch. 3.3-3.5, 5.2). Bouchard and Touzi (2011, SIAM JControlOptim).

Variational Inequality: Necessity

Theorem (Necessary Conditions)

Suppose DPP holds and $V \in \mathcal{C}^{1,2}(C) \cap \mathcal{C}^0(\bar{C})$. Under local regularity permitting DPP,

$$\sup_{a \in A} \left[(\partial_t + \mathcal{L}_t^a - r)V + f \right] = 0, \quad \text{on } C; \qquad v - V = 0, \quad \text{on } \partial C.$$

$$\text{Specifically, } \max \left\{ v - V, \sup_{a \in A} \left[(\partial_t + \mathcal{L}_t^a - r)V + f \right] \right\} = 0, \text{ on } t < T \qquad V(T, x) = v(T, x).$$

Proof Intuition

No profitable local deviation: Use any control briefly, then act optimally. Instantaneous gain cannot exceed zero.

Optimal continuation: Within C, best instantaneous gain must equal zero. If strictly negative locally, uniformly over controls, continuing would contradict DPP.

Stopping: $V \geq v$ everywhere; optimal stopping requires $V = v$.

Variational inequality holds classically where V sufficiently differentiable

(Otherwise interpreted in viscosity sense. V need not be globally $\mathcal{C}^{1,2}$).

Variational Inequality: Necessity

Theorem (Necessary Conditions)

Suppose DPP holds and $V \in \mathcal{C}^{1,2}(C) \cap \mathcal{C}^0(\bar{C})$. Under local regularity permitting DPP,

$$\sup_{a \in A} \left[(\partial_t + \mathcal{L}_t^a - r)V + f \right] = 0, \quad \text{on } C; \qquad v - V = 0, \quad \text{on } \partial C.$$

$$\text{Specifically, } \max \left\{ v - V, \sup_{a \in A} \left[(\partial_t + \mathcal{L}_t^a - r)V + f \right] \right\} = 0, \text{ on } t < T \qquad V(T, x) = v(T, x).$$

Sufficient (not necessary) conditions for $V \in \mathcal{C}^{1,2}(C) \cap \mathcal{C}^0(\bar{C})$:

regularity: A compact; v bounded, uniformly continuous; μ, σ, f bounded, jointly continuous, smooth in (t, x) with derivatives bounded uniformly in a; μ, σ globally Lipschitz in x, uniformly in (t, a); non-degeneracy: $\sigma \geq \varepsilon > 0$;

bounded value: V bounded; given other conditions, immediate if $T < \infty$ or $r > 0$.

Conditions give smoothness *on* C. Continuity gives value matching.

In controlled learning, allowing zero effort can violate non-degeneracy; requires separate argument, e.g., restrict domain.

Pure control has HJB equation; option to stop adds $V \geq v$ and variational inequality.

Verification: Regularity and Terminal Limits

Suppose W satisfies necessary conditions. Is $W = V$?

Let $W \in \mathcal{C}^0$. For classical argument, assume $W \in \mathcal{C}^{1,2}$ for $t < T$. If $T < \infty$, require

$$W(T, x) = v(T, x).$$

Integrability: for every admissible pair (α, τ), assume

$$\mathbb{E}_{t,x}\left[\sup_{\substack{t \le s \le \tau \\ s < \infty}} e^{-r(s-t)}|W(s, X_s)| + \int_t^{\tau} e^{-r(s-t)}|f(s, X_s, \alpha_s)|\, ds\right] < \infty.$$

Provides integrable bound.

Infinite horizon: if $T = \infty$, also require, for every admissible pair,

$$\lim_{u \to \infty} \mathbb{E}_{t,x}\left[\mathbf{1}_{\{\tau > u\}} e^{-r(u-t)}|W(u, X_u)|\right] = 0.$$

Holds if W is bounded and $r > 0$. Under integrability condition, also holds if $\tau < \infty$ a.s.

Sufficient conditions, can be weakened. Global $\mathcal{C}^{1,2}$ can also be relaxed.

Variational Inequality: Verification

Theorem (Sufficient Conditions, I: Upper Bound)

Under the preceding regularity, integrability, and terminal conditions, suppose

$$W \geq v, \qquad W_t + \mathcal{L}_t^a W + f(t,x,a) - rW \leq 0 \quad \text{for every } a \in A,\ t < T.$$

Then $W \geq V$.

Theorem (Sufficient Conditions, II: Attainment)

Suppose additionally that a measurable feedback $a^*(t,x)$ satisfies

$$W_t + \mathcal{L}_t^{a^*(t,x)} W + f(t,x,a^*(t,x)) - rW = 0 \quad \text{on } C^W := \{(t,x) : t < T,\ W(t,x) > v(t,x)\}.$$

Let X^* be the state under $\alpha_s^* = a^*(s, X_s^*)$, and define

$$\tau^* = \inf\{s \in [t,T) : W(s, X_s^*) = v(s, X_s^*)\} \wedge T, \qquad \inf \varnothing := \infty.$$

If (α^*, τ^*) is admissible, then $W = V$ and (α^*, τ^*) is optimal.

Inequalities establish an upper bound; proposed policy attains it.

Pure control: omit $W \geq v$, impose equality in HJB along proposed control, set $\tau^* = T$.

Why Verification Works

Proof Intuition

For control and stopping rule, set $M_s^W = e^{-r(s\wedge\tau - t)} W(s\wedge\tau, X_{s\wedge\tau}) + \int_t^{s\wedge\tau} e^{-r(u-t)} f(u, X_u, \alpha_u)\, du$.

Itô's lemma. Before stopping,

$$dM_s^W = e^{-r(s-t)} \left[(\partial_s + \mathcal{L}_s^{\alpha_s} - r) W + f(s, X_s, \alpha_s) \right](s, X_s)\, ds + e^{-r(s-t)} W_x(s, X_s)\sigma(s, X_s, \alpha_s)\, dB_s.$$

Drift of M^W is nonpositive $\implies$ supermartingale.

Take expectations: being supermartingale implies

$$W(t,x) \geq \mathbb{E}_{t,x}\left[\int_t^{t'\wedge\tau} e^{-r(s-t)} f(s, X_s, \alpha_s)\, ds + e^{-r(t'\wedge\tau - t)} W(t'\wedge\tau, X_{t'\wedge\tau}) \right].$$

Pass to terminal limit: Recall $W \geq v$ and let $t' \uparrow T$:

$$W(t,x) \geq \mathbb{E}_{t,x}\left[\int_t^{\tau} e^{-r(s-t)} f(s, X_s, \alpha_s)\, ds + \mathbf{1}_{\{\tau<\infty\}} e^{-r(\tau - t)} v(\tau, X_\tau) \right].$$

Take the supremum: $W \geq V$.

Why Verification Works

Proof Intuition

For control and stopping rule, set $M_s^W = e^{-r(s\wedge\tau - t)} W(s\wedge\tau, X_{s\wedge\tau}) + \int_t^{s\wedge\tau} e^{-r(u-t)} f(u, X_u, \alpha_u)\, du$.

Under proposed policy: $(\partial W_s + \mathcal{L}_s^{a^*(s,X_s)} - r)W(s, X_s) + f(s, X_s, a^*(s, X_s)) = 0 \implies$ drift zero before τ^*. Same argument yields a martingale. At stopping, $W = v$.
Inequalities become equalities:

$$W(t,x) = \mathbb{E}_{t,x}\left[\int_t^{\tau^*} e^{-r(s-t)} f(s, X_s, \alpha_s)\, ds + 1_{\{\tau^* < \infty\}} e^{-r(\tau^* - t)} v(\tau^*, X_{\tau^*})\right],$$

policy feasible and attains upper bound $W \geq V$, hence $W = V$.

Overview

1. Optimal Learning

2. Optimal Control: Necessity and Verification

3. Effort, Ability, and Decision Time

4. Final Remarks

Back to SAC: Endogenous Effort and Ability

Known information quality $m > 0$, ability $\lambda > 0$, and no discounting.

Effort scales speed of same posterior process:

$$\mathrm{d}p_t = 2m\sqrt{e_t}\, p_t(1 - p_t)\, \mathrm{d}W_t.$$

Higher ability lowers the cost of a given intensity:

$$V(p) = \sup_{e,\tau} \mathbb{E}_p \left[v(p_\tau) - \int_0^\tau c(e_t/\lambda)\, \mathrm{d}t \right].$$

Take $c(0) > 0$, $c' > 0$, $c'' > 0$, and $ec'(e) - c(e) \to \infty$.

Admissible plans have finite expected total cost.

Ability affects behaviour through chosen effort; it does not change primitive m.

Optimality Conditions

Variational inequality:

$$\max\left\{v - V,\ \sup_{e\geq 0}\{e\mathcal{L}_m V - c(e/\lambda)\}\right\} = 0.$$

In continuation region, HJB and first-order condition are

$$e^*\mathcal{L}_m V = c(e^*/\lambda), \qquad \text{value of continuing,}$$

$$\mathcal{L}_m V = \frac{1}{\lambda}c'(e^*/\lambda), \qquad \text{choice of intensity.}$$

First equates total information benefit to total flow cost.

Second equates marginal information benefit to marginal effort cost.

Eliminate $\mathcal{L}_m V$: same intensity solves both equations at every continuation belief.

Constant Optimal Effort

Proposition (SAC: Effort Choice)

There is unique $\bar{e} > 0$ satisfying

$$\bar{e}c'(\bar{e}) = c(\bar{e}).$$

Optimal policy uses $e^*_\lambda = \lambda\bar{e}$ until stopping, with constant flow cost $c(\bar{e})$.

Proof

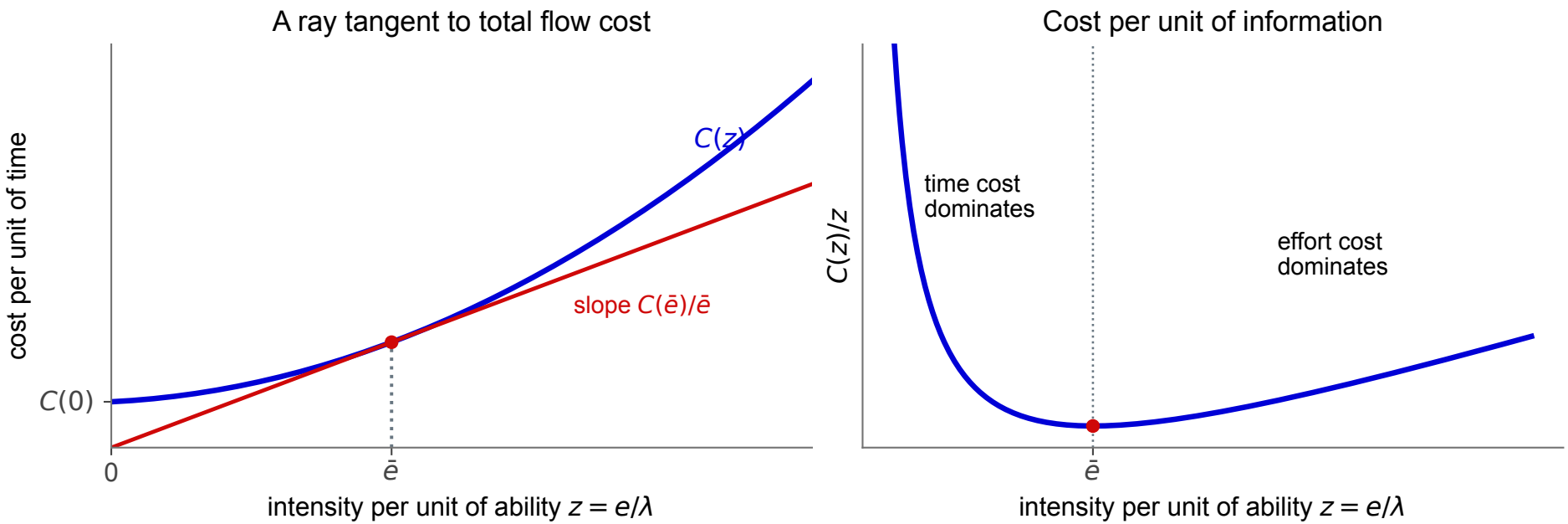
$-c(0) < 0$ and $ec'(e) - c(e)$ strictly increasing in $e > 0$, divergent.

Equivalently, $\bar{e}$ uniquely minimises $c(e)/e$:

$$\frac{\mathrm{d}}{\mathrm{d}e}\frac{c(e)}{e} = \frac{ec'(e) - c(e)}{e^2}.$$

$\implies$ $\lambda\bar{e}$ minimises $c(e/\lambda)/e$: cost per unit of information.

Constant Optimal Effort



Slow learning wastes cost of time; very fast learning incurs high marginal effort cost.

At optimum, average cost per unit of information equals marginal cost.

Information Clock and Separating Problems

Intuition: Time Change

Let $s(t) := \int_0^t e_u \, du$. In information time, posterior follows

$$dP_s = 2mP_s(1 - P_s)\, d\widetilde{W}_s, \qquad p_t = P_{s(t)}.$$

Set $K_\lambda := c(\bar{e})/(\lambda\bar{e})$. For every intensity,

$$c(e/\lambda) \geq K_\lambda e.$$

$\Longrightarrow$ any plan ending at information time $S = s(\tau)$ pays total cost at least $K_\lambda S$.

Constant effort $e = \lambda\bar{e}$ attains bound and implements any stopping rule in information time.

How quickly to buy information separates from how much to buy.

Zero-effort intervals add cost without information; omitting them improves undiscounted plan.

Recast as Pure Stopping Problem

At optimal effort,

$$\mathrm{d}p_t = 2m\sqrt{\lambda \bar{e}}\, p_t(1 - p_t)\, \mathrm{d}W_t, \qquad \text{flow cost} = c(\bar{e}).$$

Replace earlier primitives with

$$m_\lambda := m\sqrt{\lambda \bar{e}}, \qquad \bar{c} := c(\bar{e}).$$

Characterisations and comparative statics apply.

Higher ability raises effective information quality whilst keeping ‘cost of time’ unaffected.

Continuation region expands: the DM chooses to become better informed.

At fixed ability, higher m changes boundaries, but leaves optimal effort unchanged.

Optimal Stopping

Theorem (Gonçalves (WP) SAC, Theorem 2, Proposition 9)

Let $\theta = \pm 1$, $m_\lambda \geq 0$, and $X_t := m_\lambda \theta t + B_t$, where B is a Brownian motion. Let B, θ be independent, with $p_0 = \mathbb{P}(\theta = 1) \in (0,1)$, $p_t := \mathbb{P}(\theta = 1 \mid \mathcal{F}_t^X)$. Suppose $\alpha = \pm 1$, $v(p) := \max_\alpha u(\alpha, p)$, and let τ be earliest optimal stopping time solving $\sup_\tau \mathbb{E}_p[v(p_\tau) - \bar{c}\tau]$. Then, (1) $\mathbb{E}[\tau]$ is non-monotone and quasiconcave in m_λ, and (2) $\mathbb{P}(\alpha_\tau = \theta)$ increases in m_λ when $p_0 = 1/2$ or $V(1/2) > v(1/2)$.

Proof Intuition

$\mathbb{E}[\tau]$: scale functions + implicit function theorem.

$\mathbb{P}(\alpha_\tau = \theta)$: boundary expansion + $p_0 = 1/2 \implies$ Blackwell-more information.

Under optimal stopping, stopping time becomes non-monotone.

Conclusion generalises to uncertain drift $M \sim \mu$: under conditions identified earlier for μ, comparative statics under prior $\mu \downarrow 0$ and $\mu \uparrow \infty$.

Ability: More Accurate, Not Always Faster

Theorem (Gonçalves, SAC, Theorem 3)

Under optimal effort and stopping:

(i) Accuracy weakly increases in ability λ, strictly while learning occurs.

(ii) Expected stopping time is non-monotone and quasiconcave in ability λ.

(iii) Expected stopping time is quasi-supermodular (ability, complexity) (λ, m^{-1}) on $\tau > 0$.

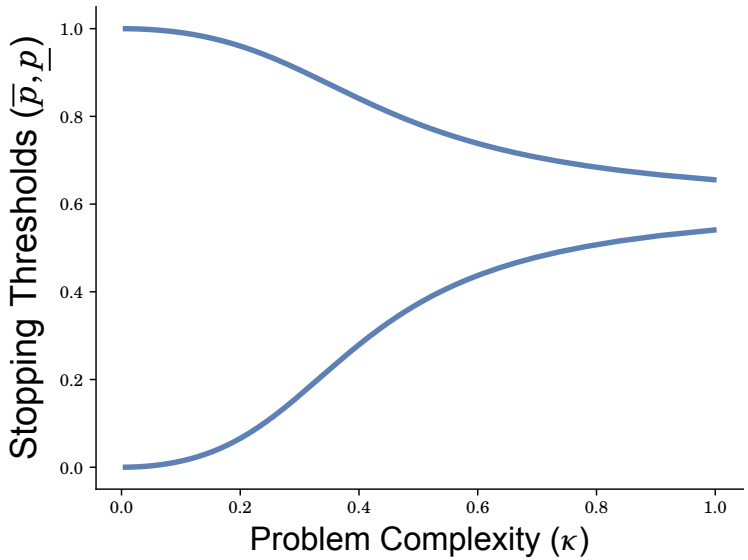
Intuition: Comparative statics carry over to ability. Single-crossing property follows immediately from $\mathbb{E}[\tau]$ being quasiconcave in m_λ and $\partial_m m_\lambda, \partial_\lambda m_\lambda > 0$ (Lemma).

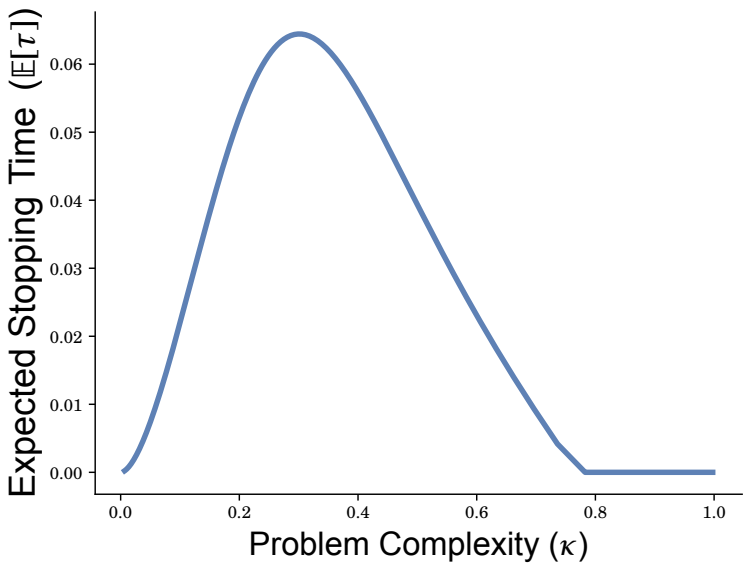
Higher ability $\equiv$ Simpler problem. Higher accuracy, but single-crossing:

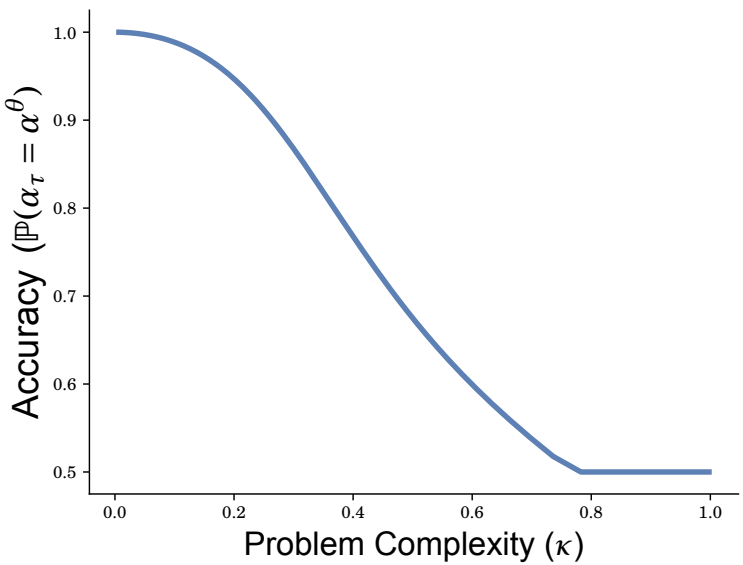
in simple problems, DM with higher ability chooses faster and better;

in complex problems, DM with higher ability chooses slower and better.

Speed-Accuracy Non-Monotonicity under Optimal Stopping







$\kappa = m^{-1}$.

Ability: More Accurate, Not Always Faster

Theorem (Gonçalves, SAC, Theorem 3)

Under optimal effort and stopping:

(i) Accuracy weakly increases in ability λ, strictly while learning occurs.

(ii) Expected stopping time is non-monotone and quasiconcave in ability λ.

(iii) Expected stopping time is quasi-submodular (ability λ, complexity m^{-1}) on $\tau > 0$.

Intuition: Comparative statics carry over to ability. Single-crossing property follows immediately from $\mathbb{E}[\tau]$ being quasiconcave in m_λ and $\partial_m m_\lambda, \partial_\lambda m_\lambda > 0$ (Lemma).

Higher ability $\equiv$ Simpler problem. Higher accuracy, but single-crossing:

in simple problems, DM with higher ability chooses faster and better;

in complex problems, DM with higher ability chooses slower and better.

Reconciling Empirical Evidence:

Inverted-*U* shaped speed-IQ (Lindley et al. 95).

Faster responses more indicative of higher ability in easier tasks (Dodonova Dodonov 13; Goldhammer et al. 15).

Overview

1. Optimal Learning

2. Optimal Control: Necessity and Verification

3. Effort, Ability, and Decision Time

4. Final Remarks

What Now?

What can you do with this?

Big hammer: several properties and tools at significant level of generality.

Unclear if further work deriving general properties payoffs off.

Apply them! Lots of gaps in literature waiting for applications.

Importantly: make time the central piece.

Inference, screening, dynamic frictions, growth, experience, search, repeated interaction.

If time is there for technicality, there's no point: can probably get similar results with static model.

Controlled Learning in Continuous Time

Duarte Gonçalves

University College London

Topics in Economic Theory

References

Pham (2009), *Continuous-time Stochastic Control and Optimization with Financial Applications*. Ch. 3, 5.

Øksendal (2003), *Stochastic Differential Equations*, Ch. 10-11.

Fleming and Soner (2006), *Controlled Markov Processes and Viscosity Solutions*. Ch. IV-V.

Liptser and Shiryaev (2001), *Statistics of Random Processes*. Ch. 8-10.

Peskir and Shiryaev (2006), *Optimal Stopping and Free-Boundary Problems*.

Moscarini & Smith (2001 Ecta), The Optimal Level of Experimentation.

Gonçalves, Speed, Accuracy, and Complexity.