No. 2026-010

The Economics of Assisted Transition:
Learning Dividends, Re-closure Risk, and Optimal Withdrawal

September 4, 2026

Takanori Suzuki

The Economics of Assisted Transition:

Learning Dividends, Re-closure Risk, and Optimal Withdrawal

Takanori Suzuki[1]

September 4, 2026

Research Institute of Business Administration (Sanken), Waseda University
Working Paper Series, No. 2026-010

Working paper note. This paper is the fifth in a research program on mental quieting, self-referential psychological memory, entry into attention, experience-dependent learning, and assisted transition. The preceding papers are RIBA Working Paper No. 2026-006 (*Two Dynamics toward Mental Quieting: A Comparative Model of Faith-Mediated and Self-Observation Pathways*), No. 2026-007 (*When Memory Governs Memory: An Endogenous-Operator Model of Recursive Psychological Activity*), No. 2026-008 (*The Dynamics of Entry into Attention: A Point Process of Candidate Openings toward Functional Non-Closure*), and No. 2026-009 (*When Openings Change Future Openings: A Fast–Slow Model of Experience-Dependent Transition Laws in Self-Referential Cognition*). The immediate predecessor develops the experience-dependent transition technology taken as given here. The present paper adds an individual first-best value problem and derives the signed Learning Dividend, post-entry consolidation, support tapering, and optimal withdrawal. This public working-paper version is based on the canonical full manuscript, Draft Version 0.4.1. A journal-submission version is in preparation. The paper consolidates the main text and technical appendices in a single document. Reproduction code, machine-readable numerical outputs, complete finite-grid audit files, and vector figures accompany the paper.

Abstract

Many forms of assistance are designed to end, yet a first qualified transition does not reveal when support should be withdrawn. This paper studies an individual first-best problem in which current assistance changes the transition law governing later entry, fusion, re-closure, and unsupported persistence. An aligned planner observes (J, ℓ), with $J \in \{R, A\}$, chooses support intensity, and may irreversibly withdraw after entry. A joint kernel carries the next coarse state and operator legibility; a companion fast–slow model generates that kernel in the numerical analysis. The Bellman operator is a contraction, and values are monotone under explicit value-relevant order conditions. A policy-fixed signed Learning Dividend compares the same contingent policy under endogenous and clamped legibility, allowing a dividend, neutrality, or penalty. A one-cycle benchmark identifies when future transition-law value creates positive post-entry consolidation despite variable cost and fixed overhead. An upper-set stopping condition yields an economic withdrawal threshold ℓ_W that generally differs from the structural persistence threshold $\ell^\dagger$. Decreasing differences provide a continuum route to a state taper, while pairwise action single crossing provides the weaker route verified by the finite numerical

[1]Faculty of Commerce, Waseda University. Email: suzuki@waseda.jp. ORCID: 0009-0008-8955-7003.

action set. Together with endpoint conditions, these results generate Starter–Consolidation–Taper–Withdrawal regions. Micro-founded stress tests produce positive and negative learning, all three orders of ℓ_W relative to $\ell^\dagger$, finite taper and withdrawal, and an information-created trial under type uncertainty. For management accounting, support should be evaluated by transition-law value, timing losses, and Cost to Durable Independence rather than first entry or service volume. The analysis is theoretical and individual-level; efficacy, provider incentives, population allocation, and solvency are outside its scope.

Keywords: assisted transition; learning dividend; optimal stopping; consolidation; support tapering; re-closure; durable independence; management accounting.
JEL classification: C61; D83; D91; M41.

1 Introduction

Many assistance relationships are valuable precisely because they are not meant to become permanent. A tutor, facilitator, clinician, adviser, or digital support system may create the conditions for an initially fragile transition, while the underlying mission is fulfilled only when the relevant function can be sustained without the active relation. Dynamic treatment regimes and just-in-time adaptive interventions formalize support as a contingent sequence of decisions rather than a fixed dose, and the scaffolding literature emphasizes contingency, fading, and transfer of responsibility (Murphy 2003; Nahum-Shani et al. 2018; Wood et al. 1976; van de Pol et al. 2010). These traditions make one point especially important here: the first observed success is not, by itself, a stopping rule. A qualified entry into functional non-closure can occur while the transition remains fragile, so re-closure may follow and later episodes may still depend heavily on external support. Immediate withdrawal can therefore sacrifice value that one more period of consolidation would preserve or create. Indefinite retention is not the answer, however. Once active support ceases to improve the individual transition, it continues to consume resources, preserves an evaluative frame, and may turn a temporary aid into a continuing dependency. The economic problem is consequently not only whether support can produce an entry. It is when to start, whether to retain support after entry, how to reduce its intensity as the individual state changes, and when to withdraw the active relation altogether.

That problem is dynamic because present assistance can change the technology of later episodes. Learning-by-doing and skill-formation models treat current activity as an input into future productive capacity (Arrow 1962; Ben-Porath 1967; Cunha and Heckman 2007). The relevant capacity here is not accumulated attention, a stock of compliant behavior, or a stronger observing agent. It is *operator legibility* ℓ: the ease with which a recurrent self-referential operator structure becomes recognizable when it reappears. Qualified experience can raise ℓ and thereby improve later entry, reduce competing fusion, delay re-closure, and enlarge the set of post-withdrawal states from which positive persistence is feasible. Assistance also has an adverse channel. Resource use and current psychological burden are direct costs, while evaluative capture can make observation into an achievement, ownership, or performance project and lower future legibility. Motivational-crowding and hidden-cost-of-control research likewise shows why an intervention cannot be ranked only by the amount supplied (Frey and Jegen 2001; Falk and Kosfeld 2006; Patzak and Zhang 2025). More support can create contact and simultaneously contaminate the learning process. The return on assistance is therefore

signed: experience may improve the future transition law, leave it unchanged, or make it worse. A one-shot evaluation that records current entry and current cost but freezes the future transition law can consequently err in either direction.

The neighboring literatures already contain nearly every generic ingredient of this argument. Adaptive-intervention research supplies state-contingent decision rules; scaffolding research supplies fading; learning models supply investment in future capability; optimal-stopping and real-option models supply the continue-versus-stop margin; Bayesian experimentation supplies the value of informative trials; and management accounting supplies the warning that incomplete measures can redirect behavior (McDonald and Siegel 1986; Blackwell 1953; Baker 1992; Feltham and Xie 1994; Ebrahim and Rangan 2014). The gap claimed here is therefore deliberately narrow. A companion transition paper develops the specific fast–slow technology used below: higher operator legibility orders attention entry, competing fusion, and matched-state re-closure, and above a structural threshold $\ell^\dagger$ a positive post-withdrawal persistence basin becomes available; the same architecture permits flat and reverse learning (Suzuki 2026). That paper intentionally stops before assigning values, costs, or an optimal intervention. For this transition technology, the open question is how current service benefit, future transition-law value, re-closure risk, resource cost, evaluative capture, and withdrawal should be combined in one individual decision problem. The present paper supplies that bridge. It does not re-prove the psychology of entry, and it does not jump ahead to provider incentives or population solvency.

The baseline model is an infinite sequence of discrete episodes with state $x = (J, \ell)$. The coarse state $J \in \{R, A\}$ distinguishes an unresolved or re-closed condition R from a post-entry active condition A; fusion remains an adverse within-episode event rather than a third persistent economic state. The planner observes (J, ℓ), chooses a scalar support intensity u, and in state A may instead select irreversible withdrawal. Withdrawal yields an autonomous continuation value $W(\ell)$ and is distinct from choosing $u = 0$ while keeping the active relation, because continued enrollment preserves future control options and carries a positive overhead. A joint Markov kernel $\mathsf{K}_J(d\ell', J' \mid \ell, u)$ carries both next-period legibility and the next coarse state. First-entry, fusion, and re-closure probabilities are current summaries generated from the same episode law, not substitutes for terminal-state probabilities. Rewards are stated in normalized functional units: explicit current benefits are reduced by fusion and re-closure damage, resource cost, and evaluative burden. No intrinsic price is assigned to attention or to the label A. The complete-information, benevolent-planner formulation is a first-best benchmark; provider rent, strategic retention, population allocation, capacity, liabilities, reserves, and solvency are excluded.

The analytical results follow the dependency structure of that problem. A one-cycle benchmark first identifies when current durability plus future transition-law value covers the variable and fixed costs of one more post-entry block. The full Bellman operator is then shown to be a contraction, yielding a unique value pair and a stationary optimal support-and-stopping policy. The policy-fixed Learning Dividend compares the same contingent policy under an endogenous legibility update and a clamped-legibility counterfactual. An exact performance-difference identity expresses this signed value as the discounted occupation of a one-step transition-law wedge, while a formation–natural-loss–capture ledger separates its channels. The stopping result is deliberately stated at the weakest economically relevant level: an upper stopping set and opposite endpoint signs yield ℓ_W; global strict decrease is a clean sufficient benchmark, whereas the

numerical model satisfies a one-crossing, eventual-decline condition rather than global decline. The sign of the continue–withdraw gap at $\ell^\dagger$ determines whether economic withdrawal occurs before, at, or after structural persistence becomes feasible. For the intensive margin, decreasing differences provide a continuum monotone-taper theorem, and pairwise action single crossing provides the finite-action theorem matched by the computation. Additional endpoint conditions generate a taper-onset boundary ℓ_T and four policy regions. A chronological Starter–Consolidation–Taper–Withdrawal path follows only along histories with entry, no intervening re-closure, and improving legibility; phase revisits remain possible. In the baseline numerical stress test, $\ell^\dagger = 0.2376$, the first observed taper action occurs at the grid point $\widehat{\ell}_T = 0.26$ with bracket $[0.24, 0.26]$, and withdrawal begins at $\widehat{\ell}_W = 0.64$ with sign bracket $[0.62, 0.64]$. The otherwise identical clamped-legibility economy withdraws at 0.24 yet, from the same initial state, never progresses to a finite exit, so a static transition law simultaneously understates the value of support and misprices its timing. From $(R, 0.06)$, the optimized signed learning value is 8.838 normalized units. These are internal realizability diagnostics, not estimates or intervention recommendations.

The paper makes four contributions at that bounded level. First, it converts an experience-dependent entry/fusion/re-closure technology into an individual dynamic value and stopping problem without collapsing the stochastic transition kernel to a mean update. Second, it introduces a policy-disciplined signed Learning Dividend and separates three boundaries that answer different questions: $\ell^\dagger$ for structural persistence feasibility, ℓ_T for the intensive-margin onset of tapering, and ℓ_W for the extensive-margin termination of active support. Third, it derives rather than assumes the four-region policy and records the exact logical boundary between the continuum sufficient theorem and the weaker finite-action condition verified numerically. Fourth, it translates the dynamic results into individual management-accounting objects. One-shot evaluation error is the negative of the policy-fixed Learning Dividend; premature withdrawal, over-retention, and nonoptimal intensity generate separate losses; Cost to Durable Independence uses durable withdrawal rather than service volume as its denominator; and the type-uncertainty extension separates the physical value of a trial from its information premium. These contributions do not amount to an asset-recognition claim, an efficacy claim, or a solution to provider agency. Section 2 positions the mechanism. Section 3 fixes the economic and transition interface. Section 4 develops the analytical results. Section 5 constructs the micro-founded numerical implementation and audits its assumptions. Section 6 develops the economic and management-accounting implications, and Section 7 concludes.

One scope statement belongs at the outset. The analytical results apply to any assisted-transition system satisfying the reduced kernel, payoff, and stopping conditions of Sections 3 and 4. The numerical model supplies one mechanism-based realization of those conditions, drawn from the self-referential cognition environment of the companion transition paper. It reproduces the finite-grid conditions and policy shapes reported in Section 5 and demonstrates internal realizability of the headline economic mechanisms; it does not verify every continuum sufficient condition, and it does not validate the framework for the educational, clinical, advisory, or digital-support settings that motivate this introduction, each of which would require its own identification of the transition law before any threshold derived here could be used.

2 Related Work and Positioning

2.1 Positioning by mechanism rather than application label

This paper sits at the intersection of five literatures that ask related but nonidentical questions. Dynamic intervention research studies how treatment or support should respond to an evolving state. Scaffolding research studies contingent assistance, fading, and transfer of responsibility. Learning-by-doing and human-capital models study present activity as an input into future productive capacity. Experimentation and real-option models study the value of preserving future choice while information arrives. Motivational-crowding and autonomy-support research studies how an intervention can help through one channel while becoming controlling or need-thwarting through another. Management accounting studies how measures, incentives, and denominators redirect organizational attention and behavior.

Each tradition supplies an essential ingredient. None of the paper's building blocks should therefore be presented as an isolated invention. State-contingent intervention is established; gradual withdrawal is established; investment in future capability is established; the value of information is established; adverse motivational effects of control are established; and distortion from incomplete performance measures is established. The contribution claimed here is the construction and solution of a particular joint object: an individual first-best problem in which a qualified opening changes an experience-dependent transition technology for later entry, fusion, and re-closure, while active support simultaneously carries resource cost and evaluative-capture cost. That object yields a signed transition-law value, a post-entry consolidation decision, an intensive-margin taper boundary, an extensive-margin withdrawal boundary, and an individual-side accounting system for durable independence.

Positioning Principle 2.1 (Mechanism before domain). The paper is not positioned by claiming that no prior field studies "support," "withdrawal," or "learning." It is positioned by identifying which mechanism each literature already supplies, which additional mechanism is required by the transition technology inherited from the companion paper, and which result follows only after those mechanisms are placed in the same dynamic value problem.

The central positioning statement is accordingly narrow:

> Existing work provides mature theories of adaptive intervention, fading, learning investment, experimentation, motivational crowding, and performance measurement. This paper combines those ingredients around a micro-founded transition state, operator legibility ℓ, whose evolution changes later entry, fusion, re-closure, and support dependence. The resulting contribution is not a new general theory of any one ingredient; it is the derivation of the signed Learning Dividend, positive consolidation, monotone taper, and economic withdrawal from one internally consistent individual transition problem.

This formulation also preserves the division of labor across the three-paper sequence. The upstream learning paper supplies the transition technology and the structural persistence boundary $\ell^{\dagger}$. The present paper values that technology and derives ℓ_T and ℓ_W. The downstream population paper will add provider incentives, heterogeneous aggregation, capacity, liabilities, reserves, and solvency. Related work is used here to sharpen that middle contribution, not to erase the boundaries among the three papers.

2.2 Dynamic interventions, optimal stopping, and the fading of support

The dynamic-treatment-regime literature formalizes an intervention as a sequence of decision rules that maps evolving information into stage-specific actions. In the canonical formulation, current state variables summarize the history relevant to the next treatment decision, and an optimal regime maximizes a cumulative outcome over the sequence (Murphy 2003). Just-in-time adaptive intervention research develops a closely related architecture for ongoing support, separating decision points, tailoring variables, intervention options, and proximal outcomes (Nahum-Shani et al. 2018). These frameworks establish that support need not be a fixed dose or a one-time assignment; it can be a contingent rule indexed by a changing individual state.

The present paper belongs to that broad family but asks a narrower economic question. Its state is not an unrestricted vector of clinical or behavioral covariates. It is the reduced pair (J, ℓ), where $J \in \{R, A\}$ records whether a qualified entry is currently operative and ℓ records the legibility of the recurrent operator structure. The action is a scalar support intensity, and withdrawal is an irreversible stopping action in state A. More importantly, current support does not only affect a proximal outcome; through qualified experience, natural loss, and evaluative capture, it changes the law governing later episodes. The intervention is therefore both a current input and an investment in—or damage to—the future transition technology.

The educational-scaffolding literature supplies the nearest substantive language for assistance that should eventually recede. The original tutoring formulation emphasizes contingent support that enables performance beyond the learner's unaided capacity (Wood et al. 1976). Subsequent reviews identify contingency, fading, and transfer of responsibility as central characteristics rather than optional embellishments (van de Pol et al. 2010). Worked-example research formalizes staged transitions from high guidance to independent problem solving (Renkl and Atkinson 2003), and later work shows that the fading of teacher and material scaffolds must be coordinated rather than treated as a mechanically uniform withdrawal schedule (Martin et al. 2019). Recent synthesis further indicates that structure and autonomy support can be complementary rather than simple opposites (Patzak and Zhang 2025).

This literature makes a broad priority claim about tapering untenable. The paper does not discover that support can be faded. Its contribution is to make fading an endogenous economic policy shape. The state-taper theorem follows from decreasing differences in the solved continuation return: as ℓ rises, the incremental value of a higher support intensity falls. A distinct taper-onset threshold ℓ_T is then derived when the optimum leaves a positive consolidation plateau. The result is more limited than a universal calendar schedule. It orders support across states; it becomes a chronological taper only on paths along which ℓ does not fall. Re-closure or reverse learning may return the process to a lower state and rationally re-escalate support.

Optimal-stopping theory supplies the comparison between a continuation value and an outside or terminal value. Search models generate reservation rules by comparing another period of search with stopping (McCall 1970); real-option models characterize the value of waiting when action is irreversible and information or project value evolves (McDonald and Siegel 1986). The present stopping problem has the same formal backbone but a different economic content. Continuing support changes the state law that determines its own future necessity, and the stopping payoff $W(\ell)$ is an autonomous regime rather than a sale price or accepted wage. The economic

withdrawal threshold ℓ_W is therefore derived from the optimized active-support branch and the autonomous value. It is not imposed as a maximum duration, and it need not equal the structural threshold $\ell^\dagger$ at which positive persistence geometry first becomes available. Structured stopping problems in medical decision making are the nearest applied relatives: optimal-timing models such as living-donor transplantation derive control-limit policies from monotone Markov-decision structure (Alagoz et al. 2004); the present withdrawal problem shares that control-limit logic but derives the threshold from an experience-dependent transition law rather than a fixed disease process.

Three distinctions locate the paper within, rather than outside, these literatures. First, adaptive treatment establishes the policy form but not the state-transition law used here. Second, scaffolding establishes fading as a pedagogical principle but not the value margin that identifies ℓ_T and ℓ_W. Third, optimal stopping establishes threshold logic but not the structural–economic gap $\ell_W - \ell^\dagger$. The individual paper's contribution is the bridge among these objects.

2.3 Learning by doing, human capital, and a changing transition technology

The economic idea that present activity can improve future performance is longstanding. Learning by doing treats cumulative production experience as a source of future productivity improvement (Arrow 1962). Human-capital models treat investment as a choice that changes the stock governing later earnings or productive capacity (Ben-Porath 1967). Models of technology choice under learning emphasize that experience affects future performance and therefore feeds back into current allocation (Jovanovic and Nyarko 1996). The skill-formation literature adds dynamic complementarity and stage dependence: earlier capabilities can raise the productivity of later investments, and the timing of inputs matters for the resulting capability path (Cunha and Heckman 2007; Cunha, Heckman, and Schennach 2010).

The present model shares the intertemporal investment logic but uses a different state object. Operator legibility is not a general stock of cognitive skill, output capacity, or human capital. It is a boundary state that changes how a recurrent operator becomes structurally visible in later episodes. Its economic importance appears through the joint distribution of future coarse state and future legibility, and through the resulting probabilities of entry, fusion, and re-closure. The model therefore treats learning as a change in transition conditions rather than as an additive payoff stock.

That distinction motivates the policy-fixed Learning Dividend. The learned and frozen environments retain the same current episode law, current payoff accounting, terminal coarse-state marginal, and contingent policy rule. They differ only in whether realized qualified experience, natural loss, and evaluative capture update ℓ. The value difference

$$\mathrm{LD}_j^\pi(\ell_0) = V_{L,j}^\pi(\ell_0) - V_{F,j}^\pi(\ell_0)$$

is therefore a transition-law value, not the current benefit of entry and not the value of choosing a different policy after learning. Reoptimization is reported separately as an adaptation premium.

The Learning Dividend is signed. This is another boundary from standard favorable-learning language. Qualified experience can raise future legibility, but natural decay can offset formation, and evaluative capture can make the same episode worsen later transition conditions. The exact performance-difference identity values the resulting one-

step kernel wedge over the learned discounted occupation measure. Positive expected movement of ℓ is not, by itself, enough when destination states and continuation values are nonlinear. The paper thus does not claim that learning by doing is always beneficial; it develops an accounting of formation benefit minus natural-loss and capture charges within a specific transition system.

The relationship to human-capital economics is therefore one of structure, not equivalence. Both approaches value a current input partly through future capability. The present model adds a stopping relation that is central to assisted transition: the accumulated state is valuable precisely because it can make external support unnecessary. Its key output is not simply a larger stock but an endogenous reduction and eventual liquidation of the active-support relation.

2.4 Experimentation, Bayesian learning, and the trial option

The experimentation literature separates the immediate payoff of an action from the information it creates for later choices. Blackwell's comparison of experiments orders information structures by the decision problems in which one signal can replicate another (Blackwell 1953). Bandit and research-allocation models formalize the tradeoff between current return and learning about an uncertain alternative (Gittins 1979; Roberts and Weitzman 1981). Real-option reasoning adds the value of preserving a future decision while uncertainty resolves (McDonald and Siegel 1986).

The type-uncertainty extension uses these ideas conservatively. A one-episode starter produces a signal about a persistent response type θ, including learning efficiency, capture sensitivity, and re-closure risk. The continuation menu contains withdrawal and selected complete-information policies. If the trial is decision-reversible, the signal may be ignored and the precommitted continuation remains feasible. Conditional on holding the physical trial fixed, the gross information premium is therefore weakly nonnegative.

The paper's trial accounting is deliberately stricter than the statement that "information has value." It distinguishes

$$O = (C - B) + I,$$

where B is the no-trial benchmark, $C - B$ is the signed physical value of conducting the trial while ignoring its signal, I is the information premium from using the signal, and O is the total net trial option. A trial can be informative and still be undesirable because it consumes resources, delays another action, exposes the individual to fusion, or damages future legibility. Conversely, information can justify a bounded starter that a signal-ignored commitment problem would reject when posterior realizations cross a consequential continuation boundary.

This extension does not claim a new general value-of-information theorem. Nor does it model strategic experimentation among competing agents. Its contribution is to place a standard Bayesian option around the already solved complete-information policy, while preserving the distinction between statistical informativeness and a participant-facing intervention. A Blackwell comparison is valid only when the physical trial, current payoff, and state transition are held fixed. If greater measurement visibility changes evaluative capture, the comparison is no longer purely informational; the altered signal technology must be valued as a different intervention.

The trial option remains secondary to the complete-information results. It gives

an additional reason to start small and preserve withdrawal, but it does not replace the signed Learning Dividend, the taper theorem, or the withdrawal threshold. This ordering protects the paper from becoming a generic experimentation model detached from the transition technology that motivates it.

2.5 Motivational crowding, autonomy support, and dependency-oriented help

A separate literature explains why more external input need not monotonically improve the outcome it targets. Early experiments showed that externally mediated rewards can alter intrinsic motivation (Deci 1971), and a broad meta-analysis examined when extrinsic rewards undermine or preserve intrinsic motivation (Deci et al. 1999). Economic formulations describe crowding as a change in motivation induced by intervention rather than as an ordinary resource cost (Frey and Jegen 2001); experiments and theory show that weak incentives or explicit control can reduce performance or voluntary effort in some environments (Gneezy and Rustichini 2000; Bénabou and Tirole 2003; Falk and Kosfeld 2006). Later synthesis emphasizes that incentives and social preferences can be substitutes or complements depending on information, framing, and institutional context (Bowles and Polania-Reyes 2012).

Self-determination research supplies a closely related distinction between need-supportive and need-thwarting environments (Ryan and Deci 2000). Recent meta-analytic evidence associates supportive behaviors with engagement and well-being and thwarting behaviors with the opposite pattern, while also showing that the constructs require careful operational separation (Howard et al. 2025). The educational evidence that structure and autonomy support can reinforce one another is especially important for the present model (Patzak and Zhang 2025). It prevents a false binary in which all support is controlling and all autonomy requires immediate withdrawal.

Research on helping relations adds the dependence margin. Assistance can be autonomy oriented, enabling the recipient to solve future problems without the helper, or dependency oriented, preserving the helper's role and the recipient's reliance (Nadler 2002; Nadler and Chernyak-Hai 2014). This distinction is substantively close to the paper's concern with a support relation whose success should eventually reduce its own necessity.

The present paper does not identify evaluative capture with any one psychological mechanism in these literatures. Evaluative capture is a model-specific reduced-form channel: support can increase contact while the active frame turns observation into an achievement, ownership, or evaluation project. It has two economic incidences. The current payoff records contemporaneous pressure or burden. The slow update records a possible loss of future legibility, which lowers the value of later episodes. Keeping both effects is not double counting because they occur at different dates and through different state variables.

This channel explains why the support action is not globally ordered by numerical intensity. A higher u can raise contact and entry while also raising pressure and capture. The model therefore imposes neither "more support is better" nor "support crowds out intrinsic motivation." It derives a safe interior starter under concavity, allows a learning penalty when capture dominates formation, and derives tapering from the declining marginal value of support as ℓ rises. The contribution is the dynamic valuation of a dual-channel support technology, not a new universal crowding proposition.

2.6 Management accounting, mission metrics, and service dependence

Management-accounting research has long emphasized that a performance measure can be informative yet incongruent with the principal objective. Multitask models show that rewarding a measured task can redirect effort away from less measured dimensions (Holmström and Milgrom 1991). Performance-measure congruity and diversity determine whether a metric tracks the desired action profile (Baker 1992; Feltham and Xie 1994). Reviews of nonfinancial and contemporary performance measurement likewise stress that system design changes behavior and can generate both intended and unintended consequences (Ittner and Larcker 1998; Franco-Santos et al. 2012).

The public and mission-oriented settings nearest to the present setting make the denominator problem especially visible. Explicit targets can induce gaming when measured outputs are easier to improve than the underlying mission (Courty and Marschke 2004; Bevan and Hood 2006). The effect of public-sector performance measurement depends on how the system is used rather than on the mere presence of indicators (Speklé and Verbeeten 2014). Social-performance research distinguishes operational scale from the scope and depth of mission outcomes (Ebrahim and Rangan 2014), while work on hybrid organizations identifies mission drift and accountability conflicts when financial and social objectives coexist (Ebrahim, Battilana, et al. 2014).

The individual paper enters this literature before agency. The baseline planner receives no continuation rent and does not manipulate measures. Accordingly, the model does not explain why a provider might strategically retain an individual, hide progress, or choose a metric that protects revenue. Those are questions for the downstream population paper. The present model supplies the first-best benchmark against which such behavior can later be identified: when continuation is individually valuable, what intensity is optimal, and at what state should withdrawal occur?

Within that boundary, the model contributes a set of dynamic accounting objects. The one-shot valuation error is exactly the negative of the policy-fixed Learning Dividend, so a static scorecard can be conservative under positive learning and optimistic under adverse learning. Re-closure-adjusted handoff treats first entry as an intermediate output rather than a durable outcome. The threshold gap $\ell_W - \ell^\dagger$ separates structural readiness from economic withdrawal. Premature-withdrawal and over-retention accounts price departures from the continue–withdraw margin, while the intensive-margin loss prices the wrong dose conditional on continuation. Cost to Durable Independence divides discounted resource cost by the probability of a durable withdrawal rather than by episodes delivered or openings recorded.

These measures refine, rather than reject, activity accounting. Tokens, staff time, episodes, and monitoring remain necessary resource inputs. The problem is that activity is observable while future support avoided appears as an absence. A scorecard that rewards only delivered service can therefore make dependence look productive; a scorecard that rewards only rapid exit can make abandonment look efficient. The three-ledger architecture—resource, transition, and exit—keeps both errors visible.

The paper calls the resulting control relation individually self-liquidating. This means that the policy supplies support when its incremental value is positive, reduces support as that value declines, and reaches a state at which withdrawal is optimal. It does not mean self-financing, and it does not propose recognition of operator legibility as an accounting asset. The term describes the mission logic of a control relation that removes its own economic reason for continuing. Aggregate solvency and the

organization's willingness to implement that logic remain downstream problems.

Visible measurement receives the same treatment. A back-end estimator that improves the planner's information while leaving the individual transition law unchanged can be valued through the information account. A participant-facing score, ranking, or attention KPI that changes pressure or capture is part of the intervention. Management accounting must then charge both the current burden and the shadow value of the damaged future transition law. This is the paper's precise connection between measurement design and the individual dynamic model.

2.7 Dynamic-programming and monotone-comparative-statics foundations

Several mathematical results used in the paper are standard tools rather than substantive contributions. Discounted dynamic programming supplies the contraction architecture used to establish a unique bounded value function and stationary optimal policy (Blackwell 1965). Puterman (1994) is the standard modern treatment of the discounted Markov-decision architecture used here. Monotone comparative statics supplies the order logic behind the state taper when the continuation return has decreasing differences (Milgrom and Shannon 1994; Topkis 1998). Structural properties of stochastic dynamic programs—monotone values and control-limit policies derived from ordered transition kernels—are systematized by Smith and McCardle (2002). Blackwell's experiment order supplies the pure information comparison in the trial extension (Blackwell 1953). The performance-difference identity behind the Learning Dividend is, in algebraic form, the policy-difference lemma familiar from reinforcement-learning analysis (Kakade and Langford 2002), applied here to one policy across two transition technologies rather than to two policies under one technology.

The paper's contribution is therefore not Bellman contraction, Topkis monotonicity, the policy-difference lemma, or nonnegative value of information. It lies in identifying economically defensible primitives under which those tools apply to the companion transition system, and in keeping the boundaries among the results explicit. Inherited first-event order does not automatically imply order of the full joint kernel. Value monotonicity does not automatically imply a stopping threshold. A threshold on the stop/continue margin does not order positive support intensities. A state taper does not automatically produce a one-way calendar path. An informative trial does not automatically have positive total value. The analytical sequence below is designed precisely to prevent those logical substitutions.

2.8 Contribution ledger and priority discipline

The review supports four contribution claims at the individual-paper level.

1. **A value interface for an experience-dependent transition technology.** The paper carries a joint kernel for next coarse state and next legibility from the upstream model into a discounted individual decision problem. It does not replace the kernel by a mean update or revalue the same future state change as an additional current reward.

2. **A signed transition-investment value.** The policy-fixed Learning Dividend compares learned and frozen transition technologies under the same contingent rule and admits dividend, neutrality, and penalty. Its performance-difference

identity and formation–loss–capture ledger separate transition-law value from current benefit and from policy adaptation.

3. **Three distinct boundaries and one derived policy architecture.** The structural persistence threshold $\ell^{\dagger}$, taper-onset threshold ℓ_T, and economic withdrawal threshold ℓ_W answer different questions. Interior starter, positive consolidation, decreasing-differences taper, and threshold withdrawal combine into four optimal policy regions without placing those labels in the state definition.
4. **Individual-side accounting for durable independence.** Learning-sensitive valuation, timing-error accounts, Cost to Durable Independence, and separate physical and informational trial accounts translate the dynamic model into decision and control objects while leaving provider agency and population finance for the next paper.

These claims are intentionally narrower than several tempting alternatives. The paper does not claim that support withdrawal has been unstudied, that fading is new, that learning has never been valued as investment, that information value is original, or that performance measures have not previously been shown to distort behavior. Nor does it treat the numerical stress test as calibration or the transition technology as a demonstrated therapeutic effect.

Positioning discipline. The novelty claim is the coupling of a specific experience-dependent entry/fusion/re-closure technology with an individual economic stopping and intensity problem. General concepts are credited to their home literatures; novelty is claimed for the resulting configuration, distinctions, and derived accounting objects.

The paper also makes four negative delimitations explicit. First, operator legibility is not priced as an intrinsic commodity; only specified functional consequences receive mission weights. Second, the first-best planner has no provider rent, so strategic retention and mission drift are not solved here. Third, the baseline observes (J, ℓ); endogenous measurement and a full belief-state control problem are extensions. Fourth, population allocation, capacity, token liability, reserve adequacy, and solvency are not part of the individual model.

2.9 Literature-to-model crosswalk

Table 1 records the division of labor. Its purpose is not to rank fields, but to distinguish established components from the paper's contribution.

Table 1: Literature-to-model crosswalk

Literature	**Established object used here**	**Specific role and remaining difference in this paper**
Dynamic treatment regimes and JITAIs	History- or state-contingent intervention rules over repeated decision points	Restricts the state to (J, ℓ) and makes support alter the future transition law; adds irreversible withdrawal and the distinction between state threshold and calendar time.
Scaffolding and fading	Contingent assistance, fading, and transfer of responsibility	Does not claim fading as new; derives a consolidation plateau and a state taper from the solved continuation return, allowing phase revisits after deterioration.
Optimal stopping and real options	Continuation versus stopping; value of preserving an irreversible choice	Defines the autonomous value $W(\ell)$ and derives ℓ_W; separates it from the structural feasibility threshold $\ell^\dagger$.
Learning by doing and skill formation	Current activity or investment changes future capability or productivity	Uses operator legibility as a transition parameter rather than a generic skill stock; permits adverse learning and defines a policy-fixed signed value.
Bayesian experimentation	Information can improve contingent policy selection	Keeps the physical trial fixed for the pure information comparison; separates I, $C - B$, and total option value O.
Motivational crowding and autonomy support	External intervention can support, control, crowd out, or become need-thwarting	Represents support as a dual-channel technology that can increase contact and evaluative capture; prices current burden and future transition damage separately.
Management accounting and mission measurement	Measure congruity, multitasking, gaming, social-performance scope, and mission drift	Supplies an individual first-best benchmark: Learning Dividend, threshold gap, timing losses, and Cost to Durable Independence; provider incentives remain downstream.
Dynamic programming and monotone comparative statics	Contraction, stationary policy, information order, and monotone policy selection	These are proof tools, not contribution claims; the paper's task is to state the primitives of the present transition system that make each tool applicable without conflating their conclusions.

The crosswalk reveals why the paper is not merely a Bellman extension of the learning model. The economic contribution is not the presence of a value function by itself. It is the set of counterfactuals and boundaries that the value function makes possible: learned versus frozen transition technology; current durability versus future formation; structural readiness versus economic withdrawal; state taper versus chronological taper; physical trial value versus information value; and delivered activity versus durable independence.

2.10 Bridge to the economic model

The model that follows therefore takes a specific route through the literatures reviewed above. It begins with an aligned planner and the reduced state (J, ℓ). It receives from the upstream paper a joint kernel for next economic state and next operator legibility, together with first-event summaries for entry, fusion, and re-closure. It values current functional consequences without assigning an intrinsic price to attention, distinguishes remaining active at zero intensity from irreversible withdrawal, and solves the support and stopping problem before introducing provider incentives or population

aggregation. The analytical results then ask four questions that the related literatures motivate but do not answer for this transition technology: What is the signed value of allowing experience to change the future law? When is positive post-entry consolidation warranted? Why should optimal support intensity decline with legibility? At what economic boundary should the active relation end, and how does that boundary differ from structural persistence?

Those questions define the manuscript's contribution and the appropriate reading of Section 3. The transition technology is not a metaphor appended to a generic intervention model, and the economic objective is not a welfare label appended to the upstream dynamics. The paper is the bridge between them.

3 Economic Model and Transition Interface

3.1 Economic question and division of labor

The companion transition model specifies how a recurrent operator becomes more or less legible across episodes and how that state changes entry, fusion, re-closure, and post-withdrawal persistence. The present paper takes that technology as an input and assigns value to its consequences. The decision question is

> For one individual, when is external support worth starting, retaining after qualified entry, tapering, and finally withdrawing once current functioning, future transition-law value, re-closure risk, resource cost, and evaluative capture are considered together?

The separation is substantive. Economic weights evaluate the outputs of the transition technology; they do not redefine attention entry, fusion, re-closure, or operator legibility. Conversely, favorable transition probabilities do not settle the economic decision because the active relation has costs and an autonomous outside branch.

Principle 3.1 (Separation of transition technology and valuation)**.** *The transition law is an input to the individual economic problem. Mission weights value specified functional consequences of that law; they do not alter the meaning of the underlying states or events.*

3.2 Planner, episodes, and economic states

Time is indexed by episodes $n = 0, 1, 2, \ldots$ and discounted by $\beta \in (0, 1)$. The baseline decision maker is a benevolent planner aligned with one individual's long-run functional objective. The planner may be interpreted as the individual acting with commitment or as an adviser, educator, clinician, or algorithm operating under a fully aligned mandate. There are no transfers between strategic parties, no provider rent, and no distributional weights.

The beginning-of-episode state is

$$x_n = (J_n, \ell_n), \qquad J_n \in \{R, A\}, \quad \ell_n \in [0, 1]. \tag{3.1}$$

The unresolved or re-closed state R contains both the pre-entry condition and the condition reached after re-closure. Support in R is therefore starter or re-entry support. The state A follows a qualified entry while the active-support problem remains open.

Support in A is consolidation or tapering support. The label A does not imply cure, permanent attention, or completed independence.

Fusion F remains an adverse within-episode event rather than a third persistent economic state. Its incidence enters current damage and the transition kernel. An episode ending in fusion returns to R at the next economic decision point. If fusion has persistent consequences not summarized by (J, ℓ), the state must be enlarged; that is outside the baseline.

Assumption 3.2 (Complete information and reduced Markov sufficiency)**.** *At the beginning of each episode the planner observes (J_n, ℓ_n) without error. Conditional on the current state and current support action, this pair is sufficient for the law of current outcomes and the next economic state.*

As in the closed-loop illustration of the companion transition paper, this reduction is purchased by a structural device rather than assumed for free: the fast state is reset to a fixed episode-initial value at each economic decision point, so operator legibility is the only cross-episode carrier. Residual fast-state carryover—in curiosity, visibility, pressure, or object sensitivity—would introduce a second economic state variable and is the first extension recorded in Section 7.

3.3 Support, withdrawal, and timing

The direct action is a scalar support intensity

$$u_n \in \mathcal{U}_{J_n} := [0, \bar{u}_{J_n}], \qquad 0 < \bar{u}_{J_n} < \infty. \tag{3.2}$$

Support may represent tokens, access, structured contact, facilitation, or another externally supplied occasion for engaging the target task. The scalar action isolates the canonical intensity margin; a multidimensional action could later distinguish contact, measurement, personalization, and participant-facing evaluation.

The central model sets a separate participant-facing attention KPI to zero. This does not eliminate pressure. Support itself may generate an episode-level evaluative-capture burden $G_j(\ell, u)$. No global quality order is imposed on u: support can improve contact while simultaneously increasing resource use and evaluative capture, so an interior safe region is possible.

Withdrawal is available only in state A. It terminates the active-support regime and transfers the individual to an autonomous continuation value $W(\ell)$. Withdrawal is irreversible within the baseline contract. It is distinct from selecting $u = 0$ while remaining active: continued enrollment preserves future intervention options and may consume scheduling, monitoring, organizational, or psychological resources. We therefore write

$$C_A(\ell, u) = f_A + c_A(\ell, u), \qquad f_A > 0, \qquad c_A(\ell, u) \geq 0, \tag{3.3}$$

so that a zero current dose is not a costless option to retain the program forever.

3.4 Episode experiment and the joint kernel

Fix current state (j, ℓ) and action u. The micro-founded episode produces a terminal coarse state $J^+ \in \{R, A\}$, qualified experience $Y \in [0, 1]$, and evaluative capture

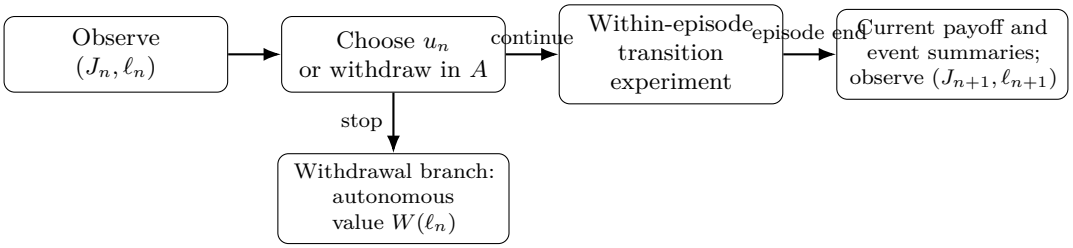

Figure 1: Timing of one economic episode. The within-episode fast dynamics are compressed into a joint transition experiment; only (J, ℓ) is carried to the next decision point.

$G \in [0,1]$. Let

$$\mathsf{P}_{j,\ell,u}(dk, dy, dg) \tag{3.4}$$

be their joint law. The cross-episode update is

$$\Phi(\ell, y, g) = \ell + \eta y(1-\ell) - (\delta_\ell + \chi g)\ell = \eta y + \{1 - \eta y - \delta_\ell - \chi g\}\ell, \tag{3.5}$$

where $\eta \geq 0$ is formation efficiency, $\delta_\ell \geq 0$ natural loss, and $\chi \geq 0$ sensitivity to evaluative capture.

Proposition 3.3 (State-space invariance)**.** *If $\ell, y, g \in [0,1]$, the parameters are nonnegative, and*

$$\eta y + \delta_\ell + \chi g \leq 1, \tag{3.6}$$

then $\Phi(\ell, y, g) \in [0,1]$.

Proof. Condition (3.6) makes the coefficient on ℓ in the second expression of (3.5) nonnegative, so $\Phi \geq \eta y \geq 0$. Since $\ell \leq 1$,

$$\Phi \leq \eta y + 1 - \eta y - \delta_\ell - \chi g = 1 - \delta_\ell - \chi g \leq 1.$$

This direct lower–upper bound is the maintained invariance argument; no literal convex-combination claim is needed. □

For Borel $B \subseteq [0,1]$ and $k \in \{R, A\}$, define the induced joint kernel

$$\mathsf{K}_j(B, k \mid \ell, u) := \int \mathbf{1}_{\{k'=k,\, \Phi(\ell,y,g)\in B\}} \mathsf{P}_{j,\ell,u}(dk', dy, dg). \tag{3.7}$$

For a bounded vector function $v = (v_R, v_A)$, write

$$\mathcal{K}_j^u v(\ell) := \sum_{k\in\{R,A\}} \int_{[0,1]} v_k(\ell')\, \mathsf{K}_j(d\ell', k \mid \ell, u). \tag{3.8}$$

The kernel, rather than a conditional mean of ℓ^+, is the economic interface. It preserves downside mass and dependence between the next coarse state and next legibility.

3.5 First-event summaries and inherited order

The transition paper supplies three current-episode summaries that enter payoffs:

$$p_A^R(\ell,u) := \Pr\{\text{qualified entry wins the first-resolution competition} \mid R,\ell,u\}, \tag{3.9}$$

$$p_F^R(\ell,u) := \Pr\{\text{fusion wins the first-resolution competition} \mid R,\ell,u\}, \tag{3.10}$$

$$r_H^A(\ell,u) := \Pr\{\text{first re-closure by horizon } H \mid A,\ell,u\}. \tag{3.11}$$

These are not generally terminal-state probabilities. Entry may be followed by re-closure before the episode ends, and fusion is mapped to R at the next economic boundary. Replacing the joint terminal kernel by these first-event summaries would therefore misprice continuation.

Under matched fast paths, the upstream model supplies the order

$$\ell_2 \geq \ell_1 \quad \Longrightarrow \quad p_A^R(\ell_2,u) \geq p_A^R(\ell_1,u), \quad p_F^R(\ell_2,u) \leq p_F^R(\ell_1,u), \quad r_H^A(\ell_2,u) \leq r_H^A(\ell_1,u). \tag{3.12}$$

This order is inherited rather than re-claimed here. It does not by itself imply that the full joint kernel is monotone for every continuation value. The dynamic results below therefore state the stronger value-relevant kernel order explicitly when they need it.

Assumption 3.4 (Feller interface)**.** *For every bounded continuous vector function v, the map $(\ell,u) \mapsto \mathcal{K}_j^u v(\ell)$ is continuous for $j \in \{R,A\}$.*

Continuous dependence of the fast ODE on (ℓ,u), bounded continuous intensities, a finite episode horizon, and a continuous state update provide a standard micro-founded route to this condition.

3.6 Current payoff and autonomous value

All benefits and harms are placed on a common normalized functional scale. The scale is cardinal for the planner's choices but need not initially be monetary. The active one-episode payoff is

$$r_j(\ell,u) = B_j(\ell,u) - D_j(\ell,u) - C_j(\ell,u) - d_G G_j(\ell,u), \qquad d_G \geq 0. \tag{3.13}$$

A canonical reduced form is

$$r_R(\ell,u) = b_A p_A^R(\ell,u) - d_F p_F^R(\ell,u) - C_R(\ell,u) - d_G G_R(\ell,u), \tag{3.14}$$

$$r_A(\ell,u) = B_A(\ell,u) - d_R r_H^A(\ell,u) - C_A(\ell,u) - d_G G_A(\ell,u). \tag{3.15}$$

The state label A receives no positive payoff merely because it is called attention. Every positive or negative weight attaches to a specified functional consequence.

Qualified experience consequently generates two distinct value consequences at different dates and through different state variables, exactly as evaluative capture does: a contemporaneous functional-flow consequence and a future transition-law consequence carried exclusively through the kernel. Within the episode, qualified engagement is a current functional flow and may carry a current weight (the b_Y terms in the numerical reduced forms of Section 5); across episodes, the same realized Y drives the slow update and is priced only through continuation values at the induced ℓ^+. There is no double counting: the flow consequence records what the episode was worth while it occurred,

and the kernel consequence records how it changed the law of later episodes. The policy-fixed Learning Dividend of Section 4.3 isolates the kernel consequence by construction, because the learned and frozen technologies share an identical current-payoff accounting, including any b_Y flow.

The autonomous withdrawal value $W(\ell)$ is the discounted value of the regime beginning when active support is terminated. It may incorporate unsupported persistence, natural loss, re-closure risk, and exogenous future re-entry cost, but it contains no discretionary support action by the planner in the baseline. A reduced benchmark is

$$W(\ell) = \frac{b_0 - d_R \bar{r}(\ell)}{1-\beta}, \qquad \bar{r}'(\ell) < 0, \tag{3.16}$$

though the analytical results require only the regularity and order conditions stated when used.

3.7 Policies, objective, and Bellman system

An admissible policy is nonanticipative. Before withdrawal it selects an action permitted by the current state; in state A it may instead stop. Let

$$\tau_W := \inf\{n \geq 0 : \text{the policy withdraws in state } A\}, \tag{3.17}$$

with $\tau_W = \infty$ if withdrawal never occurs. For initial state (j, ℓ), policy value is

$$V_j^\pi(\ell) := \mathbb{E}_{j,\ell}^\pi \left[\sum_{n=0}^{\tau_W - 1} \beta^n r_{J_n}(\ell_n, u_n) + \mathbf{1}_{\{\tau_W < \infty\}} \beta^{\tau_W} W(\ell_{\tau_W}) \right]. \tag{3.18}$$

The optimal value is $V_j(\ell) = \sup_\pi V_j^\pi(\ell)$. Its Bellman representation is

$$V_R(\ell) = \max_{u \in \mathcal{U}_R} \left\{ r_R(\ell, u) + \beta \mathcal{K}_R^u[V_R, V_A](\ell) \right\}, \tag{3.19}$$

$$V_A(\ell) = \max \left\{ W(\ell), \max_{u \in \mathcal{U}_A} \left[r_A(\ell, u) + \beta \mathcal{K}_A^u[V_R, V_A](\ell) \right] \right\}. \tag{3.20}$$

Section 4 establishes existence, uniqueness, verification, learning value, stopping shape, and policy shape.

3.8 Three boundaries and the paper's scope

The paper separates three state boundaries and one ex post performance endpoint.

- The structural persistence threshold $\ell^\dagger$ is inherited from the transition paper. Above it, a positive post-withdrawal persistence basin is geometrically available. It is not an economic stopping rule.
- The taper-onset threshold ℓ_T is an intensive-margin boundary at which the optimal positive post-entry action leaves a consolidation plateau.
- The economic withdrawal threshold ℓ_W is an extensive-margin boundary at which the optimized active branch and autonomous value are equal.
- A durable-independence hitting time τ_G is an ex post certification endpoint. It need not coincide with the optimal stopping time.

No primitive identity requires ℓ_T, $\ell^\dagger$, and ℓ_W to coincide. In particular,

$$\ell_W < \ell^\dagger, \qquad \ell_W = \ell^\dagger, \qquad \ell_W > \ell^\dagger \tag{3.21}$$

are all admissible economic regimes.

Table 2 records the frozen baseline.

Table 2: Model constitution for the individual economic paper

Object	Baseline choice
Decision maker	Benevolent planner aligned with one individual's long-run functional objective; first-best and no provider rent.
Time	Infinite sequence of episodes with discount factor $\beta \in (0,1)$.
State	(J,ℓ) with $J \in \{R,A\}$ and $\ell \in [0,1]$; fusion is a within-episode damage event.
Information	Complete observation and reduced Markov sufficiency.
Action	Scalar support intensity; withdrawal is a separate action available in A.
Policy phases	Starter, consolidation, taper, and withdrawal are derived policy regions, not primitive stages.
Value unit	Normalized cardinal functional units; no intrinsic price is assigned to attention.
Learning	Signed: dividend, neutrality, and penalty are all admissible.
Paper boundary	No empirical efficacy claim, provider agency, population allocation, capacity, liability, reserve, solvency, or baseline POMDP.
Downstream output	Type-specific optimal policy and life-cycle cost, support, duration, re-closure, and benefit summaries.

4 Economic Results

4.1 A one-cycle benchmark for post-entry consolidation

A qualified entry is not yet a stopping theorem. To expose the economic mechanism before returning to the infinite horizon, consider a state (A,ℓ) and compare immediate withdrawal with one additional consolidation block followed by withdrawal. Let $h \in [0,\bar{h}]$ index a one-dimensional family of post-entry support templates. The block produces current durability benefit $B(h;\ell)$, variable resource cost $c(h;\ell)$, current pressure damage $D(h;\ell)$, fixed active-program overhead $f_A > 0$, and terminal legibility law $\mathsf{L}_A^h(d\ell' \mid \ell)$. The autonomous branch has one-cycle kernel L^0 and satisfies

$$W(\ell) = r^0(\ell) + \beta \int W(\ell')\, \mathsf{L}^0(d\ell' \mid \ell). \tag{4.1}$$

Define the variable consolidation surplus

$$\Psi(h;\ell) := B(h;\ell) - c(h;\ell) - D(h;\ell) + \beta \left\{ \int W(\ell')\, \mathsf{L}_A^h(d\ell' \mid \ell) - \int W(\ell')\, \mathsf{L}^0(d\ell' \mid \ell) \right\}. \tag{4.2}$$

The second integral is the *no-consolidation/autonomous reference*. It is not the clamped-legibility counterfactual used later to define the Learning Dividend.

Proposition 4.1 (Consolidation hurdle)**.** *Suppose $h \mapsto \Psi(h;\ell)$ is continuous and*

strictly concave. Let

$$S(\ell) := \max_{h \in [0,\bar{h}]} \Psi(h;\ell).$$

A positive consolidation block is optimal in the one-cycle benchmark if and only if

$$S(\ell) > f_A. \tag{4.3}$$

If the inequality is reversed, immediate withdrawal is optimal; equality permits indifference. The fixed overhead changes whether the positive branch is selected, not the optimal positive h conditional on selecting it.

Proof. The incremental value of a positive block is $\Psi(h;\ell) - f_A$. Strict concavity gives a unique maximizer of the variable branch. Comparing its maximum with zero gives (4.3). Full curvature and boundary details are in Appendix C. □

Under a common-random-number representation $\ell_h^+ = \mathcal{T}_A(h,\ell,\varepsilon)$ and standard differentiation conditions,

$$\begin{aligned}\Psi_h(h;\ell) = \underbrace{B_h(h;\ell)}_{\text{current durability}} - \underbrace{\{c_h(h;\ell) + D_h(h;\ell)\}}_{\text{resource and current pressure}} \\ + \underbrace{\beta \mathbb{E}\left[W'(\ell_h^+)\frac{\partial \mathcal{T}_A}{\partial h}(h,\ell,\varepsilon)\right]}_{\text{future transition-law value}}.\end{aligned} \tag{4.4}$$

Thus an extra block can be worthwhile even when current benefit alone is insufficient, because it changes the state inherited by the autonomous regime. The same formula permits the opposite case: evaluative capture can make the future term negative and destroy an otherwise attractive consolidation decision.

4.2 Bellman values and conditional value monotonicity

Let $\mathbb{V} = C([0,1]) \times C([0,1])$ with the sup norm. For $v \in \mathbb{V}$, define

$$Q_j(\ell,u;v) := r_j(\ell,u) + \beta \mathcal{K}_j^u v(\ell), \tag{4.5}$$

$$(\mathcal{T}v)_R(\ell) := \max_{u \in \mathcal{U}_R} Q_R(\ell,u;v), \tag{4.6}$$

$$(\mathcal{T}v)_A(\ell) := \max\left\{W(\ell), \max_{u \in \mathcal{U}_A} Q_A(\ell,u;v)\right\}. \tag{4.7}$$

Assumption 4.2 (Bellman regularity)**.** *The action sets are nonempty compact intervals; r_j and W are bounded and continuous; Assumption 3.4 holds; and $\beta \in (0,1)$.*

Theorem 4.3 (Well-posed individual value problem)**.** *Under Assumption 4.2, $\mathcal{T}$ is a contraction of modulus β. It has a unique fixed point $V = (V_R, V_A) \in \mathbb{V}$, value iteration converges uniformly to V, and V equals the supremum of the policy values in* (3.18)*. An optimal stationary deterministic support-and-stopping policy exists.*

Proof sketch. A probability kernel is nonexpansive in the sup norm. Maximization preserves the bound, so $\|\mathcal{T}v - \mathcal{T}w\|_\infty \le \beta \|v - w\|_\infty$. The contraction mapping theorem gives the fixed point. Compactness and continuity give measurable maximizing selectors; iterating the Bellman inequalities along any policy and then along a selector verifies optimality. Appendix A supplies the residual, selector, verification, and perturbation arguments. □

Event order (3.12) alone is insufficient for value monotonicity because a continuation value depends jointly on J^+ and ℓ^+. Order the joint state by the product of the coarse order $R \prec A$ and the legibility order, and let

$$\mathbb{V}_\uparrow := \{v \in \mathbb{V} : v_R \text{ and } v_A \text{ are nondecreasing in } \ell \text{ and } v_A(\ell) \geq v_R(\ell) \text{ for all } \ell\},$$

the closed cone of functions nondecreasing in that product order.

Assumption 4.4 (Value-relevant product order)**.** *(i) For every $v \in \mathbb{V}_\uparrow$, state j, and action u,*

$$\ell_2 \geq \ell_1 \quad \Longrightarrow \quad \mathcal{K}_j^u v(\ell_2) \geq \mathcal{K}_j^u v(\ell_1). \tag{4.8}$$

(ii) The Bellman operator preserves the cross-state ranking: $(\mathcal{T}v)_A(\ell) \geq (\mathcal{T}v)_R(\ell)$ for every $v \in \mathbb{V}_\uparrow$ and every ℓ. In addition, W and $r_j(\cdot, u)$ are nondecreasing in ℓ.

Clause (i) is first-order dominance of the joint kernel in the product order; its finite-grid analogue is exactly the upper-set audit reported in Appendix G. Clause (ii) is maintained at the operator level; Appendix A records a primitive sufficient route through $\mathcal{U}_R \subseteq \mathcal{U}_A$ that the numerical action domains do not satisfy, which is why the clause is stated directly rather than derived. The quantifier in clause (i) deliberately runs over the product cone and not over the larger cone that drops the cross-state ranking: applied to the componentwise-constant test functions of that larger cone, kernel monotonicity would force the terminal coarse-state marginal to be invariant in legibility. On the solved grid that invariance fails directly: the terminal-A probability strictly increases in legibility at every maintained action with positive support (Appendix G, Table 12), so the unranked order cannot hold for the micro-founded kernel. Appendix A states the calculation and retains the same-state order only as a benchmark.

Proposition 4.5 (Value monotonicity and value ranking)**.** *Under Assumptions 4.2 and 4.4, V_R and V_A are nondecreasing in operator legibility and $V_A(\ell) \geq V_R(\ell)$ at every ℓ.*

Proof. For $v \in \mathbb{V}_\uparrow$, clause (i) makes every action-specific return nondecreasing in ℓ; maximization and the stop option with nondecreasing W preserve monotonicity, and clause (ii) preserves the ranking, so $\mathcal{T}\mathbb{V}_\uparrow \subseteq \mathbb{V}_\uparrow$. The zero vector lies in the cone; value iteration and the uniform limit prove the claim because the cone is closed. □

The content of clause (i) can be located exactly. Under the maintained no-overshoot condition (3.6), the slow map (3.5) satisfies $\partial\Phi/\partial\ell = 1 - \eta y - \delta_\ell - \chi g \geq 0$ and is nondecreasing in y and nonincreasing in g, so monotonicity of Φ in its own state costs nothing beyond the maintained primitives; what the clause adds is stochastic monotonicity of the episode outputs (J^+, Y, G) in legibility at a fixed action, with the terminal coarse state permitted to upgrade from R to A. Lemma A.11 in Appendix A formalizes this through a monotone-destination coupling and identifies the residual as the object whose theorem-level status the companion paper leaves open: the post-entry state is endogenous to the entry time, so the required couplings are not supplied by the matched-state ordering results. Clause (i) is thus weaker than a claim inherited from the transition paper, stronger than the inherited first-event order (3.12), and strictly weaker than the same-state benchmark, which in addition freezes the destination marginal. On the solved grid the conclusion of the ranking clause is realized with $\min_\ell\{V_A(\ell) - V_R(\ell)\} = 1.015$.

4.3 The policy-fixed signed Learning Dividend

The canonical learning comparison holds the current episode and the contingent policy rule fixed and changes only the cross-episode legibility update. Let

$$\Phi^{\mathrm{L}}(\ell, y, g) = \ell + \eta y(1-\ell) - (\delta_\ell + \chi g)\ell, \qquad \Phi^{\mathrm{F}}(\ell, y, g) = \ell. \tag{4.9}$$

The superscript F denotes *frozen legibility*: the terminal coarse-state marginal, current payoff, action set, and current episode law remain the same, while legibility is clamped at its current value. Let $\mathsf{K}_j^{\mathrm{L}}$ and $\mathsf{K}_j^{\mathrm{F}}$ be the induced joint kernels.

For a fixed stationary policy π, let P_m^π be its substochastic continuation operator under technology $m \in \{\mathrm{L}, \mathrm{F}\}$, with continuation mass terminated at withdrawal, and let V_m^π be its fixed-policy value.

Definition 4.6 (Policy-fixed signed Learning Dividend)**.** *For initial state* (j, ℓ_0)*,*

$$\mathrm{LD}_j^\pi(\ell_0) := V_{\mathrm{L},j}^\pi(\ell_0) - V_{\mathrm{F},j}^\pi(\ell_0). \tag{4.10}$$

A positive value is a dividend, a negative value a learning penalty, and zero a neutral transition-law effect under that policy.

The same contingent mapping π is used in both technologies. Once state histories diverge, realized actions may differ because the same state-contingent rule is applied to different states; this is not reoptimization.

Theorem 4.7 (Performance-difference identity)**.** *For every fixed policy,*

$$\mathrm{LD}^\pi = \beta(I - \beta P_{\mathrm{L}}^\pi)^{-1}(P_{\mathrm{L}}^\pi - P_{\mathrm{F}}^\pi)V_{\mathrm{F}}^\pi, \tag{4.11}$$

$$= \beta \sum_{n=0}^{\infty} \beta^n (P_{\mathrm{L}}^\pi)^n \omega^\pi, \qquad \omega^\pi := (P_{\mathrm{L}}^\pi - P_{\mathrm{F}}^\pi)V_{\mathrm{F}}^\pi. \tag{4.12}$$

Equivalently, the Learning Dividend is the learned discounted occupation of the one-step transition-law wedge ω^π *before withdrawal.*

Proof sketch. Subtract the two fixed-policy equations, add and subtract $\beta P_{\mathrm{L}}^\pi V_{\mathrm{F}}^\pi$, and invert $I - \beta P_{\mathrm{L}}^\pi$. The Neumann series gives (4.12). Appendix B gives the stopped-path formulation and integrability details. □

Identity (4.12) is, in algebraic form, the discounted policy-difference (performance-difference) lemma of reinforcement-learning analysis (Kakade and Langford 2002), with the two continuation operators indexed by transition technology at a fixed policy rather than by policy at a fixed technology. The contribution is the counterfactual the identity prices, not the algebra.

The identity signs mixed cases correctly. A favorable wedge in a state never reached has no value; a small adverse wedge in a frequently occupied state can dominate a large favorable wedge in a rare state. If the frozen policy values are nondecreasing in ℓ, the pathwise condition $\Phi^{\mathrm{L}} \geq \Phi^{\mathrm{F}}$ at every reachable continuation state is sufficient for $\mathrm{LD}^\pi \geq 0$; reversing the order gives a nonpositive value. Positive mean drift alone is not sufficient when continuation values are nonlinear or correlated with the destination state.

The slow map also yields an exact net ledger. Insert formation, then natural loss, then capture as sequential counterfactual channels. The one-step wedge and its

dynamic propagation can be written

$$\mathrm{LD}^{\pi} = \mathscr{B}_Y^{\pi} - \mathscr{C}_{\delta}^{\pi} - \mathscr{C}_G^{\pi}, \tag{4.13}$$

where the three gross terms are nonnegative under monotone frozen-policy values. The net identity is canonical; the allocation of nonlinear interaction terms among gross components depends on the stated insertion order. Current pressure damage already included in r_j is distinct from $\mathscr{C}_G^{\pi}$, which prices the future-value loss transmitted through next-period legibility.

Separate optimization in the two technologies produces

$$\mathrm{LD}^{*} := V_{\mathrm{L}}^{*} - V_{\mathrm{F}}^{*}, \tag{4.14}$$

and the bounds

$$\mathrm{LD}^{\pi_{\mathrm{F}}^{*}} \le \mathrm{LD}^{*} \le \mathrm{LD}^{\pi_{\mathrm{L}}^{*}}. \tag{4.15}$$

Moreover,

$$\mathrm{LD}^{*} = \mathrm{LD}^{\pi_{\mathrm{F}}^{*}} + \underbrace{\left(V_{\mathrm{L}}^{*} - V_{\mathrm{L}}^{\pi_{\mathrm{F}}^{*}}\right)}_{\text{adaptation premium } \ge 0}. \tag{4.16}$$

Reoptimization can mitigate an adverse technology effect but does not make learning valuable by definition.

4.4 Economic withdrawal and the structural–economic gap

At the Bellman fixed point define the optimized active branch and the continue–withdraw gap

$$Q_A^{*}(\ell) := \max_{u \in \mathcal{U}_A} \{ r_A(\ell, u) + \beta \mathcal{K}_A^{u} V(\ell) \}, \tag{4.17}$$

$$\Delta_A(\ell) := Q_A^{*}(\ell) - W(\ell). \tag{4.18}$$

Then

$$V_A(\ell) = W(\ell) + \max\{0, \Delta_A(\ell)\}. \tag{4.19}$$

Continuity makes the continuation set open and the stopping set closed, but closure alone does not imply a threshold.

Theorem 4.8 (Global and eventual crossing routes)**.** *Suppose Δ_A is continuous and $\Delta_A(0) > 0 > \Delta_A(1)$.*

(a) If Δ_A is strictly decreasing on $[0,1]$, it has a unique root $\ell_W \in (0,1)$.

(b) More generally, suppose there is $\bar{\ell} < 1$ such that $\Delta_A(\ell) > 0$ for $\ell \le \bar{\ell}$ and Δ_A is strictly decreasing on $[\bar{\ell}, 1]$. Then it again has a unique root $\ell_W \in (\bar{\ell}, 1)$.

In either case, continuation is optimal below ℓ_W and withdrawal at and above it under the withdrawal tie rule; the stopping set is the upper interval $[\ell_W, 1]$.

Proof. Part (a) is immediate from continuity, endpoint signs, and strict decrease. In part (b), the gap is positive through $\bar{\ell}$ and negative at one; strict decline on the upper region gives one and only one crossing. Appendix C states the corresponding boundary regimes and primitive upper-region slope conditions. □

Global strict decrease is a transparent sufficient benchmark, not a necessary

property of the numerical model. The micro-founded baseline has eleven low-legibility local increases in Δ_A, is strictly decreasing from $\ell = 0.22$ onward on the disclosed grid, and crosses zero once. Its threshold conclusion is therefore matched by part (b), not by a claim that the solved gap is globally decreasing.

The economic boundary is distinct from the structural persistence threshold. Define

$$\mathfrak{M}_\dagger := \Delta_A(\ell^\dagger). \tag{4.20}$$

Proposition 4.9 (Structural–economic threshold gap)**.** *Under either crossing route in Theorem 4.8,*

$$\mathfrak{M}_\dagger \begin{cases} > 0 & \iff \ell_W > \ell^\dagger, \\ = 0 & \iff \ell_W = \ell^\dagger, \\ < 0 & \iff \ell_W < \ell^\dagger. \end{cases} \tag{4.21}$$

Proof. The gap is positive below its unique root and negative above it. Evaluate it at $\ell^\dagger$. □

Equality is an additional economic balance condition, not an implication of the transition geometry. A positive margin at $\ell^\dagger$ means that persistence has become structurally feasible while further active support still has net learning, durability, risk-reduction, or option value. A negative margin means that active support is too costly or damaging to justify waiting for the structural boundary. In the numerical baseline the margin is strictly positive, so $\ell_W > \ell^\dagger$ (Section 5.6).

Comparative statics are most safely stated as pointwise shifts of Δ_A. Any parameter change that raises the gap at every ℓ weakly raises the unique threshold; a downward shift weakly lowers it. Parameter names alone do not determine the sign. Higher re-closure damage delays withdrawal only when active support reduces the relevant discounted burden, and greater patience delays withdrawal only when the delayed active-over-autonomous differentials are nonnegative.

4.5 Monotone taper and four policy regions

Withdrawal is the extensive margin. Conditional on continuation, the intensive margin solves

$$u_A^*(\ell) \in \arg\max_{u \in \mathcal{U}_A} Q_A(\ell, u), \qquad \ell < \ell_W. \tag{4.22}$$

A *state taper* is a nonincreasing selection $u_A^*(\ell)$. It becomes a chronological taper only along realized paths on which legibility does not fall.

Assumption 4.10 (Decreasing differences)**.** *For $\ell_2 > \ell_1$ and $u_2 > u_1$,*

$$Q_A(\ell_2, u_2) - Q_A(\ell_2, u_1) \le Q_A(\ell_1, u_2) - Q_A(\ell_1, u_1). \tag{4.23}$$

Theorem 4.11 (Continuum monotone taper)**.** *Under Assumption 4.10, the minimum and maximum optimal continuation selections are nonincreasing in ℓ. With a unique maximizer, $u_A^*(\ell)$ is nonincreasing on the continuation region.*

Proof. Suppose the maximal selection rose between ℓ_1 and ℓ_2. Optimality at ℓ_2 and decreasing differences imply that the higher action is also optimal at ℓ_1, contradicting maximality of the lower selected action there. The minimum-selection argument is symmetric. Appendix D gives the full contradiction proof and the strict differentiable case. □

For a finite action set, the computation requires less than global decreasing differences. Let $\mathcal{U}_A = \{u_1 < \cdots < u_m\}$ and $d_{hl}(\ell) := Q_A(\ell, u_h) - Q_A(\ell, u_l)$ for $u_h > u_l$.

Assumption 4.12 (Pairwise action single crossing)**.** *For every $u_h > u_l$ and $\ell_2 > \ell_1$,*

$$\textit{(i)}\ \ d_{hl}(\ell_1) \le 0 \implies d_{hl}(\ell_2) \le 0, \qquad \textit{(ii)}\ \ d_{hl}(\ell_1) < 0 \implies d_{hl}(\ell_2) < 0. \quad (4.24)$$

Once a higher action is weakly (strictly) no better than a lower action, it remains weakly (strictly) no better as legibility rises. The pair is the decreasing form of the Milgrom–Shannon single-crossing property applied to each pairwise advantage; decreasing differences imply both clauses, but not conversely.

Theorem 4.13 (Finite-action monotone taper)**.** *On a finite ordered action set: (a) clause (i) of Assumption 4.12 makes the minimum optimal selection nonincreasing in legibility; (b) clauses (i) and (ii) together make the maximum optimal selection nonincreasing as well. When the maximizer is unique, the two selections coincide.*

Proof. (a) If the minimum selection rose from u_l at ℓ_1 to u_h at ℓ_2, minimality of u_h at ℓ_2 would force $d_{hl}(\ell_2) > 0$, since otherwise u_l would also be optimal there; the contrapositive of clause (i) gives $d_{hl}(\ell_1) > 0$, contradicting optimality of u_l at ℓ_1. (b) If the maximum selection rose from u_l at ℓ_1 to u_h at ℓ_2, maximality of u_l at ℓ_1 would force $d_{hl}(\ell_1) < 0$; clause (ii) gives $d_{hl}(\ell_2) < 0$, contradicting optimality of u_h at ℓ_2. Appendix D restates the argument and records why clause (i) alone does not order the maximum selection. □

Clause (i) alone leaves the maximum selection unordered: $d_{hl}(\ell_1) < 0 = d_{hl}(\ell_2)$ is compatible with clause (i), and the tie at the higher state admits an upward jump of the maximal selection. In the numerical audit, 32 of 192 adjacent decreasing-differences checks fail, while all 21 pairwise action comparisons satisfy clause (i) of (4.24), and the reported post-entry policy—the minimum optimal selection, with ties broken toward the smaller action—has no upward step. Clause (i) is therefore exactly the condition the computation instantiates; the continuum theorem, the weak clause, and the strict clause perform different jobs and are reported separately.

A distinct consolidation region requires the optimizer initially to remain at the upper support boundary. Suppose Q_A is strictly concave in u, $Q_{A,u\ell} < 0$, the upper-bound marginal is positive at low ℓ and negative before ℓ_W, and $Q_{A,u}(\ell, 0) > 0$ below ℓ_W. Then a unique taper-onset state ℓ_T satisfies $Q_{A,u}(\ell_T, \bar{u}_A) = 0$, and

$$\pi_A^*(\ell) = \begin{cases} \bar{u}_A, & 0 \le \ell \le \ell_T, \\ u_A^*(\ell) \in (0, \bar{u}_A), & \ell_T < \ell < \ell_W, \\ \mathsf{W}, & \ell \ge \ell_W. \end{cases} \quad (4.25)$$

The left limit at withdrawal may be positive because withdrawal is a regime change, not the action $u = 0$ within the active program.

In state R, strict concavity and opposite endpoint derivatives produce a unique interior starter. If the upper action is worse than no support, a finite safe window exists: some positive support is valuable, but sufficiently intense support is not.

Corollary 4.14 (Four-region policy)**.** *Under the interior-starter conditions, either crossing route in Theorem 4.8, and the plateau/taper conditions above, an optimal stationary policy has four state regions:*

$J = R$	*interior positive action*	*Starter,*
$J = A,\ \ell \le \ell_T$	*upper positive action*	*Consolidation,*
$J = A,\ \ell_T < \ell < \ell_W$	*declining positive action*	*Taper,*
$J = A,\ \ell \ge \ell_W$	*stop*	*Withdrawal.*

The labels are outputs of the solved policy problem. Along a path with entry, no intervening re-closure, increasing ℓ, and visitation of the taper interval, they appear chronologically in the same order. Otherwise phases may collapse or be revisited.

The exact numerical ℓ_T is a maintained-domain diagnostic. Enlarging the feasible action interval can eliminate a ceiling plateau, while coarsening the action grid can postpone the first observed drop. Monotone state taper is the robust policy-shape claim; the reported taper-onset grid point is action-resolution and domain dependent.

4.6 Type uncertainty and the trial option

The complete-information benchmark treats the persistent response type θ as known. To isolate one informational extension without turning the baseline into a POMDP, suppose (J, ℓ) remains observed while θ is unknown under prior μ. A one-episode starter u produces a current payoff, a physical post-trial state, and a signal Z. After observing Z, the planner chooses from a fixed continuation menu Π_T that includes withdrawal and every policy that could have been precommitted. This *decision reversibility* means that information may be ignored; it does not make the physical trial costless or reversible.

Let $\mathcal{B}$ be the best no-trial value, $\mathcal{C}(u)$ the value of conducting the physical trial while forcing the continuation decision to ignore its signal, and $\mathcal{T}(u)$ the value when continuation may adapt to Z. Define

$$\mathcal{I}(u) := \mathcal{T}(u) - \mathcal{C}(u), \qquad \mathcal{O}(u) := \mathcal{T}(u) - \mathcal{B}. \tag{4.26}$$

Theorem 4.15 (Pure information value and net trial option)**.** *Under decision reversibility and integrability,*

$$\mathcal{I}(u) \ge 0, \qquad \mathcal{O}(u) = \{\mathcal{C}(u) - \mathcal{B}\} + \mathcal{I}(u). \tag{4.27}$$

The inequality is strict when positive-probability signal realizations select different continuation policies. The trial is chosen only if its information premium plus its physical value is positive.

Proof. For every signal realization, the conditional maximum over Π_T weakly exceeds the conditional value of any fixed precommitted policy. Take expectations and cancel the common physical trial payoff. The decomposition is algebraic. Full strictness, Blackwell, and binary-type proofs are in Appendix F. □

Thus $\mathcal{I} > 0$ does not imply $\mathcal{O} > 0$. A statistically useful trial may remain economically undesirable because of resource cost, delay, fusion exposure, or evaluative capture. Conversely, information can create a bounded starter that the signal-ignored problem would reject when posterior realizations cross a continuation boundary. This is an additional reason to start small, not a fifth phase appended to the complete-information policy.

5 Micro-founded Numerical Analysis

5.1 Purpose and evidence status

The analytical results separate the within-episode transition technology, the cross-episode learning law, and the individual objective. The numerical exercise reconnects them in that order. It asks whether one internally consistent parameterization can jointly produce favorable and adverse learning, positive consolidation, a finite stopping boundary, a state taper, and a trial option. It does not estimate an intervention effect or select a real-world policy. Every reported number is a normalized theoretical stress test.

The computation has three layers. First, it integrates the inherited fast dynamics and validates the direction of entry, fusion, re-closure, and qualified-experience summaries. Second, it constructs the full joint law of (J^+, ℓ^+) rather than substituting a mean next state. Third, it assigns normalized mission weights, solves the Bellman–stopping system, and computes policy and management-accounting objects. The reproduction code supplied with this manuscript implements all three layers and the separate audit tests in Appendix G.

5.2 Source transition technology

For each economic episode of length $H = 12$, legibility ℓ and support u are held fixed while the fast state (q, v, g, κ) follows[2]

$$\dot{q} = a_q\{1 - e^{-\alpha u}\}(1 - q) + \rho v q(1 - q) - (d_q + b_q g)q, \tag{5.1}$$
$$\dot{v} = \{a_{v0} + a_{v1}\ell\}q(1 - v) - d_v v, \tag{5.2}$$
$$\dot{g} = a_g u - d_g g, \tag{5.3}$$
$$\dot{\kappa} = a_\kappa q v(1 - \kappa) - (d_\kappa + b_\kappa g)\kappa. \tag{5.4}$$

The cause-specific entry, fusion, and re-closure intensities are

$$\lambda_A(t; \ell, u) = \lambda_0 \exp\{\theta_q q + \theta_v v + \theta_\kappa \kappa - \theta_g g\}\kappa, \tag{5.5}$$
$$\nu_F(t; \ell, u) = \nu_0 \exp\{\phi_g g - \phi_v v - \phi_\kappa \kappa\}, \tag{5.6}$$
$$\mu_R(t; \ell, u) = \mu_0 \exp\{\psi_g g - \psi_v v - \psi_\kappa \kappa\}. \tag{5.7}$$

The visibility coefficient is $a_v(\ell) = 0.50 + 0.40\ell$. The associated analytical fold is

$$a_v^* = 0.59504, \qquad \ell^\dagger = 0.237603. \tag{5.8}$$

The unresolved-state fast path begins at $(0, 0, 0, 0)$ and the post-entry path at $(0.18, 0.25, 0.05, 0.12)$. These are common theoretical initial conditions, not estimates.

Qualified experience is the Markov reward

$$I_A = \int_0^H \mathbf{1}_{\{S(t)=A\}} v(t)\kappa(t)\, dt, \qquad Y = 1 - e^{-I_A}, \tag{5.9}$$

[2]The economic episode length $H = 12$, the support-intensity scale $u \in [0, 0.25]$, and the post-entry initial fast state deliberately differ from the closed-loop illustration in the companion paper (episode length 40, support 0.8, common initial state $(0, 0, 0, 0)$). The two exercises share the near-critical visibility mapping $a_v(\ell) = 0.50 + 0.40\,\ell$ but stress different layers of the same architecture—the learning loop there, the economic decision problem here—and no cross-paper numerical identity is implied.

and evaluative capture is

$$G = 1 - \exp\left\{-0.0018 \int_0^H g(t)\, dt\right\}. \tag{5.10}$$

The baseline cross-episode update uses $(\eta, \delta_\ell, \chi) = (0.22, 0.02, 0.70)$ in (3.5).

5.3 Joint-kernel construction and economic layer

Fast paths are integrated on 51 equally spaced legibility values and 21 diagnostic support values. Over intervals of length $\Delta t = 0.10$, the competing-risk probabilities are generated from the cause-specific intensities, and probability mass is propagated jointly over the terminal coarse state and an exposure grid for I_A. Each exposure bin is mapped through the slow update and placed on the economic legibility grid by mass-preserving linear interpolation. The resulting object is the joint kernel

$$\mathsf{K}_J(d\ell', J' \mid \ell, u), \qquad J, J' \in \{R, A\}, \tag{5.11}$$

used in the Bellman equations.

The economic action sets are

$$\mathcal{U}_R = \{0, 0.025, \ldots, 0.25\}, \qquad \mathcal{U}_A = \{0, 0.025, \ldots, 0.15\}. \tag{5.12}$$

They are maintained safe domains, not an ordering assumption. In particular, $\mathcal{U}_R \not\subseteq \mathcal{U}_A$, which is why the ranking clause of Assumption 4.4 is maintained at the operator level rather than derived from the primitive inclusion route of Appendix A; its conclusion is realized on the solved grid with minimum gap $V_A - V_R = 1.015$.

The asymmetric ceilings are substantive rather than incidental: creating first contact from the unresolved state can require intensities that would be counterproductive after entry, where the legibility–support interaction $c_\ell \ell u$ and capture exposure make high doses unattractive, and the lower post-entry ceiling encodes that maintained safe region. A matched-ceiling variant would repair only the action-set inclusion required by the primitive route to the ranking clause; the cross-state payoff and kernel-dominance components need their own audit, which Appendix G now reports on the common actions: the dominance holds at every grid point up to and including the first withdrawal point $\widehat{\ell}_W$ and fails only at high-legibility points where withdrawal is already strictly optimal. The variant remains listed among the extensions in Section 7.

The numerical one-episode rewards are

$$r_R(\ell, u) = b_E p_A^R + b_{YR}\mathbb{E}[Y \mid R] - d_F p_F^R - d_U(1 - p_A^R) - C_R(u) - d_G G_R(\ell, u), \tag{5.13}$$

$$r_A(\ell, u) = b_{YA}\mathbb{E}[Y \mid A] - d_R r_H^A - C_A(\ell, u) - d_G G_A(\ell, u), \tag{5.14}$$

with

$$C_R(u) = c_{R0}\mathbf{1}_{\{u>0\}} + c_1 u + c_2 u^2, \tag{5.15}$$

$$C_A(\ell, u) = c_{A0} + c_1 u + c_2 u^2 + c_\ell \ell u. \tag{5.16}$$

The b_{YR} and b_{YA} terms price the current service value of qualified engagement; the same realized Y separately drives the slow update that produces ℓ^+, and the two channels are priced at different dates, as recorded in Section 3.6.

Withdrawal value is

$$W(\ell) = \frac{w_0 + b_P P_{\text{persist}}(\ell) - d_R r_0^A(\ell)}{1-\beta}. \tag{5.17}$$

The persistence curve propagates 10,000 fixed-seed illustrative withdrawal states without support. It converts the structural basin into a value-function input; it is not a population estimate.

Two consistency notes attach to (5.17). First, its weights price the same functional categories as the active-regime reward: w_0 prices baseline unsupported functioning, $b_P P_{\text{persist}}(\ell)$ prices persistent qualified functioning without support—the autonomous analogue of the b_Y flows—and d_R is the same re-closure damage weight used in (5.14), so exit is not made attractive by construction through a unit change. Constructing W instead as the fixed-policy value of the reward function (5.14) at $u = 0$ under the unsupported kernel, with the overhead removed, is an equivalent-units refinement listed among the extensions in Section 7. Second, if persistence were certified by basin membership, P_{persist} and hence W could jump upward at $\ell^\dagger$, because the positive basin appears with positive extent at a saddle–node fold. None of the stopping conclusions rests on continuity there: the unique sign change of Δ_A lies in $[0.62, 0.64]$, where W varies smoothly on the disclosed grid, and the fold-discontinuity paragraph in Appendix C records the corresponding weakening of the continuity hypotheses.

Table 3: Baseline economic weights and learning parameters

Parameter	Value	Role	Parameter	Value	Role
β	0.94	episode discount factor	b_E	1.60	qualified-entry weight
b_{YR}	0.35	qualified-experience flow in R	b_{YA}	1.30	qualified-experience flow in A
d_F	1.00	fusion damage	d_R	1.00	re-closure damage
d_U	0.12	unresolved-episode burden	d_G	1.00	current pressure weight
c_{R0}	0.035	starter fixed cost	c_{A0}	0.10	active-program overhead
c_1, c_2	0.04, 0.10	variable support costs	c_ℓ	0.80	legibility–support interaction
w_0	0.45	autonomous baseline	b_P	1.80	persistence weight
η, δ_ℓ, χ	0.22, 0.02, 0.70	baseline learning type			

Notes. All entries are normalized stress-test values. They are neither monetary valuations nor empirical estimates.

5.4 Source validation

Table 4 and Figure 2 show the first audit layer at $u = 0.20$. Holding support and initial fast states fixed, higher legibility raises the probability that attention entry wins the first-resolution competition, lowers fusion, lowers matched-state re-closure, and raises expected qualified experience in both economic states. The structural fold is not used to impose a discontinuity on these curves.

Table 4: Selected transition summaries at support intensity $u = 0.20$

ℓ	p_A^R	p_F^R	r_H^A	$\mathbb{E}[Y \mid R]$	$\mathbb{E}[Y \mid A]$
0.00	0.346	0.530	0.491	0.171	0.701
0.24	0.414	0.493	0.471	0.240	0.768
0.60	0.492	0.448	0.447	0.336	0.837
1.00	0.554	0.410	0.427	0.423	0.884

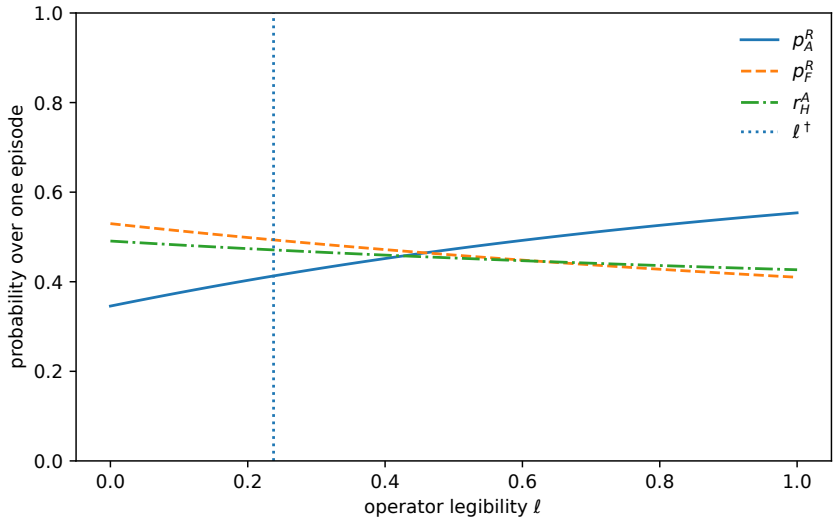

Figure 2: Layer-1 source validation at $u = 0.20$. Higher operator legibility raises first attention entry and lowers first fusion and matched-state re-closure.

5.5 Learning value, withdrawal, and taper

The Bellman equations are solved by value iteration to a sup-norm tolerance of 10^{-10}. The baseline converges in 68 iterations. The clamped-legibility economy preserves the same current episode technology and rewards but sets $\ell^+ = \ell$; it converges in 358 iterations.

Figure 3 compares the optimized post-entry continuation branches with W. The learned economy first selects withdrawal at

$$\widehat{\ell}_W = 0.64, \qquad \Delta_A(0.62) > 0 \geq \Delta_A(0.64), \tag{5.18}$$

so $[0.62, 0.64]$ is the grid sign bracket. The clamped-legibility economy withdraws at $\widehat{\ell}_W^F = 0.24$. The learned economy therefore uses a higher state threshold, yet from the initial state it reaches a finite exit because legibility moves. The clamped state remains below its lower threshold and the active relation remains open-ended.

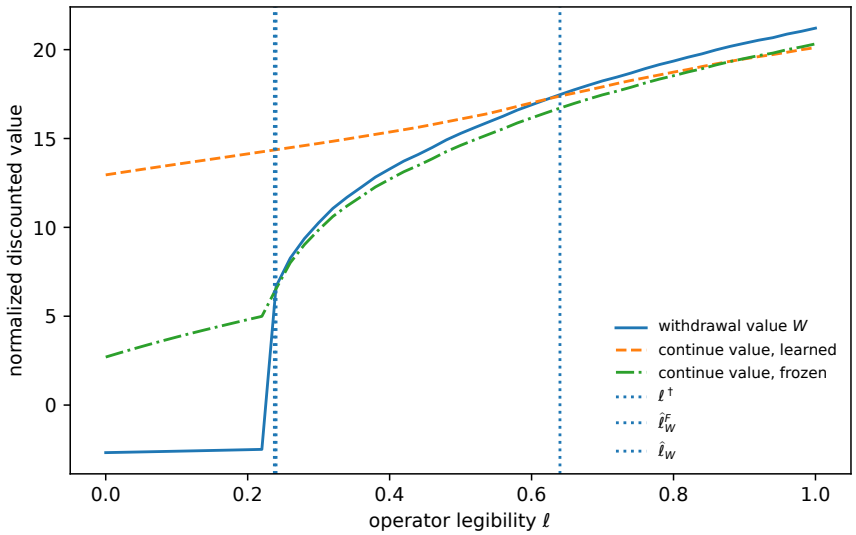

Figure 3: Withdrawal value and optimized continuation branches. The learned economy withdraws at the grid point 0.64, while the clamped-legibility economy withdraws at 0.24.

At $(R, \ell_0 = 0.06)$, the policy-fixed technology effect under the clamped-optimal rule is

$$\mathrm{LD}_R^{\pi_F^*}(0.06) = 5.786428. \tag{5.19}$$

Policy adaptation contributes

$$AP_R(0.06) = 3.051756, \tag{5.20}$$

so

$$\mathrm{LD}_R^*(0.06) = 8.838184, \qquad V_R(0.06) = 11.662283. \tag{5.21}$$

The decomposition is theoretical bookkeeping: approximately two thirds of the value difference comes from the transition technology under the fixed counterfactual rule and one third from adapting the policy to that technology.

Figure 4 shows the policy. The starter at $\ell_0 = 0.06$ is $u_R^* = 0.20$. Post-entry support remains at its numerical ceiling 0.15 through $\ell = 0.24$, first drops at

$$\widehat{\ell}_T = 0.26, \qquad \widehat{\ell}_T \in [0.24, 0.26], \tag{5.22}$$

then falls to 0.125 and 0.10 before withdrawal. The bracket records the interval between the last ceiling action and the first lower action. It is an action-grid and domain diagnostic, not a technology invariant: doubling the action step moves the first observed drop to 0.46 even though the selected policy remains nonincreasing and $\widehat{\ell}_W$ remains 0.64.

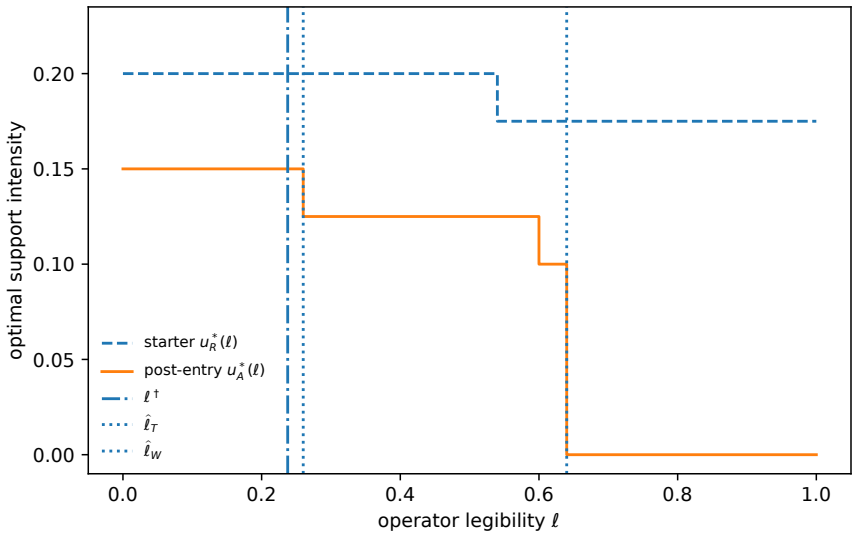

Figure 4: Optimal starter and post-entry support. The reported taper-onset and withdrawal values are finite-grid diagnostics with brackets $[0.24, 0.26]$ and $[0.62, 0.64]$.

The baseline ordering is therefore

$$\ell^\dagger = 0.237603 < \widehat{\ell}_T = 0.26 < \widehat{\ell}_W = 0.64. \tag{5.23}$$

This ordering is produced by the chosen economic weights and action domain, not imposed by the transition geometry.

5.6 Theorem–numerics audit

The whole-grid audit matters because a monotone selected policy does not certify every sufficient theorem used in the analytical development. Figure 5 displays the gap together with the selected post-entry action. The gap has eleven local increases at low legibility, with a largest increment of 0.1076, and becomes strictly decreasing only from $\ell = 0.22$ onward. It nevertheless has one sign change and an upper stopping interval. The numerical threshold is therefore supported by the eventual-crossing theorem, not by global strict decrease.

The low-legibility increases have an economic reading rather than a numerical one. Below $\ell^\dagger$ the autonomous branch contains no persistence geometry, so W is low and flat over the relevant range, while the continuation branch prices both current functioning and the learning it generates; as legibility approaches the fold from below, continuation appropriates the approaching persistence opportunity and the gap widens. Once the fold is crossed, the autonomous value begins to absorb the same opportunity and the gap turns decisively downward. Before structural feasibility, learning therefore pushes economic withdrawal further away—an implication of the model, not an anomaly of the grid. Because the unique sign change lies in $[0.62, 0.64]$, above $\ell^\dagger$, Proposition 4.9 applies with $\mathfrak{M}_\dagger = \Delta_A(\ell^\dagger) > 0$: the baseline sits strictly inside the $\ell_W > \ell^\dagger$ regime.

Similarly, 32 of 192 adjacent global decreasing-differences checks fail. By contrast, every one of the 21 pairwise action comparisons satisfies the weak single-crossing clause (i) of Assumption 4.12, and the reported post-entry policy—the minimum optimal selection, with ties broken toward the smaller action—has no upward step. The computation therefore instantiates the minimum-selection clause of the finite-action

theorem in Section 4.5; it is not reported as a verification of global decreasing differences or of the strict clause.

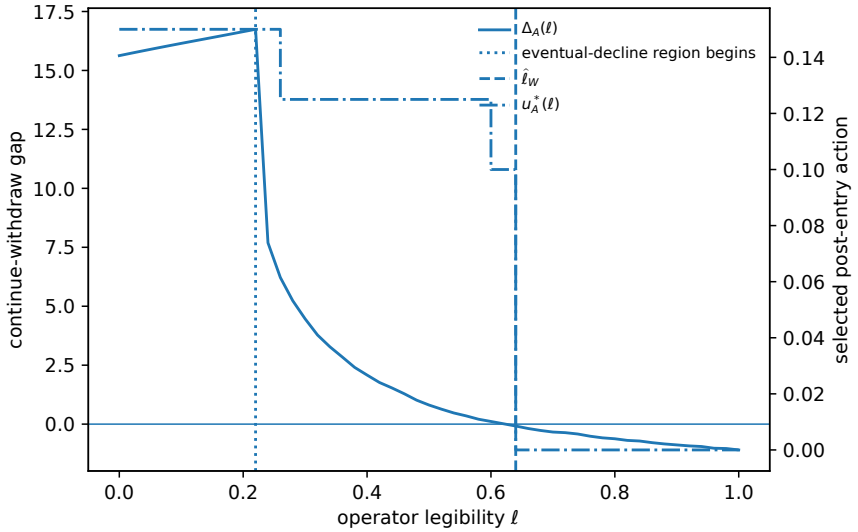

Figure 5: Continue–withdraw gap and selected action. The gap is eventually decreasing and crosses once, while the selected action is nonincreasing.

The audit also verifies probability mass to machine precision, inherited event order on the diagnostic grid, nondecreasing solved values, and $V_A \geq V_R$ in the chosen parameterization. These solved conclusions do not retroactively verify the strongest optional primitive route, whose action-set inclusion condition fails.

5.7 Signed learning under bounded and projected adverse cases

Figure 6 and Table 5 vary only the slow formation and capture parameters. Two bounded cases satisfy the global no-overshoot condition $\eta + \delta_\ell + \chi \leq 1$ and still produce negative optimized Learning Dividends. The first reverses sign at $\ell_0 = 0.38$ and withdraws at 0.36; the second reverses at 0.24 and withdraws at 0.24. Thus the signed counter-result does not depend on clipping.

The earlier high-capture specification $(0.05, 0.02, 12.0)$ is retained only as a *projected extreme stress case*: the code projects ℓ^+ to $[0, 1]$, and the primitive no-overshoot condition is not satisfied. Its negative value is illustrative but not needed for the theoretical counterexample.

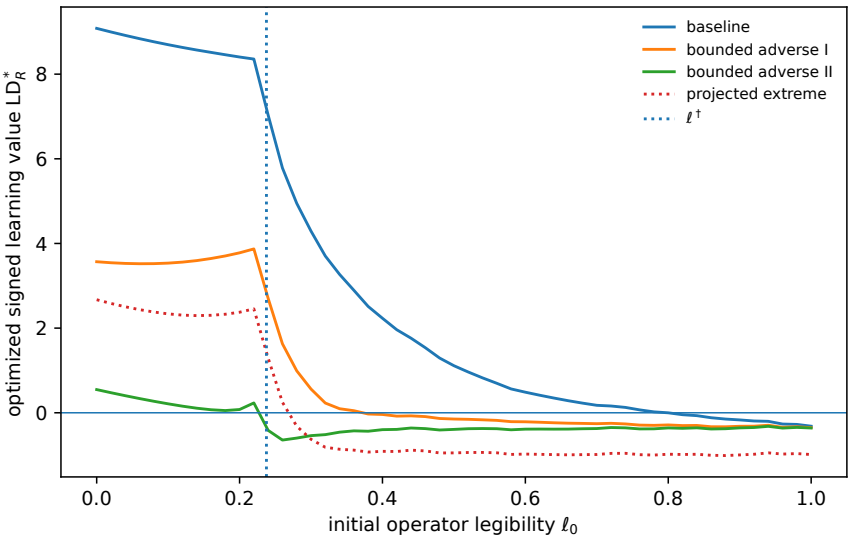

Figure 6: Optimized signed Learning Dividends for the baseline, two bounded adverse cases, and a separately labeled projected extreme.

Table 5: Learning/capture cases and numerical sign reversals

Case	η	δ_ℓ	χ	First negative ℓ_0	$\widehat{\ell}_W$
Baseline	0.22	0.02	0.70	0.80	0.64
Bounded adverse I	0.05	0.02	0.90	0.38	0.36
Bounded adverse II	0.01	0.02	0.97	0.24	0.24
Projected extreme	0.05	0.02	12.00	0.28	0.30

5.8 Individual accounting and an information-created trial

For the baseline accounting exercise, withdrawal is classified as durable when

$$\ell \geq \ell^\dagger, \qquad P_{\text{persist}}(\ell) \geq 0.55, \qquad r_0^A(\ell) \leq 0.58. \tag{5.24}$$

These are illustrative certification cutoffs. Table 6 separates cheap exit from durable exit. The learning-optimal policy uses a temporary active relation, reaches durable withdrawal with probability essentially one, and has CDI $= 0.659$. Withdrawal at the first A state is inexpensive in the numerator but almost never durable, so its cost per durable unit is enormous. The clamped-legibility optimum spends resources without reaching a finite withdrawal time from the initial state.

Table 6: Policy accounting from $(R, \ell_0 = 0.06)$

Policy	Initial value	Durable withdrawal probability	Expected active episodes	Discounted resource cost	CDI
Learning-optimal	11.662	1.000	9.79	0.659	0.659
Clamped-legibility optimum	2.824	0.000	∞	1.356	–
Withdraw on first A	-1.290	< 0.001	3.16	0.132	$> 25{,}000$

The trial calculation uses two persistent types, prior probability $p = 0.20$ of the

high-learning type, and a one-episode starter $u = 0.10$ at $(R, 0.30)$. The signal satisfies

$$\Pr(g \mid H) = 0.4073, \qquad \Pr(g \mid L) = 0.0207, \tag{5.25}$$

which yields posteriors $\Pr(H \mid g) = 0.8310$ and $\Pr(H \mid b) = 0.1314$. A good signal selects the long-learning policy; a bad signal selects the conservative policy.

Table 7: Trial values at prior probability $p = 0.20$

Object	Normalized value
No-trial benchmark $\mathcal{B}$	7.6012
Signal-ignored trial $\mathcal{C}$	7.5103
Adaptive trial $\mathcal{T}$	7.6128
Information premium $\mathcal{I} = \mathcal{T} - \mathcal{C}$	0.1025
Net trial option $\mathcal{O} = \mathcal{T} - \mathcal{B}$	0.0116

The physical trial is unattractive when its signal is ignored: $\mathcal{C} - \mathcal{B} = -0.0909$. Adaptive use of the signal contributes 0.1025, creating a small positive net option. This is an information-created starter in the precise sense of Theorem 4.15.

Two qualifications keep the trial result in proportion. The net option equals 0.15 percent of the value level and is computed at a single prior–dose point. At a degenerate prior the posterior equals the prior and, with the persistent type known and the state observed, a single continuation policy is conditionally optimal almost surely, so $\mathcal{I}(u) = 0$ by the second clause of Proposition F.2; with $\mathcal{C}(u) - \mathcal{B} < 0$ there, the net option is negative near $p \in \{0, 1\}$, and any information-created region lies in the interior of the prior interval and inside the posterior-crossing region of Proposition F.5; whether it is a single band around p^* depends on additional shape conditions recorded in the sign-structure paragraph of Appendix F. A full (p, u) map is listed among the extensions in Section 7. The likelihoods in (5.25) are computed from the trial episode's outcome distribution under the disclosed episode law; they are not free parameters.

5.9 Robustness and reproducibility

Table 8 reports matched numerical comparisons. The fine-exposure value is 8.823390, computed with both the learned and clamped economies on the fine exposure grid. The withdrawal grid point remains 0.64 in every row.

Table 8: Matched numerical convergence checks

Specification	Action step	Exposure step	$(\widehat{\ell}_W, \mathrm{LD}^*_R(0.06))$
Baseline	0.025	0.050	(0.64, 8.838184)
Coarse economic action grid	0.050	0.050	(0.64, 8.824599)
Fine exposure grid, matched	0.025	0.025	(0.64, 8.823390)

Positive rescaling of every payoff coefficient by factors from 0.1 to 10 leaves both learned and clamped policy arrays, $\widehat{\ell}_T$, and $\widehat{\ell}_W$ unchanged; values and resource ledgers scale proportionally. Additive shifts are not policy neutral because policies differ in active duration.

Table 9 gives a compact mission-weight check. The threshold and duration move in economically sensible but not mechanically predetermined ways. In particular,

greater patience raises the state threshold in this parameterization, while a larger active overhead lowers it. The numerical exercise illustrates gap-shift logic; it does not establish universal comparative statics.

Table 9: Selected mission-weight and discount robustness

Scenario	$\widehat{\ell}_T$	$\widehat{\ell}_W$	$\mathrm{LD}^*_R(0.06)$	Active episodes	CDI
Baseline	0.26	0.64	8.838	9.79	0.659
Post-entry flow low	0.26	0.58	10.327	8.73	0.579
Post-entry flow high	0.26	0.72	7.546	11.73	0.797
Active overhead high	0.28	0.60	9.640	9.07	1.042
$\beta = 0.90$	0.24	0.54	3.946	8.04	0.571
$\beta = 0.97$	0.26	0.74	23.055	12.30	1.012

Finally, the CDI denominator is substantive. A stricter illustrative durability rule first certifies at $\ell = 0.96$, above the baseline withdrawal boundary; durable probability is then zero and CDI is undefined, not zero. Appendix G reports the complete scale, weight, learning, grid, threshold, and denominator audits. The accompanying code rebuilds the model without external data and records hashes for the reproduction materials.

6 Economic and Management-Accounting Implications

6.1 Decision-accounting boundary

The model changes the accounting question. A one-period evaluation asks whether current support produces enough current benefit to cover current cost. The present model asks whether the same support also changes the law governing later entry, fusion, re-closure, and eventual independence. Once that channel is admitted, activity, success, and cost cannot be read from one episode or one scorecard.

Principle 6.1 (Economic investment is not accounting recognition)**.** *Calling part of support an investment is an economic statement: current expenditure can alter future transition conditions and create future functional value. It is not a claim that operator legibility, attention, or the Learning Dividend satisfies financial-reporting recognition criteria. The model supplies shadow values for planning and control, not a balance-sheet asset.*

Normalized cardinal units allow dynamic alternatives to be ranked before monetary calibration. Monetary cost–benefit analysis is a later empirical exercise. The contribution here is to specify the relevant decision objects, cost boundaries, denominators, and timing errors before asking how they should be monetized.

6.2 Current service and transition investment

For a fixed policy, the learned and clamped technologies have the same current episode and current payoff accounting. Their difference

$$\mathrm{LD}^{\pi}_{j}(\ell) = V^{\pi}_{\mathrm{L},j}(\ell) - V^{\pi}_{\mathrm{F},j}(\ell) \tag{6.1}$$

is therefore the future transition-law component of value.

Accounting Identity 6.2 (One-shot evaluation error)**.** *If a one-shot scorecard reports* $V^{\pi}_{\mathrm{F},j}$ *when the relevant technology is* $V^{\pi}_{\mathrm{L},j}$*, its valuation error is*

$$\mathcal{E}^{\pi}_{\mathrm{OS},j}(\ell) := V^{\pi}_{\mathrm{F},j}(\ell) - V^{\pi}_{\mathrm{L},j}(\ell) = -\mathrm{LD}^{\pi}_{j}(\ell). \tag{6.2}$$

Positive learning makes the one-shot scorecard conservative; adverse learning makes it optimistic.

The same expenditure can therefore have a current-service consequence and a transition-investment consequence. Current contact, entry, fusion avoidance, maintained functioning, and current pressure belong to the service account. The effect of the episode on ℓ^{+} belongs to the transition account. These are not two cash payments; they are two value consequences of one payment.

Where the value and transition map are differentiable, the local shadow return on transition investment contains

$$\mathcal{M}_{\ell,j}(\ell, u) := \beta \mathbb{E}\left[\frac{\partial V^{\pi}_{J^{+}}(\ell^{+})}{\partial \ell}\frac{\partial \mathcal{T}_{j}(\ell, u, \varepsilon)}{\partial u}\right]. \tag{6.3}$$

It is positive when marginal support improves value-relevant future legibility, zero when the future law is unchanged, and negative when evaluative capture damages that law.

6.3 Outcome chain and denominator discipline

Management systems tend to privilege what is observed earliest and most cheaply: tokens, completed episodes, or first entry. The model supplies a longer outcome chain.

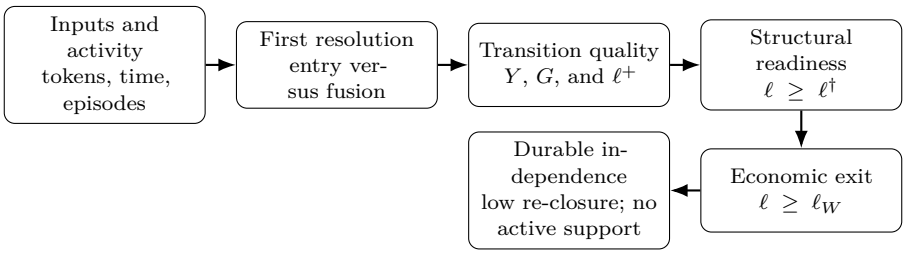

Figure 7: Outcome chain for individual transition accounting. Activity and first entry are leading indicators; durable independence is the terminal mission outcome.

The layers should not be collapsed. First entry can occur at low legibility and be followed by rapid re-closure. Positive persistence geometry can become available without making withdrawal economically optimal. Withdrawal can occur before a chosen durability standard is met. A first-entry rate is therefore a process indicator, not a complete performance measure. A durability-adjusted handoff measure is

$$HA^{\pi}(H) := \overset{\pi}{\Pr}\{\tau_A < \infty,\ T_R - \tau_A > H\}, \tag{6.4}$$

where τ_A is qualified entry and T_R the subsequent first re-closure time. Two programs with the same entry rate can have different $HA(H)$ because their post-entry transition laws differ.

6.4 Threshold accounting and timing losses

The three state thresholds and the durability endpoint answer different questions.

Object	Control margin	Interpretation
$\ell^\dagger$	Structural readiness	Positive unsupported-persistence geometry becomes available; no economic action is implied by geometry alone.
ℓ_T	Intensive margin	The optimal positive action leaves the consolidation plateau and begins to taper.
ℓ_W	Extensive margin	The optimized active branch and autonomous withdrawal value are equal.
τ_G	Ex post performance	A disclosed composite durability standard is first met.

The threshold gap

$$\Gamma_W := \ell_W - \ell^\dagger \tag{6.5}$$

is a management-accounting diagnostic. A positive gap means that persistence has become structurally feasible before active support loses economic value; a negative gap means that the active relation is too costly or damaging to justify waiting for the structural boundary. The same information is carried by $\mathfrak{M}_\dagger = \Delta_A(\ell^\dagger)$.

The continue–withdraw gap also prices timing errors. For an observed post-entry action a, define

$$L_{\text{time}}(\ell, a) := \mathbf{1}_{\{a=\mathsf{W}\}}[\Delta_A(\ell)]_+ + \mathbf{1}_{\{a\neq\mathsf{W}\}}[-\Delta_A(\ell)]_+. \tag{6.6}$$

Withdrawing where $\Delta_A > 0$ creates a premature-withdrawal loss; remaining active where $\Delta_A < 0$ creates an over-retention loss. Conditional on continuation, nonoptimal intensity creates

$$L_{\text{int}}(\ell, u) := Q_A^*(\ell) - Q_A(\ell, u) \geq 0. \tag{6.7}$$

Thus a wrong decision to remain active and a wrong choice of intensity are distinct control failures.

For an observed policy π, discounted benchmark-deviation accounts can be defined as

$$PW^\pi := \mathbb{E}^\pi\left[\sum_{n\geq 0} \beta^n \mathbf{1}_{\{J_n=A, a_n=\mathsf{W}\}}[\Delta_A(\ell_n)]_+\right], \tag{6.8}$$

$$OR^\pi := \mathbb{E}^\pi\left[\sum_{n\geq 0} \beta^n \mathbf{1}_{\{J_n=A, a_n\neq\mathsf{W}\}}[-\Delta_A(\ell_n)]_+\right]. \tag{6.9}$$

These are first-best deviation measures. Under partial observation they would become expected losses relative to a belief-state benchmark.

6.5 Cost to Durable Independence

Let D^π indicate that withdrawal satisfies the adopted durability criterion, and let discounted resource cost be

$$C^\pi := \mathbb{E}^\pi\left[\sum_{n=0}^{\tau_W - 1} \beta^n C_{J_n}(\ell_n, u_n)\right]. \tag{6.10}$$

The Cost to Durable Independence is

$$\mathrm{CDI}^{\pi} := \frac{C^{\pi}}{\Pr^{\pi}(D^{\pi} = 1)}, \tag{6.11}$$

when the denominator is positive. It is undefined, not zero, when durable withdrawal never occurs.

The denominator is the substantive part of the measure. A low numerator may reflect efficient support or premature abandonment. A high numerator may reflect wasteful retention or a temporary investment that eliminates future support. In the numerical stress test, withdrawal on first entry spends only 0.132 in discounted resource cost but has durable probability below 0.001; the resulting CDI exceeds 25,000. The learning-optimal policy spends 0.659, reaches durable withdrawal with probability essentially one, and has CDI = 0.659. The figures are not forecasts, but they show why neither low cost nor high activity is a sufficient performance criterion. Across the disclosed experiments the denominator spans its three regimes: near-certain durability in the baseline (CDI = 0.659), a small interior durable probability in the lower-formation environment of Appendix G (CDI $\approx$ 47.8), and an undefined ratio when certification is unreachable, as under the stricter illustrative durability rule. The measure is informative precisely because it separates these regimes rather than collapsing them.

A related break-even object can be defined without pricing attention. Let $\mathcal{D}^{\pi}$ be expected durable-transition units and K^{π} the full specified resource, pressure, and mission cost. If b is the normalized mission value per durable unit,

$$NMV^{\pi}(b) = b\mathcal{D}^{\pi} - K^{\pi}, \qquad b^{*,\pi} = \frac{K^{\pi}}{\mathcal{D}^{\pi}}. \tag{6.12}$$

External evidence may later calibrate b; the decision architecture need not wait for that calibration.

6.6 Activity bias and self-liquidating control

Tokens issued, staff time, measurement frequency, and continued enrollment are visible activities. Future support avoided is less visible because it appears as an absence. An activity-based scorecard can therefore reward keeping the relationship active even when the mission is to make it safely unnecessary.

The model does not imply that fewer episodes are always better: immediate withdrawal can destroy value. Nor does it imply that more episodes are better: a clamped transition law can absorb resources without reaching durable exit. Each additional episode should be judged by the continue–withdraw margin, and its intensity by the continuation action gap.

Definition 6.3 (Individual self-liquidating control)**.** *A support policy is individually self-liquidating from a state if it supplies positive support when the active relation has positive incremental value, reduces support as its marginal value declines, and reaches a state at which withdrawal is optimal with positive probability. Success is not defined as indefinite retention.*

Self-liquidating does not mean self-financing. It means that the control relation is designed to remove its own economic reason for continuing. Financial solvency and provider incentives belong downstream.

Three linked ledgers are needed:

- a *resource ledger* for support, staff time, platform cost, monitoring, fixed overhead, and participant burden;
- a *transition ledger* for entry, fusion, re-closure, qualified experience, evaluative capture, and legibility;
- an *exit ledger* for taper onset, structural readiness, economic withdrawal, durable certification, later re-entry, and support avoided.

A resource ledger without the transition ledger treats learning value as current expense only. A transition ledger without the exit ledger can turn improvement into an excuse for permanent supervision. An exit ledger without the earlier two cannot distinguish efficient independence from premature abandonment.

6.7 Visible evaluation and the cost of measurement

Evaluative pressure has two incidences. Current payoff includes a term such as $-d_G G_j(\ell, u)$. The same episode may also lower future legibility through $-\chi G\ell$ in the transition law. These are not duplicate charges: one is a current burden and the other is the future value lost through a changed transition state.

If participant-facing measurement intensity z raises capture, a local future shadow cost contains

$$\mathcal{C}_G^{\text{future}} = \beta \mathbb{E}\left[\frac{\partial V_{J^+}(\ell^+)}{\partial \ell}\chi\ell\frac{\partial G}{\partial z}\right]. \tag{6.13}$$

The full marginal cost of visible evaluation therefore includes resource cost, current burden, and the shadow value of the damaged transition law.

Principle 6.4 (Measurement neutrality is substantive)**.** *A more informative signal can be ranked by pure information value only when changing the signal leaves current payoff and the physical transition kernel unchanged. Participant-facing scoring, ranking, or alerts that change evaluative capture are interventions and must be evaluated by full net value.*

The practical implication is not a universal instruction to measure less. It is to distinguish backend information used by the planner from evaluation displayed to the participant. The former may improve decisions without changing the person; the latter may alter the transition being measured.

6.8 Experimentation and phase-specific accountability

Trial accounting should keep three objects separate:

$$\underbrace{\mathcal{O}}_{\text{net trial option}} = \underbrace{(\mathcal{C} - \mathcal{B})}_{\text{physical trial value}} + \underbrace{\mathcal{I}}_{\text{information premium}}. \tag{6.14}$$

This prevents two opposite errors. A pilot that does not immediately improve the participant need not be wasteful if it creates decision-relevant information. But positive information value does not make a harmful physical trial desirable. Ex post continuation is also an invalid stand-alone success measure: a trial can be valuable precisely because an adverse signal supports stopping.

The four policy regions imply different control questions rather than one uniform KPI.

Table 10: Phase-specific accounting in the individual model

Phase	Primary decision	Appropriate accounting focus	Characteristic control failure
Starter	Whether and how much to initiate contact	Entry/fusion odds, safe window, trial option, access cost	No contact, or excessive support that raises pressure
Consolidation	Whether post-entry support should remain positive	Current durability plus transition-law return relative to overhead	Immediate exit that sacrifices positive learning value
Taper	How rapidly positive support should fall	Marginal support value, ℓ_T, intensity loss, re-closure risk	Fixed support after marginal value has declined
Withdrawal	Whether the active relation should end	Δ_A, ℓ_W, durability certification, later re-entry	Premature withdrawal or over-retention

A high opening rate can indicate successful access while saying nothing about re-closure. Low support can indicate efficient taper or neglect. A low caseload can indicate independence or exclusion. Phase-specific interpretation is part of the control design.

6.9 Downstream life-cycle interface

The individual paper ends before population aggregation. For type θ, its core output is

$$\Omega(\theta) := \{\pi^*(\theta), C(\theta), U(\theta), T(\theta), r(\theta), B(\theta)\}, \tag{6.15}$$

where the entries record the optimal policy, discounted resource cost, support issuance, active duration, re-closure or return risk, and normalized functional benefit. The diagnostic vector is

$$\Xi(\theta) := \{\mathrm{LD}^*(\theta), \ell_T(\theta), \ell_W(\theta), \Pr(D = 1 \mid \theta), \mathrm{CDI}(\theta), \mathcal{O}(\theta)\}. \tag{6.16}$$

A population paper can aggregate $\Omega(\theta)$ over heterogeneous types and then add capacity, provider behavior, liabilities, reserves, and solvency. The diagnostics in $\Xi(\theta)$ explain why individuals with similar current service cost can generate different life-cycle burdens; they do not replace the core quantities required for aggregation.

7 Conclusion

This paper asked when an assistance relation should remove itself. The answer cannot be read from the first qualified entry, from a structural bifurcation threshold, or from the amount of support already delivered. A first entry may be too fragile to survive withdrawal; a positive post-withdrawal basin may be available before the current state is safely inside it; and continued service may remain valuable for a finite interval even after unsupported persistence becomes feasible. At the same time, support is not valuable merely because the transition is incomplete. It consumes resources, keeps an active program frame in place, and can alter observation through evaluative capture. The individual problem is therefore a dynamic investment-and-stopping problem: assistance buys current functioning and may rewrite the transition law of later episodes, but the same assistance can also damage that law. Optimality requires valuing what remains

when the aid is gone, not only what occurs while it is present.

The model closes that problem in a sequence of distinct results. A reduced joint kernel preserves the destination state and the distribution of future operator legibility without carrying the full fast dynamics through every Bellman update. The discounted Bellman system has a unique continuous solution and a stationary optimal policy under compactness, boundedness, and Feller conditions. The signed Learning Dividend identifies the value of allowing experience to rewrite the future transition law under the same contingent policy. Its performance-difference identity and channel ledger show why qualified experience, natural loss, and evaluative capture belong in one net account and why positive mean drift need not imply positive value. The one-cycle benchmark identifies the fixed-overhead hurdle for positive consolidation. The stopping problem adds a separate boundary: an upper-set crossing yields ℓ_W, and $\Delta_A(\ell^\dagger)$ determines whether the economic boundary lies below, at, or above the structural threshold. The intensive margin then yields a state taper. Decreasing differences supply a clean continuum theorem; the weak clause of pairwise action single crossing supplies the finite-action theorem that orders the minimum optimal selection reported by the implementation. With endpoint conditions, a separate ℓ_T marks the transition from a consolidation plateau to tapering. Starter–Consolidation–Taper–Withdrawal is thus a derived regime, not a schedule inserted into the state space.

The numerical analysis shows that these regions are jointly realizable within one transition architecture while also revealing the limits of the strongest sufficient conditions. The baseline produces $\ell^\dagger = 0.237603$, a first taper drop at the grid point $\widehat{\ell}_T = 0.26$ with bracket $[0.24, 0.26]$, and withdrawal at $\widehat{\ell}_W = 0.64$ with sign bracket $[0.62, 0.64]$. Clamping legibility while preserving the current episode technology moves withdrawal to 0.24 and prevents the initial state from progressing toward a finite exit. The optimized learning value of 8.838 at $(R, 0.06)$ separates into a policy-fixed technology effect and an adaptation premium. The continue–withdraw gap is not globally decreasing at low legibility, but it becomes strictly decreasing on the relevant upper region and crosses once. The solved Q array does not satisfy global adjacent decreasing differences, but every pairwise comparison satisfies the weak single-crossing clause and the reported minimum-selection action is nonincreasing. Two bounded adverse cases produce negative learning without clipping, while a separately labeled projected extreme provides an additional stress test. These findings narrow the claims rather than weaken the core result: the internally generated kernel can support temporary consolidation, tapering, finite withdrawal, adverse learning, and information-created experimentation in one disclosed environment. None of the quantities is an empirical estimate.

The management-accounting implication is a change in the unit of accountability. Tokens, episodes, staff time, first entries, and continued enrollment are observable activities, but none is the terminal mission outcome. Current support has a service component and a signed transition-investment component; the latter is omitted when later transition conditions are frozen. Structural readiness, taper onset, economic withdrawal, and durable certification are separate events and should not be reported as one milestone. Cost to Durable Independence disciplines both cheap exit and indefinite retention by placing durable withdrawal in the denominator. Timing-loss accounts distinguish premature withdrawal from over-retention, while an intensity-loss account separates the decision to remain active from the amount supplied. Visible evaluation must be charged both as a current burden and, when it alters legibility, as a future transition cost. Trial accounting must likewise distinguish physical delivery value from

information value. These objects support managerial valuation and control; they do not imply financial-statement recognition of operator legibility or attention.

The boundaries of the result are as important as the result itself. The planner is benevolent, fully aligned with one individual, and observes (J, ℓ) without error. Mission weights are normalized rather than empirically monetized. The upstream transition law requires independent identification before any real intervention is selected. Partial observation, endogenous measurement precision, residual fast-state carryover, repeated Bayesian experimentation, a matched-ceiling action domain, an autonomous value constructed directly from the active reward function at zero intensity, a prior-and-dose map of the trial option, and empirical calibration remain extensions. Provider incentives are absent by design: the benchmark identifies when continuation, tapering, and withdrawal are individually efficient, but it does not guarantee that an organization benefits from implementing those choices. Nor does the paper aggregate people or finance a system. Its downstream interface is $\Omega(\theta)$ together with the diagnostics $\Xi(\theta)$ in (6.15)–(6.16). A population paper can add heterogeneity, capacity, provider behavior, liabilities, reserves, and solvency without reopening the individual fast dynamics. The central conclusion is narrower, but stronger for being narrow: a qualified opening can create a future learning return, yet assistance also carries resource and evaluative costs; optimal support therefore values not only entry, but the transition technology that remains when support is gone.

Transparency and Reproducibility

This is a theoretical article. No empirical data are analyzed and no empirical effect is estimated. All numerical results are generated by the accompanying reproduction code from the stated transition equations, grids, normalized weights, and fixed seed. The source package contains the frozen numerical core, the main-figure generator, the robustness script, machine-readable audit tables, the complete finite-grid upper-set audit, and cryptographic checksums. Numerical agreement establishes reproducibility and internal realizability under the disclosed model; it does not establish external validity, efficacy, or calibration.

References

Alagoz, O., L. M. Maillart, A. J. Schaefer, and M. S. Roberts (2004). "The Optimal Timing of Living-Donor Liver Transplantation". *Management Science* 50.10, pp. 1420–1430. DOI: 10.1287/mnsc.1040.0287.

Arrow, K. J. (1962). "The Economic Implications of Learning by Doing". *Review of Economic Studies* 29.3, pp. 155–173. DOI: 10.2307/2295952.

Baker, G. P. (1992). "Incentive Contracts and Performance Measurement". *Journal of Political Economy* 100.3, pp. 598–614. DOI: 10.1086/261831.

Ben-Porath, Y. (1967). "The Production of Human Capital and the Life Cycle of Earnings". *Journal of Political Economy* 75.4, Part 1, pp. 352–365. DOI: 10.1086/259291.

Bénabou, R. and J. Tirole (2003). "Intrinsic and Extrinsic Motivation". *Review of Economic Studies* 70.3, pp. 489–520. DOI: 10.1111/1467-937X.00253.

Bevan, G. and C. Hood (2006). "What's Measured Is What Matters: Targets and Gaming in the English Public Health Care System". *Public Administration* 84.3, pp. 517–538. DOI: 10.1111/j.1467-9299.2006.00600.x.

Blackwell, D. (1953). "Equivalent Comparisons of Experiments". *Annals of Mathematical Statistics* 24.2, pp. 265–272. DOI: 10.1214/aoms/1177729032.

— (1965). "Discounted Dynamic Programming". *Annals of Mathematical Statistics* 36.1, pp. 226–235. DOI: 10.1214/aoms/1177700285.

Bowles, S. and S. Polania-Reyes (2012). "Economic Incentives and Social Preferences: Substitutes or Complements?" *Journal of Economic Literature* 50.2, pp. 368–425. DOI: 10.1257/jel.50.2.368.

Courty, P. and G. Marschke (2004). "An Empirical Investigation of Gaming Responses to Explicit Performance Incentives". *Journal of Labor Economics* 22.1, pp. 23–56. DOI: 10.1086/380402.

Cunha, F. and J. J. Heckman (2007). "The Technology of Skill Formation". *American Economic Review* 97.2, pp. 31–47. DOI: 10.1257/aer.97.2.31.

Cunha, F., J. J. Heckman, and S. M. Schennach (2010). "Estimating the Technology of Cognitive and Noncognitive Skill Formation". *Econometrica* 78.3, pp. 883–931. DOI: 10.3982/ECTA6551.

Deci, E. L. (1971). "Effects of Externally Mediated Rewards on Intrinsic Motivation". *Journal of Personality and Social Psychology* 18.1, pp. 105–115. DOI: 10.1037/h0030644.

Deci, E. L., R. Koestner, and R. M. Ryan (1999). "A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation". *Psychological Bulletin* 125.6, pp. 627–668. DOI: 10.1037/0033-2909.125.6.627.

Ebrahim, A., J. Battilana, and J. Mair (2014). "The Governance of Social Enterprises: Mission Drift and Accountability Challenges in Hybrid Organizations". *Research in Organizational Behavior* 34, pp. 81–100. DOI: 10.1016/j.riob.2014.09.001.

Ebrahim, A. and V. K. Rangan (2014). "What Impact? A Framework for Measuring the Scale and Scope of Social Performance". *California Management Review* 56.3, pp. 118–141. DOI: 10.1525/cmr.2014.56.3.118.

Falk, A. and M. Kosfeld (2006). "The Hidden Costs of Control". *American Economic Review* 96.5, pp. 1611–1630. DOI: 10.1257/aer.96.5.1611.

Feltham, G. A. and J. Xie (1994). "Performance Measure Congruity and Diversity in Multi-Task Principal/Agent Relations". *The Accounting Review* 69.3, pp. 429–453.

Franco-Santos, M., L. Lucianetti, and M. Bourne (2012). "Contemporary Performance Measurement Systems: A Review of Their Consequences and a Framework for Research". *Management Accounting Research* 23.2, pp. 79–119. DOI: 10.1016/j.mar.2012.04.001.

Frey, B. S. and R. Jegen (2001). "Motivation Crowding Theory". *Journal of Economic Surveys* 15.5, pp. 589–611. DOI: 10.1111/1467-6419.00150.

Gittins, J. C. (1979). "Bandit Processes and Dynamic Allocation Indices". *Journal of the Royal Statistical Society: Series B (Methodological)* 41.2, pp. 148–177. DOI: 10.1111/j.2517-6161.1979.tb01068.x.

Gneezy, U. and A. Rustichini (2000). "Pay Enough or Don't Pay at All". *Quarterly Journal of Economics* 115.3, pp. 791–810. DOI: 10.1162/003355300554917.

Holmström, B. and P. Milgrom (1991). "Multitask Principal–Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design". *Journal of Law, Economics, and Organization* 7.Special Issue, pp. 24–52. DOI: 10.1093/jleo/7.special_issue.24.

Howard, J. L., G. R. Slemp, and X. Wang (2025). "Need Support and Need Thwarting: A Meta-Analysis of Autonomy, Competence, and Relatedness Supportive and Thwarting Behaviors in Student Populations". *Personality and Social Psychology Bulletin* 51.9, pp. 1552–1573. DOI: 10.1177/01461672231225364.

Ittner, C. D. and D. F. Larcker (1998). "Innovations in Performance Measurement: Trends and Research Implications". *Journal of Management Accounting Research* 10, pp. 205–238.

Jovanovic, B. and Y. Nyarko (1996). "Learning by Doing and the Choice of Technology". *Econometrica* 64.6, pp. 1299–1310. DOI: 10.2307/2171832.

Kakade, S. and J. Langford (2002). "Approximately Optimal Approximate Reinforcement Learning". *Proceedings of the Nineteenth International Conference on Machine Learning (ICML 2002)*. San Francisco, CA: Morgan Kaufmann, pp. 267–274.

Martin, N. D., C. Dornfeld Tissenbaum, D. Gnesdilow, and S. Puntambekar (2019). "Fading Distributed Scaffolds: The Importance of Complementarity Between Teacher and Material Scaffolds". *Instructional Science* 47.1, pp. 69–98. DOI: 10.1007/s11251-018-9474-0.

McCall, J. J. (1970). "Economics of Information and Job Search". *Quarterly Journal of Economics* 84.1, pp. 113–126. DOI: 10.2307/1879403.

McDonald, R. and D. Siegel (1986). "The Value of Waiting to Invest". *Quarterly Journal of Economics* 101.4, pp. 707–727. DOI: 10.2307/1884175.

Milgrom, P. and C. Shannon (1994). "Monotone Comparative Statics". *Econometrica* 62.1, pp. 157–180. DOI: 10.2307/2951479.

Murphy, S. A. (2003). "Optimal Dynamic Treatment Regimes". *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 65.2, pp. 331–355. DOI: 10.1111/1467-9868.00389.

Nadler, A. (2002). "Inter-Group Helping Relations as Power Relations: Maintaining or Challenging Social Dominance Between Groups Through Helping". *Journal of Social Issues* 58.3, pp. 487–502. DOI: 10.1111/1540-4560.00272.

Nadler, A. and L. Chernyak-Hai (2014). "Helping Them Stay Where They Are: Status Effects on Dependency/Autonomy-Oriented Helping". *Journal of Personality and Social Psychology* 106.1, pp. 58–72. DOI: 10.1037/a0034152.

Nahum-Shani, I., S. N. Smith, B. J. Spring, L. M. Collins, K. Witkiewitz, A. Tewari, and S. A. Murphy (2018). "Just-in-Time Adaptive Interventions (JITAIs) in Mobile Health: Key Components and Design Principles for Ongoing Health Behavior Support". *Annals of Behavioral Medicine* 52.6, pp. 446–462. DOI: 10.1007/s12160-016-9830-8.

Patzak, A. and X. Zhang (2025). "Blending Teacher Autonomy Support and Provision of Structure in the Classroom for Optimal Motivation: A Systematic Review and Meta-Analysis". *Educational Psychology Review* 37, p. 17. DOI: 10.1007/s10648-025-09994-2.

Puterman, M. L. (1994). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York: Wiley.

Renkl, A. and R. K. Atkinson (2003). "Structuring the Transition From Example Study to Problem Solving in Cognitive Skill Acquisition: A Cognitive Load Perspective". *Educational Psychologist* 38.1, pp. 15–22. DOI: 10.1207/S15326985EP3801_3.

Roberts, K. and M. L. Weitzman (1981). "Funding Criteria for Research, Development, and Exploration Projects". *Econometrica* 49.5, pp. 1261–1288. DOI: 10.2307/1912754.

Ryan, R. M. and E. L. Deci (2000). "Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being". *American Psychologist* 55.1, pp. 68–78. DOI: 10.1037/0003-066X.55.1.68.

Smith, J. E. and K. F. McCardle (2002). "Structural Properties of Stochastic Dynamic Programs". *Operations Research* 50.5, pp. 796–809. DOI: 10.1287/opre.50.5.796.365.

Speklé, R. F. and F. H. M. Verbeeten (2014). "The Use of Performance Measurement Systems in the Public Sector: Effects on Performance". *Management Accounting Research* 25.2, pp. 131–146. DOI: 10.1016/j.mar.2013.07.004.

Suzuki, T. (2026). *When Openings Change Future Openings: A Fast–Slow Model of Experience-Dependent Transition Laws in Self-Referential Cognition*. Working Paper 2026-009. Tokyo: Research Institute of Business Administration (Sanken), Waseda University.

Topkis, D. M. (1998). *Supermodularity and Complementarity*. Princeton, NJ: Princeton University Press.

van de Pol, J., M. Volman, and J. Beishuizen (2010). "Scaffolding in Teacher–Student Interaction: A Decade of Research". *Educational Psychology Review* 22.3, pp. 271–296. DOI: 10.1007/s10648-010-9127-6.

Wood, D., J. S. Bruner, and G. Ross (1976). "The Role of Tutoring in Problem Solving". *Journal of Child Psychology and Psychiatry* 17.2, pp. 89–100. DOI: 10.1111/j.1469-7610.1976.tb00381.x.

A Bellman Operator, State Invariance, and Value Monotonicity

A.1 A precise state-invariance argument for the slow update

The reduced transition technology uses

$$\Phi(\ell, y, g) = \ell + \eta y(1-\ell) - (\delta_\ell + \chi g)\ell = \eta y + \{1 - \eta y - \delta_\ell - \chi g\}\ell. \tag{A.1}$$

Section 3.4 states the required invariance conclusion. The proof is most transparent as a direct lower–upper bound rather than as a literal convex-combination argument.

Proposition A.1 (State-space invariance)**.** *Suppose $\ell, y, g \in [0,1]$, $\eta, \delta_\ell, \chi \geq 0$, and*

$$\eta y + \delta_\ell + \chi g \leq 1. \tag{A.2}$$

Then $\Phi(\ell, y, g) \in [0,1]$.

Proof. Condition (A.2) makes the coefficient of ℓ in the second expression in (A.1) nonnegative. Hence $\Phi \geq \eta y \geq 0$. Since $\ell \leq 1$,

$$\Phi(\ell, y, g) \leq \eta y + 1 - \eta y - \delta_\ell - \chi g = 1 - \delta_\ell - \chi g \leq 1.$$

□

A.2 Function space and continuation nonexpansiveness

Let

$$X = \{R, A\} \times [0,1], \qquad \mathcal{V} = C([0,1]) \times C([0,1]),$$

with

$$\|v\|_\infty = \max_{j \in \{R,A\}} \sup_{\ell \in [0,1]} |v_j(\ell)|.$$

For a probability kernel $\mathsf{K}_j(d\ell', k \mid \ell, u)$, define

$$\mathcal{K}_j^u v(\ell) := \sum_{k \in \{R,A\}} \int_{[0,1]} v_k(\ell')\, \mathsf{K}_j(d\ell', k \mid \ell, u). \tag{A.3}$$

Lemma A.2 (Continuation nonexpansiveness and order preservation)**.** *For bounded vector functions v, w,*

$$|\mathcal{K}_j^u v(\ell) - \mathcal{K}_j^u w(\ell)| \leq \|v - w\|_\infty. \tag{A.4}$$

Moreover, $v \leq w$ implies $\mathcal{K}_j^u v(\ell) \leq \mathcal{K}_j^u w(\ell)$.

Proof. Under the joint next-state law,

$$\left|\mathbb{E}[v_{J^+}(\ell^+) - w_{J^+}(\ell^+)]\right| \leq \mathbb{E}\left[|v_{J^+}(\ell^+) - w_{J^+}(\ell^+)|\right] \leq \|v - w\|_\infty.$$

The order statement follows by integrating a pointwise nonnegative difference. □

A.3 Contraction, fixed point, and verification

For $v \in \mathcal{V}$, let

$$Q_j(\ell, u; v) = r_j(\ell, u) + \beta \mathcal{K}_j^u v(\ell), \tag{A.5}$$

$$(\mathcal{T}v)_R(\ell) = \max_{u \in \mathcal{U}_R} Q_R(\ell, u; v), \tag{A.6}$$

$$(\mathcal{T}v)_A(\ell) = \max\left\{ W(\ell), \max_{u \in \mathcal{U}_A} Q_A(\ell, u; v) \right\}. \tag{A.7}$$

Assume $\beta \in (0,1)$, compact action intervals, bounded continuous rewards and W, and a Feller continuation kernel.

Theorem A.3 (Bellman contraction and residual bound)**.** *The operator $\mathcal{T}$ maps $\mathcal{V}$ into itself and is a contraction with modulus β:*

$$\|\mathcal{T}v - \mathcal{T}w\|_\infty \le \beta \|v - w\|_\infty. \tag{A.8}$$

It therefore has a unique fixed point $V = (V_R, V_A)$. For $v^{n+1} = \mathcal{T}v^n$,

$$\|v^n - V\|_\infty \le \frac{\beta^n}{1-\beta} \|\mathcal{T}v^0 - v^0\|_\infty. \tag{A.9}$$

Proof. Feller continuity and compactness imply that the action maxima are attained and continuous in ℓ. For real families (a_x) and (b_x) on a common index set,

$$|\sup_x a_x - \sup_x b_x| \le \sup_x |a_x - b_x|.$$

Apply this inequality to each action branch and then Lemma A.2. The withdrawal branch is identical under v and w, so taking the maximum with W cannot enlarge the difference. This gives (A.8). Banach's fixed-point theorem yields existence, uniqueness, and uniform convergence. Finally,

$$\|v^0 - V\|_\infty \le \|v^0 - \mathcal{T}v^0\|_\infty + \beta \|v^0 - V\|_\infty,$$

so $\|v^0 - V\|_\infty \le (1-\beta)^{-1} \|\mathcal{T}v^0 - v^0\|_\infty$; applying n contractions gives (A.9). □

Theorem A.4 (Verification and stationary optimality)**.** *The fixed point in Theorem A.3 equals the supremum of the policy values fixed by the model constitution. A stationary deterministic optimal policy exists. It chooses an element of*

$$\arg\max_{u \in \mathcal{U}_R} Q_R(\ell, u; V)$$

in state R; in state A it withdraws when

$$W(\ell) \ge \max_{u \in \mathcal{U}_A} Q_A(\ell, u; V),$$

and otherwise chooses a maximizing continuation action.

Proof. The fixed-point equations imply the Bellman inequalities

$$V_j(\ell) \ge r_j(\ell, u) + \beta \mathcal{K}_j^u V(\ell), \qquad V_A(\ell) \ge W(\ell).$$

Iterating these inequalities along any admissible nonanticipative policy up to the earlier

of episode N and its withdrawal time bounds that policy's payoff by the fixed-point value plus a terminal term of magnitude at most $\beta^N \|V\|_\infty$. Letting $N \to \infty$ proves that V dominates every policy value.

Continuity and compactness make each argmax correspondence nonempty, compact valued, and upper hemicontinuous; a Borel selector exists. Under the stationary selector and the displayed stopping rule, every Bellman inequality used above is an equality until withdrawal, and $V_A = W$ at withdrawal. Repeating the finite-horizon iteration and letting $N \to \infty$ proves attainment. □

A.4 A same-state benchmark order and its exact strength

Let

$$\mathcal{V}_\uparrow = \{v \in \mathcal{V} : v_R \text{ and } v_A \text{ are nondecreasing in } \ell\}.$$

No cross-state ranking $v_A \geq v_R$ is imposed. This subsection records what such an unranked order demands; the maintained route, stated in the next subsection, uses the narrower product-order cone.

Assumption A.5 (Same-state legibility order (benchmark))**.** *For every j, u and every $v \in \mathcal{V}_\uparrow$,*

$$\ell_2 \geq \ell_1 \quad \Longrightarrow \quad \mathcal{K}_j^u v(\ell_2) \geq \mathcal{K}_j^u v(\ell_1).$$

In addition, W and $r_j(\cdot, u)$ are nondecreasing in ℓ.

Applied to the componentwise-constant elements $(v_R, v_A) = (1, 0)$ and $(0, 1)$ of $\mathcal{V}_\uparrow$, the displayed implication forces $\Pr(J^+ = R \mid j, \ell, u)$ and $\Pr(J^+ = A \mid j, \ell, u)$ to be simultaneously nondecreasing in ℓ; because they sum to one, both must be constant. Assumption A.5 is therefore equivalent to the conjunction of (a) invariance of the terminal coarse-state marginal in legibility and (b) first-order monotonicity of the conditional legibility law at each destination reached with positive probability. Requirement (a) fails for the micro-founded kernel of Section 5. At the largest common maintained action, $u = 0.15$, the terminal-A probability rises from 0.2894 to 0.5049 for an R-start episode and from 0.7388 to 0.8671 for an A-start episode as legibility moves from zero to one; the audit in Appendix G records strictly increasing terminal-A marginals at every maintained action with positive support, the zero-support R kernel being identically zero. The moving first-event summaries ordered by (3.12) are economically consistent with this movement but are not, by themselves, proof of it: an in-episode entry can be followed by an in-episode re-closure, so first-event order does not determine the terminal marginal. This is why the maintained route of Section 4.2 quantifies over the product-order cone instead.

Theorem A.6 (Value monotonicity under the benchmark order)**.** *Under the Bellman regularity conditions and Assumption A.5, the unique fixed point satisfies*

$$V_R(\ell_2) \geq V_R(\ell_1), \qquad V_A(\ell_2) \geq V_A(\ell_1) \quad (\ell_2 \geq \ell_1).$$

Proof. For $v \in \mathcal{V}_\uparrow$, every action-specific return in (A.5) is nondecreasing in ℓ. Maximization over a common action set preserves monotonicity, as does the maximum with nondecreasing W. Hence $\mathcal{T}\mathcal{V}_\uparrow \subseteq \mathcal{V}_\uparrow$. Starting value iteration from the zero vector and using Theorem A.3, the iterates converge uniformly to V. A uniform limit of nondecreasing functions is nondecreasing. □

Remark A.7 (What is inherited and what is additional)**.** *The companion transition model directly orders matched-condition entry, fusion, and re-closure summaries. That event ordering does not by itself order the full joint law of* (J^+, ℓ^+)*, and the preceding paragraph shows that the unranked same-state order overshoots: it additionally freezes the destination marginal. The maintained condition of the paper is therefore the product-order clause of Assumption 4.4, whose finite-grid analogue is the upper-set audit of Appendix G; the present assumption is retained as a transparent benchmark. The audit checks value monotonicity and the product-order dominance on the solved grids but does not convert either finite-grid fact into a continuum theorem.*

The final subsection of this appendix gives a coupling sufficient condition for the kernel clause of the maintained product order and records why the benchmark demands strictly more.

A.5 The maintained product-order route

Order the joint state by $R \prec A$ and by the usual order on legibility. The continuous functions increasing in this product order form the closed cone

$$\mathcal{V}_X^{\uparrow} := \{v \in \mathcal{V} : v_R, v_A \text{ are nondecreasing and } v_A(\ell) \geq v_R(\ell)\ \forall \ell\}. \tag{A.10}$$

This cone is narrower than $\mathcal{V}_{\uparrow}$ because it attaches an economic cross-state order to the technological product order. In the main text this cone is denoted $\mathbb{V}_{\uparrow}$; conditions (i) and (iii) of the next theorem, together with the payoff monotonicity in (ii), constitute Assumption 4.4.

Theorem A.8 (Monotonicity through the maintained product-order cone)**.** *Suppose: (i) for every action and every* $v \in \mathcal{V}_X^{\uparrow}$*, the continuation expectation is nondecreasing in current legibility; (ii)* $r_j(\cdot, u)$ *and* W *are nondecreasing; and (iii) the Bellman operator preserves the cross-state order,*

$$(\mathcal{T}v)_A(\ell) \geq (\mathcal{T}v)_R(\ell) \qquad (v \in \mathcal{V}_X^{\uparrow},\ \ell \in [0,1]). \tag{A.11}$$

Then $\mathcal{T}\mathcal{V}_X^{\uparrow} \subseteq \mathcal{V}_X^{\uparrow}$*. The unique fixed point is nondecreasing in legibility and satisfies* $V_A(\ell) \geq V_R(\ell)$*.*

Proof. Conditions (i)–(ii) make each component of $\mathcal{T}v$ nondecreasing in legibility. Condition (A.11) supplies the remaining cross-state inequality, so the cone is invariant. Starting value iteration from the zero vector and taking the uniform limit gives the result because the cone is closed. □

A convenient but stronger primitive route to (A.11) is: $\mathcal{U}_R \subseteq \mathcal{U}_A$; $r_A(\ell, u) \geq r_R(\ell, u)$ on common actions; and $\mathcal{K}_A^u v(\ell) \geq \mathcal{K}_R^u v(\ell)$ for every $v \in \mathcal{V}_X^{\uparrow}$ and common action. The finite-grid audit in Appendix G verifies product-order legibility monotonicity of both kernels, $V_A \geq V_R$, and $r_A \geq r_R$ on common actions. It does *not* verify this strongest primitive route because the maintained numerical domains satisfy $\max \mathcal{U}_R = 0.25 > 0.15 = \max \mathcal{U}_A$. In addition, the kernel-comparison component of this route is audited directly on the common actions: the dominance $\mathcal{K}_A^u v \geq \mathcal{K}_R^u v$ over the product cone holds at every grid point up to and including the first withdrawal point $\widehat{\ell}_W = 0.64$ and fails only on a small high-legibility set ($\ell \in \{0.92, 0.94, 1.00\}$; 149 of 491,232 upper-set comparisons, maximum deficit 0.0041), where withdrawal is already strictly optimal. A

matched-ceiling action domain would therefore repair the inclusion but would not, on the present kernel, deliver the global primitive route; it would deliver it at every grid point up to and including the first withdrawal point.

Principle A.9 (Analytical–numerical boundary)**.** *Conditions (i) and (iii) of Theorem A.8—equivalently, Assumption 4.4—are the maintained abstract sufficient conditions. The same-state Assumption A.5 is a benchmark only: it is strictly stronger, and it fails on any kernel whose destination marginal moves with legibility. The finite-grid upper-set audit instantiates the product-order dominance of clause (i) on the solved kernels, and the solved values realize the conclusions of the theorem; neither finite-grid fact is a proof of the continuum conditions, and the primitive inclusion route to* (A.11) *remains unavailable because* $\max \mathcal{U}_R = 0.25 > 0.15 = \max \mathcal{U}_A$.

A.6 Perturbation of an estimated or discretized interface

Let $\widehat{\mathcal{T}}$ be another Bellman operator with the same discount factor, fixed point $\widehat{V}$, and

$$\delta_M = \sup_{\|v\|_\infty \le M} \|\mathcal{T}v - \widehat{\mathcal{T}}v\|_\infty,$$

where both fixed points lie in the radius-M ball.

Proposition A.10 (Bellman perturbation bound)**.**

$$\|V - \widehat{V}\|_\infty \le \frac{\delta_M}{1-\beta}. \tag{A.12}$$

If the stopping-payoff, current-payoff, and continuation-operator discrepancies are bounded by ε_W, ε_r, *and* ε_K, *respectively, then*

$$\|V - \widehat{V}\|_\infty \le \frac{1}{1-\beta} \max\{\varepsilon_W, \varepsilon_r + \beta\varepsilon_K\}. \tag{A.13}$$

Proof. Using the fixed-point equations and adding and subtracting $\mathcal{T}\widehat{V}$,

$$\|V - \widehat{V}\|_\infty \le \beta \|V - \widehat{V}\|_\infty + \delta_M.$$

Rearrangement gives (A.12). The maximum inequality used in Theorem A.3 bounds δ_M by the maximum in (A.13). □

A.7 A coupling route to the maintained kernel clause

The final ingredient gives a pathwise sufficient condition for clause (i) of Assumption 4.4 and separates it into a part already implied by the maintained primitives and a residual that is genuinely economic. Recall the slow map (A.1) and the no-overshoot condition (A.2).

Lemma A.11 (Monotone-destination coupling and the kernel clause)**.** *Fix* $j \in \{R, A\}$, *an action* u, *and* $\ell_2 \ge \ell_1$. *Suppose the episode outputs admit a coupling under which the terminal coarse state never downgrades,* $J^+(\ell_1) = A \Rightarrow J^+(\ell_2) = A$ *almost surely, while* $Y(\ell_2) \ge Y(\ell_1)$ *and* $G(\ell_2) \le G(\ell_1)$ *almost surely. If the no-overshoot condition* (A.2) *holds pathwise, then* $\mathcal{K}_j^u v(\ell_2) \ge \mathcal{K}_j^u v(\ell_1)$ *for every* $v \in \mathcal{V}_X^{\uparrow}$*: the kernel clause of the maintained product order holds at* (j, u).

Proof. Under (A.2), $\partial\Phi/\partial\ell = 1 - \eta y - \delta_\ell - \chi g \geq 0$, and Φ is nondecreasing in y and nonincreasing in g on the unit cube, so $\ell^+(\ell_2) \geq \ell^+(\ell_1)$ pathwise on the coupled space. On paths with a common destination, the relevant component of v is nondecreasing in legibility, so the integrand is ordered. On upgrade paths, $J^+(\ell_1) = R$ and $J^+(\ell_2) = A$, and $v_A(\ell^+(\ell_2)) \geq v_R(\ell^+(\ell_2)) \geq v_R(\ell^+(\ell_1))$ by the cross-state ranking and monotonicity. Taking expectations gives the claim. □

Remark A.12 (The free part and the residual of the maintained order)**.** *The lemma splits clause (i) into a free part and a residual. Monotonicity of the slow map in its own state costs nothing beyond the maintained no-overshoot condition, so the economic content of the clause is stochastic monotonicity of the episode outputs (J^+, Y, G) in legibility at fixed action, with the destination permitted to upgrade from R to A. That residual is the object whose theorem-level status the companion paper declines to claim (its Remark 8.1): the post-entry state is endogenous to the entry time, so the required couplings are not supplied by the matched-state ordering results. Forcing destinations to coincide instead yields, by the same argument, the benchmark clause of Assumption A.5 over the larger cone $\mathcal{V}_\uparrow$; but such a coupling exists only under the destination-marginal invariance recorded after that assumption, which the micro-founded kernel violates. The maintained clause is therefore weaker than a claim inherited from the transition paper, stronger than the inherited first-event order, and strictly weaker than the benchmark; the finite-grid audit in Appendix G instantiates its upper-set analogue but cannot certify the continuum condition.*

B The Stochastic-Kernel Learning Dividend

B.1 Counterfactual construction and fixed-policy operators

Fix the common episode law

$$\mathsf{P}_{j,\ell,u}(dk, dy, dg)$$

over the terminal coarse state $k \in \{R, A\}$, qualified experience y, and evaluative capture g. The learned and frozen maps are

$$\Phi^L(\ell, y, g) = \ell + \eta y(1 - \ell) - (\delta_\ell + \chi g)\ell, \tag{B.1}$$

$$\Phi^F(\ell, y, g) = \ell. \tag{B.2}$$

They induce joint kernels K_j^L and K_j^F with the same terminal coarse-state marginal and the same current reward. Only the cross-episode legibility update differs.

A stationary deterministic policy π specifies actions $u_j^\pi(\ell)$ and a continuation indicator $s_R^\pi = 1$, $s_A^\pi(\ell) \in \{0, 1\}$. Let

$$\varrho_R^\pi(\ell) = r_R(\ell, u_R^\pi(\ell)), \tag{B.3}$$

$$\varrho_A^\pi(\ell) = s_A^\pi(\ell) r_A(\ell, u_A^\pi(\ell)) + \{1 - s_A^\pi(\ell)\} W(\ell), \tag{B.4}$$

$$(\mathcal{P}_m^\pi v)_j(\ell) = s_j^\pi(\ell) \sum_k \int v_k(\ell') \mathsf{K}_j^m(d\ell', k \mid \ell, u_j^\pi(\ell)), \quad m \in \{L, F\}. \tag{B.5}$$

The fixed-policy value solves

$$V_m^\pi = \varrho^\pi + \beta \mathcal{P}_m^\pi V_m^\pi. \tag{B.6}$$

Definition B.1 (Policy-fixed signed Learning Dividend)**.**

$$\mathrm{LD}^{\pi} = V_L^{\pi} - V_F^{\pi}. \tag{B.7}$$

A positive value is a dividend, a negative value a learning penalty, and zero a neutral transition-law effect under the fixed contingent rule.

Remark B.2 (Same contingent rule, not necessarily the same realized action path)**.** *The same mapping π is committed in both technologies. Once state histories diverge, the mapping can select different realized actions because it is evaluated at different states; that is not reoptimization. An identical realized-action-path comparison is obtained by restricting π to an open-loop or legibility-blind rule. The sign and resolvent results below continue to apply to that subclass.*

B.2 Counterfactual naming discipline

The one-cycle benchmark and the infinite-horizon Learning Dividend use related but not automatically identical reference laws. In the one-cycle problem, $\mathsf{L}_A^0(\cdot \mid \ell)$ is the *no-consolidation reference*: it may include autonomous decay or another post-withdrawal transition during the common comparison cycle. In the canonical Learning-Dividend problem, $\Phi^F(\ell, y, g) = \ell$ is the *clamped-legibility frozen technology.* The two coincide only when the no-consolidation reference itself places unit mass on the current legibility.

Accordingly, we reserve "frozen legibility" for the clamp in (B.2). The one-cycle object is called the no-consolidation or autonomous-reference learning wedge. This prevents a reader from believing that a counterfactual benchmark has been silently changed between the local consolidation argument and the infinite-horizon decomposition.

B.3 Resolvent and stopped-occupation identities

Theorem B.3 (Fixed-policy performance-difference identity)**.** *For every fixed admissible policy,*

$$\mathrm{LD}^{\pi} = \beta(I - \beta\mathcal{P}_L^{\pi})^{-1}(\mathcal{P}_L^{\pi} - \mathcal{P}_F^{\pi})V_F^{\pi}, \tag{B.8}$$

$$= \beta(I - \beta\mathcal{P}_F^{\pi})^{-1}(\mathcal{P}_L^{\pi} - \mathcal{P}_F^{\pi})V_L^{\pi}. \tag{B.9}$$

With

$$\omega^{\pi} = (\mathcal{P}_L^{\pi} - \mathcal{P}_F^{\pi})V_F^{\pi}, \tag{B.10}$$

we have

$$\mathrm{LD}^{\pi} = \beta \sum_{n=0}^{\infty} \beta^n (\mathcal{P}_L^{\pi})^n \omega^{\pi}. \tag{B.11}$$

Equivalently, from initial state x_0 and learned-technology withdrawal time τ_π,

$$\mathrm{LD}^{\pi}(x_0) = \beta \mathbb{E}_L^{\pi}\left[\sum_{n<\tau_\pi} \beta^n \omega^{\pi}(X_n) \mid X_0 = x_0\right]. \tag{B.12}$$

Proof. Subtract the two equations in (B.6) and add and subtract $\beta\mathcal{P}_L^{\pi}V_F^{\pi}$:

$$(I - \beta\mathcal{P}_L^{\pi})(V_L^{\pi} - V_F^{\pi}) = \beta(\mathcal{P}_L^{\pi} - \mathcal{P}_F^{\pi})V_F^{\pi}.$$

Because $\|\mathcal{P}_m^\pi\| \le 1$ and $\beta < 1$, the inverse exists and has the Neumann representation

$$(I - \beta\mathcal{P}_m^\pi)^{-1} = \sum_{n\ge 0} \beta^n (\mathcal{P}_m^\pi)^n.$$

This yields (B.8) and (B.11). Adding and subtracting $\beta\mathcal{P}_F^\pi V_L^\pi$ instead yields (B.9). Since $\mathcal{P}_L^\pi$ is substochastic and kills continuation mass at withdrawal, the Neumann series is exactly the stopped discounted occupation expression (B.12). □

Corollary B.4 (Exact sign and strictness)**.** *If $\omega^\pi \ge 0$ at every state reachable before withdrawal, then $\mathrm{LD}^\pi(x_0) \ge 0$. The reverse inequality gives a nonpositive dividend. The result is strict when the one-step inequality is strict on a set with positive learned discounted occupation probability.*

B.4 Coupling and stochastic-order conditions

Suppose the frozen-policy values are componentwise nondecreasing in legibility. The one-step wedge is

$$\omega_j^\pi(\ell) = s_j^\pi(\ell)\mathbb{E}\left[V_{F,J^+}^\pi\left(\Phi^L(\ell, Y, G)\right) - V_{F,J^+}^\pi(\ell)\right]. \tag{B.13}$$

Theorem B.5 (Conditional coupling sign theorem)**.** *At every reachable continuation state, suppose the learned and frozen next states admit a coupling with the same J^+ and*

$$\ell_L^+ \ge \ell_F^+ \quad \textit{almost surely.}$$

Then $\mathrm{LD}^\pi \ge 0$. If the order is reversed, $\mathrm{LD}^\pi \le 0$. Strictness follows when strict movement reaches a region on which the relevant frozen value is strictly increasing with positive discounted occupation probability.

Proof. Under the common-J^+ coupling and componentwise monotonicity,

$$V_{F,J^+}^\pi(\ell_L^+) - V_{F,J^+}^\pi(\ell_F^+) \ge 0$$

pathwise. Hence $\omega^\pi \ge 0$. Apply Theorem B.3 and positivity of the resolvent. The reverse and strict cases are identical with signs reversed. □

Corollary B.6 (Pathwise formation-versus-loss condition)**.** *If at every reachable continuation state*

$$\eta Y(1-\ell) \ge (\delta_\ell + \chi G)\ell \quad \textit{almost surely,} \tag{B.14}$$

then $\mathrm{LD}^\pi \ge 0$. Reversing (B.14) *gives a nonpositive dividend. Equality everywhere gives zero.*

The pathwise condition is sufficient, not necessary. In mixed distributions, the value-weighted expression (B.13), rather than the sign of $\mathbb{E}[\ell^+ - \ell]$ alone, determines the Learning Dividend.

B.5 Formation, natural loss, and evaluative capture

For one realized episode define

$$a_Y = \eta Y(1-\ell), \qquad a_\delta = \delta_\ell \ell, \qquad a_G = \chi G \ell,$$

and

$$\ell^Y = \ell + a_Y, \qquad \ell^{Y\delta} = \ell + a_Y - a_\delta, \qquad \ell^L = \ell + a_Y - a_\delta - a_G.$$

Under state invariance, all arguments lie in $[0,1]$ and $\ell^Y \geq \ell^{Y\delta} \geq \ell^L$.

Theorem B.7 (Exact signed transition-law ledger)**.** *If V_F^π is componentwise nondecreasing, define the one-step gross formation, natural-loss, and capture values by the corresponding differences*

$$\begin{aligned}
\gamma_Y^\pi(k) &= V_{F,k}^\pi(\ell^Y) - V_{F,k}^\pi(\ell), \\
\gamma_\delta^\pi(k) &= V_{F,k}^\pi(\ell^Y) - V_{F,k}^\pi(\ell^{Y\delta}), \\
\gamma_G^\pi(k) &= V_{F,k}^\pi(\ell^{Y\delta}) - V_{F,k}^\pi(\ell^L).
\end{aligned}$$

After conditional expectation and discounted occupation,

$$\mathrm{LD}^\pi(x_0) = \mathcal{B}_Y^\pi(x_0) - \mathcal{C}_\delta^\pi(x_0) - \mathcal{C}_G^\pi(x_0), \tag{B.15}$$

where all three gross terms on the right are nonnegative. Consequently,

$$\mathrm{LD}^\pi \gtreqless 0 \quad \Longleftrightarrow \quad \mathcal{B}_Y^\pi \gtreqless \mathcal{C}_\delta^\pi + \mathcal{C}_G^\pi.$$

Proof. For each destination state k, telescope:

$$\begin{aligned}
V_{F,k}^\pi(\ell^L) - V_{F,k}^\pi(\ell) = &\{V_{F,k}^\pi(\ell^Y) - V_{F,k}^\pi(\ell)\} \\
&- \{V_{F,k}^\pi(\ell^Y) - V_{F,k}^\pi(\ell^{Y\delta})\} \\
&- \{V_{F,k}^\pi(\ell^{Y\delta}) - V_{F,k}^\pi(\ell^L)\}.
\end{aligned}$$

Take the episode conditional expectation, substitute the resulting one-step decomposition into (B.12), and use monotonicity for nonnegativity. □

Remark B.8 (Attribution order)**.** *The net identity* (B.15) *is exact. With nonlinear continuation values, the allocation of gross interaction terms depends on the sequential channel order. The displayed order is formation, then natural loss, then evaluative capture. A Shapley allocation could be reported as a robustness decomposition, but it would not alter the net dividend or its sign.*

B.6 Separate optimization and the adaptation premium

Let $V_m^* = \sup_\pi V_m^\pi$ and let π_L^*, π_F^* be optimal policies.

Theorem B.9 (Reoptimization bounds)**.** *Pointwise in the initial state,*

$$\mathrm{LD}^{\pi_F^*} \leq \mathrm{LD}^* := V_L^* - V_F^* \leq \mathrm{LD}^{\pi_L^*}. \tag{B.16}$$

Moreover,

$$\mathrm{LD}^* = \mathrm{LD}^{\pi_F^*} + \underbrace{(V_L^* - V_L^{\pi_F^*})}_{\textit{learned-world adaptation premium} \geq 0}, \tag{B.17}$$

$$= \mathrm{LD}^{\pi_L^*} - \underbrace{(V_F^* - V_F^{\pi_L^*})}_{\textit{frozen-world policy mismatch} \geq 0}. \tag{B.18}$$

Proof. Optimality gives $V_L^* \geq V_L^{\pi_F^*}$ and $V_F^* \geq V_F^{\pi_L^*}$. Subtract $V_F^{\pi_F^*} = V_F^*$ from the first inequality and subtract the second from $V_L^{\pi_L^*} = V_L^*$. Adding and subtracting the cross-policy values gives (B.17)–(B.18). □

Principle B.10 (Main-text retention)**.** *The main article should retain the signed definition, the same-policy counterfactual discipline, the one-step wedge, and the economic sign interpretation. The operator inversion, occupation proof, channel ledger details, and reoptimization algebra can be placed in this Appendix. Removing the sign logic itself from the main text would obscure one of the paper's primary contributions.*

C Consolidation and the Economic Withdrawal Boundary

C.1 One-cycle consolidation with a fixed overhead

Let $h \in [0, \bar{h}]$ index one additional post-entry support block followed by withdrawal. Relative to immediate withdrawal, define variable surplus

$$\Psi(h;\ell) = B(h;\ell) - c(h;\ell) - D(h;\ell) + \beta \left\{ \int W \, d\mathsf{L}_A^h - \int W \, d\mathsf{L}_A^0 \right\}, \tag{C.1}$$

where B is current durability value, c variable resource cost, D current pressure damage, and L_A^h the terminal-legibility law. Retaining the active program incurs fixed overhead $f_A > 0$.

Proposition C.1 (Marginal decomposition)**.** *Under the differentiable representation* $\ell_h^+ = \mathcal{T}_A(h, \ell, \varepsilon)$*,*

$$\Psi_h(h;\ell) = B_h - c_h - D_h + \beta \mathbb{E}\left[W'(\ell_h^+) \frac{\partial \mathcal{T}_A}{\partial h}(h, \ell, \varepsilon) \right]. \tag{C.2}$$

Proof. The autonomous reference term in (C.1) is constant in h. Differentiate the remaining terms and use the assumed integrable envelope to interchange derivative and expectation. □

Theorem C.2 (Intensive and extensive consolidation margins)**.** *Suppose* $\Psi(\cdot;\ell)$ *is strictly concave,* $\Psi(0;\ell) = 0$*, and write*

$$\widehat{h}(\ell) = \arg\max_{h \in [0,\bar{h}]} \Psi(h;\ell).$$

Then:

(i) If $\Psi_h(0^+;\ell) \leq 0$*, immediate withdrawal is uniquely optimal.*

(ii) If $\Psi_h(0^+;\ell) > 0$*, then* $\widehat{h}(\ell) > 0$ *and it satisfies the marginal condition in* (C.2) *when interior.*

(iii) Positive consolidation is actually selected if and only if

$$\Psi(\widehat{h}(\ell);\ell) > f_A. \tag{C.3}$$

Proof. Concavity and $\Psi(0) = 0$ imply $\Psi(h) \leq \Psi_h(0^+)h$ for $h > 0$. This proves (i), and the fixed overhead makes the inequality strict at the policy level. A positive initial derivative makes the unique variable-surplus maximizer positive, proving (ii). The best

positive branch has incremental value $\Psi(\widehat{h}) - f_A$ relative to immediate withdrawal, which proves (iii). □

The theorem separates a dose condition from a program-retention condition. A favorable marginal block can fail to cover the fixed overhead; conversely, future transition-law value can make a finite block cover that overhead even when current benefit alone would not.

C.2 Global and eventual single crossing

For the solved Bellman value, define

$$Q_A(\ell, u) = r_A(\ell, u) + \beta \mathcal{K}_A^u[V_R, V_A](\ell), \tag{C.4}$$

$$\Delta_A(\ell) = \max_{u \in \mathcal{U}_A} Q_A(\ell, u) - W(\ell). \tag{C.5}$$

Theorem C.3 (Upper-set threshold characterization)**.** *Let Δ_A be continuous. Suppose the stopping set*

$$\mathcal{S}_A = \{\ell \in [0,1] : \Delta_A(\ell) \leq 0\}$$

is a nonempty upper set and that continuation is optimal somewhere. Then, with

$$\ell_W := \inf \mathcal{S}_A,$$

we have $\ell_W \in (0,1)$, $\Delta_A(\ell_W) = 0$, continuation is optimal for $\ell < \ell_W$, and withdrawal is optimal for $\ell \geq \ell_W$ under withdrawal tie breaking.

Proof. Continuity makes $\mathcal{S}_A$ closed, so its infimum belongs to the set. No point below the infimum lies in $\mathcal{S}_A$, hence the gap is positive there. Because $\mathcal{S}_A$ is an upper set, every point above the infimum belongs to it. Finally, if $\Delta_A(\ell_W) < 0$, continuity would place negative-gap points strictly below ℓ_W, contradicting the definition of the infimum. Thus the boundary is an indifference point. □

This theorem separates the economic threshold from any particular sufficient shape restriction. Global strict decrease, eventual strict decrease, and finite-action monotone-stopping arguments are alternative ways to establish the upper-set premise.

Theorem C.4 (Global strict-single-crossing benchmark)**.** *If Δ_A is continuous and strictly decreasing on $[0,1]$, with $\Delta_A(0) > 0 > \Delta_A(1)$, there is a unique $\ell_W \in (0,1)$ satisfying $\Delta_A(\ell_W) = 0$. Continuation is optimal below ℓ_W and withdrawal above it.*

Proof. Existence follows from continuity and opposite endpoint signs. Strict decrease gives uniqueness and the sign on either side. □

The clean theorem is useful, but it is stronger than needed. The numerical audit shows a low-legibility region in which the active advantage rises while remaining very positive. What matters for a unique upper stopping boundary is decline after the active advantage reaches its economically relevant high point.

Theorem C.5 (Eventual single crossing)**.** *Suppose there is $\bar{\ell} \in [0,1)$ such that:*

(i) Δ_A is continuous on $[0,1]$;

(ii) $\Delta_A(\ell) > 0$ for every $\ell \in [0, \bar{\ell}]$;

(iii) Δ_A *is strictly decreasing on* $[\bar{\ell}, 1]$*; and*

(iv) $\Delta_A(1) < 0$.

Then there is a unique $\ell_W \in (\bar{\ell}, 1)$ *with* $\Delta_A(\ell_W) = 0$*, and the optimal stopping set is the upper interval* $[\ell_W, 1]$ *under withdrawal tie breaking.*

Proof. At $\bar{\ell}$ the gap is positive, while at one it is negative. Continuity gives a root in $(\bar{\ell}, 1)$. Strict decrease on that interval gives uniqueness. The gap is positive on the entire lower region by assumption and positive between $\bar{\ell}$ and the root by strict decrease; it is negative above the root. □

Remark C.6 (Why this is a refinement, not an ex post rescue)**.** *Theorem C.5 does not use the numerical threshold to define the theorem. It states the minimal shape required for an upper stopping region when low-legibility continuation can become more attractive before autonomy catches up. The global theorem remains a transparent sufficient special case with* $\bar{\ell} = 0$.

A fold discontinuity in the autonomous value. If the withdrawal value inherits a jump at the structural threshold—as it can under a basin-membership persistence criterion, since the positive basin appears with positive extent at a saddle–node fold—then Δ_A has a downward jump at $\ell^\dagger$. The upper-set characterization (Theorem C.3) and the eventual-crossing route (Theorem C.5) survive with continuity weakened to: Δ_A continuous except for downward jumps, with the sign change bracketed inside a subinterval of continuity. Downward jumps cannot create a second upper crossing under the strict-decline hypothesis on the upper region; they can only enlarge the stopping set weakly, and if the sign change occurs across a jump the boundary point is a strict-preference switch rather than an indifference point. In the numerical model the sign bracket $[0.62, 0.64]$ lies well above $\ell^\dagger$, in a region where the persistence curve and hence W vary smoothly on the disclosed grid, so the baseline threshold conclusions do not rest on continuity of W at the fold.

C.3 Primitive sufficient conditions on the relevant upper region

Let $D_A(\ell, u) = Q_A(\ell, u) - W(\ell)$ and let

$$M_V = \max\{\|V_R\|_\infty, \|V_A\|_\infty\}.$$

Use the variation norm

$$\|\mu - \nu\|_{\mathrm{TV}} = \sup_{\|f\|_\infty \le 1} \left| \int f \, d\mu - \int f \, d\nu \right|.$$

Theorem C.7 (Upper-region strong-slope criterion)**.** *Fix* $\bar{\ell} \in [0, 1)$*. Suppose that uniformly over* $u \in \mathcal{U}_A$ *and* $\bar{\ell} \le \ell_1 < \ell_2 \le 1$*,*

$$W(\ell_2) - W(\ell_1) \ge m_W(\ell_2 - \ell_1), \tag{C.6}$$

$$r_A(\ell_2, u) - r_A(\ell_1, u) \le L_r(\ell_2 - \ell_1), \tag{C.7}$$

$$\|\mathsf{K}_A(\cdot \mid \ell_2, u) - \mathsf{K}_A(\cdot \mid \ell_1, u)\|_{\mathrm{TV}} \le L_K(\ell_2 - \ell_1). \tag{C.8}$$

If

$$m_W > L_r + \beta M_V L_K, \tag{C.9}$$

then Δ_A is strictly decreasing on $[\bar{\ell}, 1]$.

Proof. For every fixed action,

$$D_A(\ell_2, u) - D_A(\ell_1, u) = r_A(\ell_2, u) - r_A(\ell_1, u) + \beta\{\mathcal{K}_A^u V(\ell_2) - \mathcal{K}_A^u V(\ell_1)\} - \{W(\ell_2) - W(\ell_1)\}.$$

The kernel term is at most $M_V L_K(\ell_2 - \ell_1)$ in absolute value. Hence

$$D_A(\ell_2, u) - D_A(\ell_1, u) \leq -\{m_W - L_r - \beta M_V L_K\}(\ell_2 - \ell_1).$$

The bound is uniform in u, so taking the maximum over the common action set preserves it. □

This localizes the primitive condition to the region in which withdrawal can actually become optimal. It does not require autonomy to dominate the active branch's slope while legibility is so low that no stopping decision is near.

C.4 Structural and economic thresholds

Assume an interior economic root and an inherited structural threshold $\ell^\dagger \in (0, 1)$. Define

$$\mathfrak{M}^\dagger = \Delta_A(\ell^\dagger).$$

Theorem C.8 (Structural–economic threshold gap)**.** *Under either Theorem C.4 or Theorem C.5, provided $\ell^\dagger$ lies in a region on which the gap has the threshold sign pattern,*

$$\mathfrak{M}^\dagger \begin{cases} > 0 & \iff \ell_W > \ell^\dagger, \\ = 0 & \iff \ell_W = \ell^\dagger, \\ < 0 & \iff \ell_W < \ell^\dagger. \end{cases} \tag{C.10}$$

Proof. The gap is positive below the unique root and negative above it. Evaluate that sign at $\ell^\dagger$. □

Equality requires the independent economic balance

$$\max_u \{r_A(\ell^\dagger, u) + \beta \mathcal{K}_A^u V(\ell^\dagger)\} = W(\ell^\dagger).$$

It is not implied by the saddle–attractor geometry.

C.5 Comparative statics through gap shifts

Proposition C.9 (Threshold movement)**.** *Let θ index an environment and suppose each $\Delta_A(\cdot; \theta)$ has a unique upper crossing. If*

$$\Delta_A(\ell; \theta_2) \geq \Delta_A(\ell; \theta_1) \quad \textit{for every } \ell,$$

then $\ell_W(\theta_2) \geq \ell_W(\theta_1)$. If differentiable at an interior root,

$$\frac{d\ell_W}{d\theta} = -\frac{\Delta_{A,\theta}(\ell_W; \theta)}{\Delta_{A,\ell}(\ell_W; \theta)}.$$

Proof. At the old root the new gap is nonnegative. Since the new gap is negative above and crosses once, its root cannot be below the old root. The derivative formula is the implicit-function theorem. □

Parameter names alone do not sign the shift. Re-closure damage delays withdrawal only when continuation reduces the relevant discounted burden relative to autonomy. Greater patience delays withdrawal only when the delayed active-over-autonomous differentials are nonnegative. Learning efficiency can raise the state threshold while shortening the calendar time needed to reach it.

Principle C.10 (Recommended main-text refinement)**.** *Section 4.4 presents the global theorem as a benchmark and the eventual-single-crossing theorem as the result matched by the numerical example. Section 5.6 reports one upper crossing and does not treat the numerical exercise as a verification of global strict decrease of* Δ_A*.*

D Monotone Taper, Starter, and Phase Realization

D.1 Extremal selections under decreasing differences

On the continuation region, let

$$\mathcal{U}_A^*(\ell) = \arg\max_{u\in[0,\bar{u}_A]} Q_A(\ell, u).$$

Define its minimum and maximum selections $\underline{u}_A^*(\ell)$ and $\overline{u}_A^*(\ell)$.

Assumption D.1 (Decreasing differences)**.** *For* $\ell_2 > \ell_1$ *and* $u_2 > u_1$*,*

$$Q_A(\ell_2, u_2) - Q_A(\ell_2, u_1) \le Q_A(\ell_1, u_2) - Q_A(\ell_1, u_1). \tag{D.1}$$

Theorem D.2 (Monotone taper for extremal selections)**.** *Under Assumption D.1, both* $\underline{u}_A^*$ *and* $\overline{u}_A^*$ *are nonincreasing in* ℓ*. If the maximizer is unique,* $u_A^*(\ell)$ *is nonincreasing.*

Proof. Take $\ell_2 > \ell_1$. Suppose, contrary to the maximal-selection claim, that $\overline{u}_A^*(\ell_2) > \overline{u}_A^*(\ell_1)$. Optimality at ℓ_2 gives

$$Q_A(\ell_2, \overline{u}_A^*(\ell_2)) - Q_A(\ell_2, \overline{u}_A^*(\ell_1)) \ge 0.$$

Decreasing differences makes the corresponding difference at ℓ_1 at least as large. Optimality of $\overline{u}_A^*(\ell_1)$ makes it at most zero. Equality follows, so $\overline{u}_A^*(\ell_2)$ is also an optimizer at ℓ_1, contradicting maximality there.

For the minimum selection, suppose $\underline{u}_A^*(\ell_2) > \underline{u}_A^*(\ell_1)$. Optimality at ℓ_1 gives a nonpositive difference between the higher and lower actions. Decreasing differences makes the corresponding difference at ℓ_2 no larger, while optimality of $\underline{u}_A^*(\ell_2)$ makes it nonnegative. Equality follows, making the lower action an optimizer at ℓ_2, contrary to minimality. □

D.2 A weaker finite-action route: pairwise action single crossing

The decreasing-differences condition is convenient and global. A finite numerical action set needs less. Let $\mathcal{U}_A = \{u_1 < \cdots < u_m\}$ and, for $u_h > u_l$, write

$$d_{hl}(\ell) = Q_A(\ell, u_h) - Q_A(\ell, u_l).$$

Assumption D.3 (Pairwise single crossing from higher to lower support)**.** *For every $u_h > u_l$ and $\ell_2 > \ell_1$,*

$$\textit{(i)}\ \ d_{hl}(\ell_1) \le 0 \implies d_{hl}(\ell_2) \le 0, \qquad \textit{(ii)}\ \ d_{hl}(\ell_1) < 0 \implies d_{hl}(\ell_2) < 0. \quad \text{(D.2)}$$

Clause (i) says that once a higher action is weakly no better than a lower action, it never becomes strictly better again as legibility rises; clause (ii) preserves the strict comparison. Together they are the decreasing form of single crossing applied to each pairwise advantage d_{hl}; decreasing differences imply both clauses because every pairwise advantage is then nonincreasing in ℓ, and the converse need not hold.

Theorem D.4 (Monotone taper under pairwise action single crossing)**.** *On a finite ordered action set: (a) under clause (i) of Assumption D.3, the minimum optimal selection is nonincreasing in legibility; (b) under clauses (i) and (ii), the maximum optimal selection is nonincreasing as well.*

Proof. Take $\ell_2 > \ell_1$.

(a) Suppose the minimum selection rises, so $u_h = \underline{u}^*_A(\ell_2) > u_l = \underline{u}^*_A(\ell_1)$. If $d_{hl}(\ell_2) \le 0$, then u_l attains at least the optimal value at ℓ_2 and is itself optimal there, contradicting minimality of u_h. Hence $d_{hl}(\ell_2) > 0$, and the contrapositive of clause (i) gives $d_{hl}(\ell_1) > 0$. Then u_h strictly improves on u_l at ℓ_1, contradicting optimality of u_l at ℓ_1.

(b) Suppose the maximum selection rises, so $u_h = \overline{u}^*_A(\ell_2) > u_l = \overline{u}^*_A(\ell_1)$. If $d_{hl}(\ell_1) \ge 0$, then u_h would also be optimal at ℓ_1, contradicting maximality of u_l there. Hence $d_{hl}(\ell_1) < 0$, and clause (ii) gives $d_{hl}(\ell_2) < 0$, contradicting optimality of u_h at ℓ_2. □

Clause (i) alone does not order the maximum selection. The configuration $d_{hl}(\ell_1) < 0 = d_{hl}(\ell_2)$ is compatible with clause (i): the higher action is strictly worse at the lower state, ties at the higher state, and the maximal selection jumps upward from u_l to u_h. The strict clause removes exactly this tie case. In the finite-grid audit, 32 of 192 adjacent global decreasing-differences checks fail, but all 21 pairwise action comparisons satisfy clause (i), and the reported post-entry policy—the minimum optimal selection, with ties broken toward the smaller action—has no upward step. Section 4.5 therefore presents decreasing differences as the clean continuum theorem, clause (i) as the route actually verified by the disclosed finite action set and instantiated by the reported selection, and clause (ii) as the addition required to order the maximum selection.

D.3 Strict interior taper

Proposition D.5 (Differentiable strict taper)**.** *Suppose Q_A is twice continuously differentiable, strictly concave in u, and its unique optimizer is interior. If*

$$Q_{A,u\ell} < 0, \qquad Q_{A,uu} < 0,$$

then

$$\frac{du^*_A(\ell)}{d\ell} = -\frac{Q_{A,u\ell}}{Q_{A,uu}} < 0. \quad \text{(D.3)}$$

Proof. Differentiate $Q_{A,u}(\ell, u^*_A(\ell)) = 0$ and apply the implicit-function theorem. □

The cross-partial applies to the solved dynamic continuation return, not only the current reward. A static decline in current marginal benefit is insufficient when a larger action still creates a sufficiently valuable future-state shift.

D.4 Consolidation plateau and taper onset

Suppose Q_A is strictly concave in u, $Q_{A,u\ell} < 0$, and

$$Q_{A,u}(0, \bar{u}_A) > 0, \qquad Q_{A,u}(\ell_W, \bar{u}_A) < 0, \qquad Q_{A,u}(\ell, 0) > 0 \quad (\ell < \ell_W).$$

Theorem D.6 (Plateau, interior taper, and withdrawal)**.** *There is a unique $\ell_T \in (0, \ell_W)$ such that $Q_{A,u}(\ell_T, \bar{u}_A) = 0$. The post-entry policy has the form*

$$\pi_A^*(\ell) = \begin{cases} \bar{u}_A, & 0 \le \ell \le \ell_T, \\ u_A^*(\ell) \in (0, \bar{u}_A), & \ell_T < \ell < \ell_W, \\ \mathsf{W}, & \ell \ge \ell_W, \end{cases} \tag{D.4}$$

with u_A^* strictly decreasing on the interior taper interval.

Proof. The upper-bound marginal $m_{\bar{u}}(\ell) = Q_{A,u}(\ell, \bar{u}_A)$ is continuous and strictly decreasing. Its opposite endpoint signs give a unique zero. Below the zero, strict concavity makes the derivative positive throughout the action interval, so the upper boundary is optimal. Above it and below ℓ_W, the derivative is positive at zero and negative at the upper boundary, so a unique interior root exists; Proposition D.5 makes it strictly decreasing. The stopping theorem supplies withdrawal at and above ℓ_W under the chosen tie rule. □

Remark D.7 (Withdrawal is a regime change)**.** *The left limit $u_A^*(\ell_W^-)$ may be positive. Withdrawal is not the action $u = 0$ inside the active program; it terminates the relation and its overhead. A smooth sunset requires the additional boundary condition $Q_{A,u}(\ell_W, 0) = 0$.*

D.5 Interior starter and safe window

For the unresolved state, let

$$\Gamma_R(\ell, u) = Q_R(\ell, u) - Q_R(\ell, 0).$$

Theorem D.8 (Interior starter and safe window)**.** *Fix ℓ. If $Q_R(\ell, \cdot)$ is strictly concave and*

$$Q_{R,u}(\ell, 0) > 0 > Q_{R,u}(\ell, \bar{u}_R),$$

there is a unique interior starter. If in addition $Q_R(\ell, \bar{u}_R) < Q_R(\ell, 0)$, there is a unique $\bar{u}_R^{\text{safe}} > u_R^*$ such that positive support beats no support exactly for $0 < u < \bar{u}_R^{\text{safe}}$.

Proof. Strict concavity makes the derivative strictly decreasing, so the endpoint signs give one interior zero and maximizer. The incremental return starts at zero with positive derivative, reaches its unique maximum, then strictly decreases. The negative upper-end comparison creates one and only one second zero. □

D.6 Four policy regions and chronological realization

Combining Theorems C.5, D.6, and D.8 yields four *state regions*: starter in R, consolidation at low post-entry legibility, taper at intermediate post-entry legibility, and withdrawal at the upper stopping boundary. A chronological four-phase path additionally requires entry, no intervening re-closure, monotone progress in ℓ, and actual visitation of the taper interval.

Corollary D.9 (State taper versus pathwise taper)**.** *If $u_A^*(\ell)$ is nonincreasing and a realized active path satisfies $\ell_{n+1} \geq \ell_n$ until withdrawal, then the realized support sequence is nonincreasing. If legibility falls, a state-contingent optimal policy may re-escalate support without violating the state-taper theorem.*

Remark D.10 (Action-domain dependence of ℓ_T)**.** *The theorem defines ℓ_T relative to an upper feasible action. Enlarging the action domain can remove a ceiling plateau because the optimizer is interior from the first post-entry state; coarsening the action grid can postpone the first observable drop. Accordingly, monotone taper is the robust policy-shape object. The exact numerical location of ℓ_T is a maintained-domain diagnostic, not a technology invariant.*

E Autonomous Re-entry and Partial Observation

E.1 A semi-Markov re-entry version of the withdrawal value

The baseline treats withdrawal as irreversible inside the current active-support contract. A robustness extension can allow a later, costly re-entry opportunity without returning provider incentives or population finance to the model.

Let an autonomous spell beginning at legibility ℓ generate a holding time $\tau > 0$, next legibility ℓ', and indicator $e \in \{0, 1\}$ from a semi-Markov kernel

$$\mathsf{Q}_0(d\tau, d\ell', de \mid \ell).$$

The indicator $e = 1$ means that a new support-contract opportunity becomes available at the end of the spell. Let $R_0(\ell, \tau, e)$ be bounded expected discounted functional flow during the spell, $\rho > 0$ the continuous discount rate, and $\kappa_R(\ell') \geq 0$ the re-entry cost. Given the active unresolved-state value V_R, define

$$(\mathcal{S}w)(\ell) = \int \Big[R_0(\ell, \tau, e) + e^{-\rho\tau} \big\{ (1-e) w(\ell') \tag{E.1}$$

$$+ e \max\left(w(\ell'), V_R(\ell') - \kappa_R(\ell') \right) \big\} \Big] \mathsf{Q}_0(d\tau, d\ell', de \mid \ell). \tag{E.2}$$

Assumption E.1 (Discounted semi-Markov contraction)**.** *There is $\bar{d} < 1$ such that*

$$\sup_{\ell} \int e^{-\rho\tau} \mathsf{Q}_0(d\tau, d\ell', de \mid \ell) \leq \bar{d}.$$

Theorem E.2 (Semi-Markov autonomous value)**.** *The operator $\mathcal{S}$ is a contraction on bounded functions with modulus at most $\bar{d}$. It has a unique bounded fixed point W^{SM}. Replacing W by W^{SM} in the active Bellman equation leaves the Bellman contraction proof unchanged, provided W^{SM} satisfies the required boundedness and continuity conditions.*

Proof. The scalar map $x \mapsto \max\{x, c\}$ is nonexpansive. Therefore, for bounded w, z,

$$|(\mathcal{S}w)(\ell) - (\mathcal{S}z)(\ell)| \le \int e^{-\rho\tau}\|w - z\|_\infty \, d\mathbb{Q}_0 \le \bar{d}\|w - z\|_\infty.$$

Banach's theorem gives the fixed point. The active Bellman operator still takes the maximum of a common stopping payoff and continuation actions, so its modulus remains β. □

Remark E.3 (Boundary against costless immediate re-entry)**.** *If a zero-delay, zero-cost re-entry option is always available, withdrawal can preserve essentially the same control option as remaining active. The positive-overhead distinction that makes stopping nondegenerate can then collapse. A meaningful re-entry extension therefore requires a new contract, positive delay, positive cost, or another friction. This is a technical condition, not a provider-agency model.*

E.2 Belief-state formulation when legibility is latent

Suppose the coarse state J is observed but ℓ is latent. Let $b \in \mathcal{P}([0,1])$ be the beginning-of-episode belief. Conditional on latent ℓ, action u, and current j, let

$$\mathsf{M}_j(dy, dk, d\ell' \mid \ell, u)$$

be the joint physical-and-observation kernel for signal y, next coarse state k, and next legibility ℓ'. The predictive law under belief b is obtained by integrating over $b(d\ell)$. A regular conditional posterior is denoted

$$\mathfrak{B}_j(b, u, y, k)(d\ell').$$

Define belief-averaged reward and withdrawal value

$$\bar{r}_j(b, u) = \int r_j(\ell, u) b(d\ell),$$
$$\bar{W}(b) = \int W(\ell) b(d\ell).$$

Definition E.4 (Belief-state Bellman system)**.**

$$\mathbb{V}_R(b) = \sup_{u \in \mathcal{U}_R} \left\{ \bar{r}_R(b, u) + \beta \mathbb{E}\left[\mathbb{V}_{J^+}(\mathfrak{B}_R(b, u, Y, J^+)) \right] \right\}, \tag{E.3}$$

$$\mathbb{V}_A(b) = \max\left\{ \bar{W}(b), \sup_{u \in \mathcal{U}_A} \left[\bar{r}_A(b, u) + \beta \mathbb{E}\left[\mathbb{V}_{J^+}(\mathfrak{B}_A(b, u, Y, J^+)) \right] \right] \right\}. \tag{E.4}$$

Proposition E.5 (Belief-state contraction)**.** *With bounded rewards and $\beta \in (0,1)$, the Bellman operator in* (E.3)–(E.4) *is a sup-norm contraction of modulus β on bounded value functions over $\{R, A\} \times \mathcal{P}([0,1])$.*

Proof. For a fixed action, the only value-dependent term is an expectation under a probability law over next beliefs and next coarse states. It is nonexpansive in the sup norm. Maximization and the common stopping branch preserve the bound. □

Remark E.6 (No automatic threshold in the posterior mean)**.** *The complete-information threshold does not imply a threshold in $\mathbb{E}_b[\ell]$. Two beliefs with the same*

mean can differ in downside mass and therefore in expected withdrawal value or continuation value. Monotone likelihood-ratio order, total positivity, or another belief order would be needed for a monotone stopping theorem. The Appendix formulates the POMDP; it does not claim that the scalar policy survives unchanged.

Principle E.7 (Measurement neutrality is substantive)**.** *A Blackwell-more-informative signal has weakly higher decision value only when the signal change leaves the current payoff and physical transition kernel unchanged. Participant-facing measurement that raises evaluative capture changes the person as well as the information set and must be evaluated through the full physical-plus-information value.*

F Type Uncertainty and the Trial Option

F.1 Commitment, adaptation, and pure information value

Let persistent type $\theta \sim \mu$, and let a one-episode trial action u produce signal Z and a bounded current payoff. Let Π_T be a finite continuation-policy menu and $L_\pi(\theta, Z)$ the post-trial continuation value of policy π. Define

$$\mathcal{C}(u) = \bar{r}_T(u) + \beta \max_{\pi \in \Pi_T} \mathbb{E}[L_\pi(\theta, Z)], \tag{F.1}$$

$$\mathcal{T}(u) = \bar{r}_T(u) + \beta \mathbb{E}\left[\max_{\pi \in \Pi_T} \mathbb{E}[L_\pi(\theta, Z) \mid Z]\right], \tag{F.2}$$

$$\mathcal{I}(u) = \mathcal{T}(u) - \mathcal{C}(u), \tag{F.3}$$

where $\mathcal{T}$ here denotes adaptive trial value, not the Bellman operator. Let $\mathcal{B}$ be the best no-trial value and $\mathcal{O}(u) = \mathcal{T}(u) - \mathcal{B}$.

Theorem F.1 (Trial information cannot reduce value)**.**
If the trial is decision-reversible—every precommitment policy remains available after every signal—then

$$\mathcal{I}(u) \geq 0.$$

Proof. For every fixed π and signal realization z,

$$\max_{\rho \in \Pi_T} \mathbb{E}[L_\rho(\theta, Z) \mid Z = z] \geq \mathbb{E}[L_\pi(\theta, Z) \mid Z = z].$$

Take expectations and then maximize the right-hand side over the ex ante commitment policy. The current trial payoff cancels. □

Proposition F.2 (Strict information value)**.** *If conditional policy values have unique maximizers almost surely and two positive-probability signal sets select different continuation policies, then $\mathcal{I}(u) > 0$. If one policy is conditionally optimal almost surely, then $\mathcal{I}(u) = 0$.*

Proof. When different unique maximizers occur, no single commitment policy attains the pointwise conditional maximum almost surely, so the inequality in Theorem F.1 is strict on a positive-probability set. If one policy is conditionally optimal everywhere, committing to it reproduces the adaptive value. □

Theorem F.3 (Blackwell order for a fixed physical trial)**.** *If signal Z_0 is a garbling of Z_1 while the current payoff, physical state law, and continuation menu are fixed, the adaptive value under Z_1 is weakly higher.*

Proof. Every decision rule based on Z_0 can be implemented after observing Z_1 by first applying the garbling kernel. The richer signal therefore offers a weak superset of feasible contingent rules. □

F.2 Net option value and information-created trials

Proposition F.4 (Exact option decomposition)**.**

$$\mathcal{O}(u) = \{\mathcal{C}(u) - \mathcal{B}\} + \mathcal{I}(u). \tag{F.4}$$

An information-created trial satisfies

$$\mathcal{C}(u) \leq \mathcal{B} < \mathcal{T}(u).$$

Proof. Add and subtract $\mathcal{C}(u)$. The displayed inequalities say that the physical trial would be rejected if its signal had to be ignored, while adaptive use of the signal reverses the decision. □

Positive information value does not imply a positive trial option. Resource cost, delay, fusion exposure, and evaluative capture are contained in $\mathcal{C} - \mathcal{B}$ and can dominate $\mathcal{I}$.

F.3 Binary-type benchmark

Let $\theta \in \{H, L\}$, prior $p = \mathbb{P}(H)$, and type-specific continuation advantages $d_H > 0 > d_L$ relative to withdrawal. Continuation is optimal at posterior p' if

$$D(p') = p' d_H + (1 - p') d_L \geq 0,$$

with threshold

$$p^* = \frac{-d_L}{d_H - d_L}.$$

For posterior realizations p_g, p_b after a binary signal, the post-trial information value is

$$I_{post} = P_g[D(p_g)]_+ + (1 - P_g)[D(p_b)]_+ - [D(p)]_+. \tag{F.5}$$

Proposition F.5 (Posterior crossing criterion)**.** *If both signals have positive probability, then $I_{post} > 0$ exactly when*

$$p_b < p^* < p_g.$$

Proof. Posterior beliefs are a martingale. The map $p \mapsto [D(p)]_+$ is convex and piecewise linear with its only kink at p^*. Jensen's inequality is strict exactly when the posterior support straddles that kink. □

Sign structure in the prior. At a degenerate prior the posterior equals the prior for every signal realization; with the persistent type known and the economic state observed, a single continuation policy is conditionally optimal almost surely, so $\mathcal{I}(u) = 0$ by the second clause of Proposition F.2. If in addition $\mathcal{C}(u) - \mathcal{B} < 0$ at those priors and the three value objects are continuous in the prior, then $\mathcal{O}(u) < 0$ in neighborhoods of $p \in \{0, 1\}$. Any information-created region therefore lies in the interior of the prior interval and inside the posterior-crossing region $\{p : p_b < p^* < p_g\}$; its existence is not guaranteed, because the premium generated by posterior crossing must exceed the

physical deficit $|\mathcal{C} - \mathcal{B}|$ somewhere in that region. Interiority and location do not by themselves make the region one interval or force it to contain p^*: the physical component $\mathcal{C}(u) - \mathcal{B}$ varies with the prior, so the positive set of $\mathcal{O}$ can in principle split or sit on one side of p^*. Under additional shape conditions—for instance, a prior-invariant physical component, or single-peakedness of $\{\mathcal{C}(u) - \mathcal{B}\} + \mathcal{I}(u)$ in the prior—the region is a single interval around p^*. The single-point computation in Section 5.8 lies inside the positive region; mapping that region over the prior and the trial dose is recorded as an extension in Section 7.

Principle F.6 (Subordinate placement)**.** *The main article should keep the economic decomposition $\mathcal{O} = (\mathcal{C} - \mathcal{B}) + \mathcal{I}$, the decision-reversibility condition, and one short numerical witness. The full Bayesian and Blackwell proofs belong here so that the trial extension does not displace the complete-information consolidation, taper, and withdrawal results.*

G Monetary, Mission-Weight, and Numerical Robustness

G.1 Positive rescaling and the limit of unit invariance

Proposition G.1 (Scale invariance of value rankings)**.** *Multiply every current payoff and the withdrawal payoff by a common scalar $a > 0$, leaving transition kernels, feasible actions, and β unchanged. Then the Bellman value becomes aV, every optimal action and stopping decision is unchanged, all state thresholds are unchanged, and every policy-fixed or optimized Learning Dividend is multiplied by a.*

Proof. If V solves the original Bellman equation, every transformed action return evaluated at aV is a times its original return, and the transformed stopping branch is aW. Hence aV is the unique fixed point of the transformed contraction. Positive multiplication preserves all rankings. □

If the separately reported resource ledger is expressed in the same normalized value unit and is also multiplied by a, CDI is multiplied by a while its denominator is unchanged. In physical currency units, by contrast, a mere change of the utility numeraire would not rescale the physical cost ledger. Positive multiplication is the only invariance claimed. Adding a constant to each active-period reward is generally *not* neutral because policies induce different active durations and withdrawal truncates the stream.

G.2 Reproduction and theorem-assumption audit

The reproduction script rebuilds the numerical model and adds separate audit layers. It rebuilds the fast paths, Markov-reward distributions, joint (J^+, ℓ^+) kernels, Bellman values, fixed-policy evaluations, and policy accounts. It then checks probability mass, inherited event order, product-order kernel inequalities, value monotonicity, stopping shape, taper shape, scaling, alternative mission weights, learning parameters, durability definitions, and matched numerical discretizations. A supplementary audit script added with this revision reports, in addition, the terminal coarse-state marginals of both kernels, the cross-state kernel comparison on the common actions, and the own-kernel product-order audit over the complete family of 1,376 nontrivial upper sets.

Table 11: Reproduction of the headline numerical solution

Quantity	Original calculation	Independent audit
Structural threshold $\ell^\dagger$	0.23760	0.237603
Taper-onset grid point $\widehat{\ell}_T$	0.26	0.26
Withdrawal grid point, learned economy $\widehat{\ell}_W$	0.64	0.64
Withdrawal grid point, clamped-frozen economy $\widehat{\ell}_W^F$	0.24	0.24
Policy-fixed $\mathrm{LD}_R^{\pi_F^*}(0.06)$	5.786	5.78643
Adaptation premium	3.052	3.05176
Optimized $\mathrm{LD}_R^*(0.06)$	8.838	8.83818
Initial value $V_R(0.06)$	11.662	11.66228
Learned / frozen Bellman iterations	68 / 358	68 / 358

Table 12: Finite-grid audit of assumptions and solved policy shape

Object	Checks	Violations	Reading
Probability mass, R and A kernels	–	$< 1.6 \times 10^{-15}$ max error	Kernels are numerically stochastic.
Inherited p_A^R, p_F^R, r_H^A legibility order	1,050 each	0	Matched-condition event order is realized.
Terminal-A marginal in legibility	900	0	Nondecreasing at every maintained action; strictly increasing except at the zero-support R kernel (identically zero). The same-state benchmark's marginal invariance fails.
Product-order upper sets, R kernel	756,800	0	Finite-grid analogue of kernel clause (i) of Assumption 4.4 holds over the complete upper-set family.
Product-order upper sets, A kernel	481,600	0	Finite-grid analogue of kernel clause (i) of Assumption 4.4 holds over the complete upper-set family.
Cross-state dominance $\mathcal{K}_A^u \succeq \mathcal{K}_R^u$, common actions	491,232	149	Holds at every grid point up to and including the first withdrawal point $\widehat{\ell}_W = 0.64$; all violations lie at $\ell \in \{0.92, 0.94, 1.00\}$, maximum deficit 0.0041.
Common-action payoff order $r_A \geq r_R$	357	0	Holds with minimum margin 0.00705; the payoff component of the primitive route.
Monotonicity of V_R, V_A, W	50 differences each	0	Solved values are nondecreasing and satisfy the cross-state conclusion, $\min(V_A - V_R) = 1.015$; the operator-invariance clause itself is maintained, not verified.
Global strict decline of Δ_A	50 differences	11	Not verified; the largest local rise is 0.1076.
Strict decline from $\ell = 0.22$ onward	39 differences	0	The 51-point baseline has one decreasing upper tail and one sign crossing.
Adjacent global decreasing differences of Q_A	192	32	The clean sufficient theorem is not numerically verified globally.
Pairwise action single crossing	21 action pairs	0	Clause (i) of the finite-action route is verified; the strict clause and the maximum selection are not separately audited.
Upward steps in selected $u_A^*(\ell)$	50 differences	0	The solved post-entry action is nonincreasing.

The product-order audit instantiates, on the solved kernels, the finite-grid analogue of clause (i) of Assumption 4.4; it is a diagnostic, not a proof of the continuum clause. The audit ranges over all 1,376 nontrivial upper sets of the 2×51 product lattice, including the pure-A tail sets $\{(A, \ell) : \ell \geq t\}$; the terminal-marginal row is the $t = 0$

member of that tail family. The solved gap $\min(V_A - V_R) = 1.015$ is the cross-state conclusion of Proposition 4.5; the operator-invariance clause itself remains a maintained sufficient condition and is not verified for every value function. The primitive inclusion route of Appendix A is doubly unavailable at the global level: $\mathcal{U}_R \subseteq \mathcal{U}_A$ fails, with domain maxima 0.25 and 0.15, and the cross-state row shows the kernel-comparison component holding at every grid point up to and including the first withdrawal point. The superseded same-state benchmark of Appendix A cannot hold on this kernel: its destination-marginal invariance is contradicted directly by the terminal-marginal row; the moving first-event summaries are economically consistent with that movement but are not, by themselves, proof of it. The complete-family counts in the two product-order rows, together with the terminal-marginal and cross-state rows, are produced by a standalone supplementary audit script distributed with the source package; the original reproduction script retains the both-parts-nonempty subfamily (729,300 and 464,100 comparisons, also with zero violations), and the supplementary script writes a separate result file while modifying no previously distributed file.

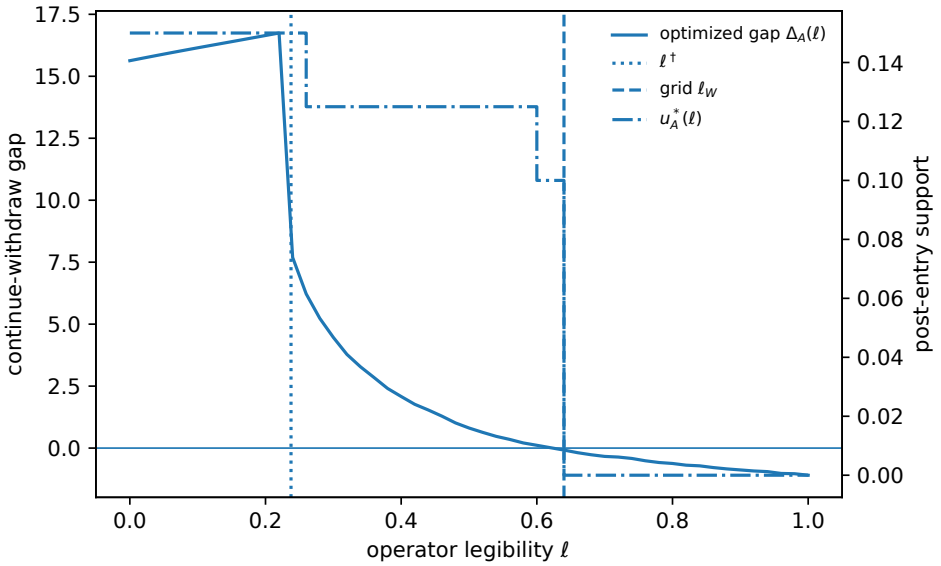

Figure 8: Gap and policy audit. The optimized gap rises at low legibility, remains positive, then declines through one upper crossing. The selected support action is nonincreasing. Vertical markers show the structural threshold and the first withdrawal point on the disclosed grid.

G.3 Threshold precision and matched discretization

On the 51-point legibility grid, the first withdrawal point is $\widehat{\ell}_W = 0.64$ and the sign-changing bracket is $[0.62, 0.64]$. Linear interpolation of the two neighboring grid gaps gives 0.623246 as a diagnostic only. Likewise, the first observed departure from the active ceiling is $\widehat{\ell}_T = 0.26$ with bracket $[0.24, 0.26]$. These quantities are reported as grid points rather than continuum estimates.

Table 13: Matched grid and Bellman-solver robustness

Specification	$\widehat{\ell}_T$	$\widehat{\ell}_W$	$\widehat{\ell}_W^F$	$\mathrm{LD}_R^*(0.06)$	$V_R(0.06)$
baseline	0.26	0.64	0.24	8.838	11.662
coarse economic action grid	0.46	0.64	0.24	8.825	11.649
looser Bellman tolerance	0.26	0.64	0.24	8.838	11.662
tighter Bellman tolerance	0.26	0.64	0.24	8.838	11.662
fine exposure grid	0.26	0.64	0.24	8.823	11.689

The coarse action grid leaves the withdrawal boundary at 0.64 but moves the first visible taper departure to 0.46 because the 0.125 action is absent. The Bellman tolerance changes iteration counts but not displayed results. The fine exposure-grid check compares a fine learned economy with a fine clamped-legibility economy. The matched comparison gives

$$\mathrm{LD}_R^*(0.06) = 8.823390.$$

The withdrawal grid point remains 0.64 and the Learning Dividend moves by less than 0.2 percent.

G.4 Alternative mission weights and discounting

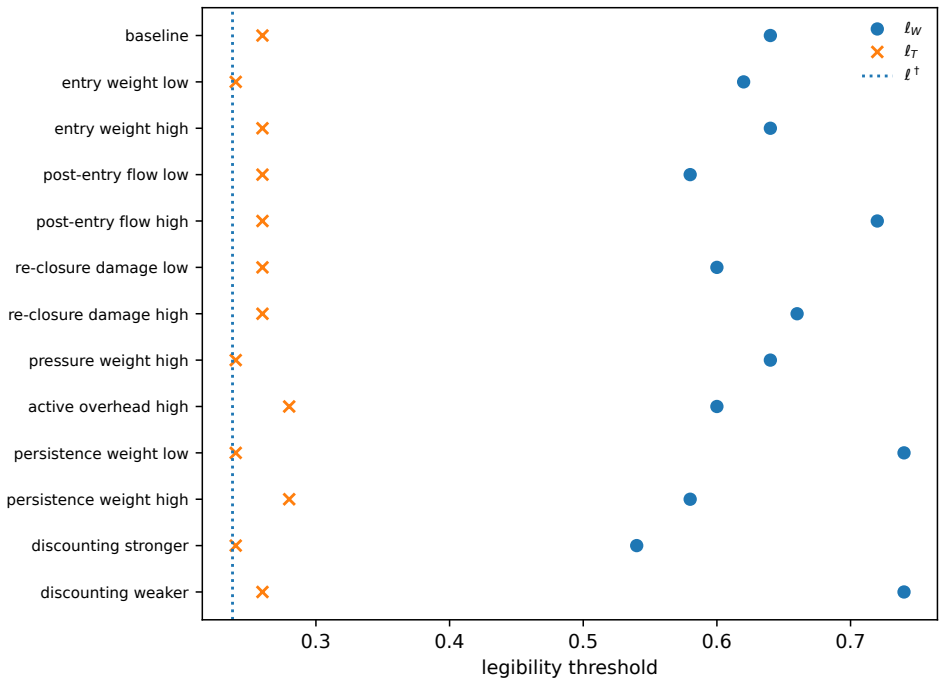

Figure 9: Mission-weight sensitivity of the taper and withdrawal grid points. The structural threshold is fixed by the upstream transition technology; the economic thresholds move with valuation and discounting.

Table 14: Selected mission-weight and discount-factor sensitivity

Scenario	$\widehat{\ell}_T$	$\widehat{\ell}_W$	$\mathrm{LD}^*_R(0.06)$	Active episodes	CDI
baseline	0.26	0.64	8.838	9.79	0.659
post-entry flow low	0.26	0.58	10.327	8.73	0.579
post-entry flow high	0.26	0.72	7.546	11.73	0.797
re-closure damage low	0.26	0.60	10.276	9.07	0.607
re-closure damage high	0.26	0.66	7.464	10.25	0.695
pressure weight high	0.24	0.64	8.879	9.79	0.658
active overhead high	0.28	0.60	9.640	9.07	1.042
persistence weight low	0.24	0.74	5.252	12.31	0.830
persistence weight high	0.28	0.58	12.597	8.73	0.579
discounting stronger	0.24	0.54	3.946	8.04	0.571
discounting weaker	0.26	0.74	23.055	12.30	1.012

The directions are conditional on the maintained technology. Raising the post-entry flow weight moves the economic withdrawal boundary outward and lengthens active duration. A higher active overhead moves withdrawal inward and raises cost per durable withdrawal. Raising the value assigned to autonomous persistence raises W and therefore moves withdrawal inward. Stronger discounting moves the boundary inward; greater patience moves it outward in this specification. The pressure-weight change barely moves $\widehat{\ell}_W$ because it affects current learned and frozen returns jointly. These are gap-shift illustrations, not universal signs attached to parameter names.

G.5 Signed learning under bounded and projected adverse cases

The state-invariance condition is pathwise: $\eta Y + \delta_\ell + \chi G \leq 1$. Because $Y, G \leq 1$, the simpler inequality $\eta + \delta_\ell + \chi \leq 1$ is a global sufficient condition, not a necessary one. The baseline satisfies it. The high-capture stress case does not; its numerical implementation projects ℓ^+ back to $[0, 1]$, so it is reported separately as a projected robustness case.

Table 15: Learning-parameter sensitivity and negative-Learning-Dividend regions

Scenario	η	δ_ℓ	χ	Global bound	First negative LD	$\widehat{\ell}_W$	Min. LD	CDI
baseline	0.22	0.02	0.70	True	0.80	0.64	-0.317	0.659
lower formation	0.12	0.02	0.70	True	0.66	0.50	-0.329	47.754
lower capture	0.22	0.02	0.30	True	0.82	0.64	-0.295	0.657
bounded adverse 1	0.05	0.02	0.90	True	0.38	0.36	-0.351	–
bounded adverse 2	0.01	0.02	0.97	True	0.24	0.24	-0.645	–
projected high-capture	0.05	0.02	12.00	False	0.28	0.30	-1.011	–

Two bounded adverse cases satisfy the global sufficient no-overshoot condition and still generate negative optimized Learning Dividends. With $(\eta, \delta_\ell, \chi) = (0.05, 0.02, 0.90)$, the first negative value occurs at $\ell = 0.38$ and $\widehat{\ell}_W = 0.36$. With $(0.01, 0.02, 0.97)$, the sign reversal occurs at 0.24 and withdrawal also begins at 0.24. Thus the signed counter-result does not depend on clipping. The projected high-capture case remains useful as an extreme stress test, but it belongs outside the unclipped admissible baseline.

G.6 Durability-denominator sensitivity

Table 16: Sensitivity of Cost to Durable Independence to the certification rule

Criterion	Persistence cut	Re-closure cut	First certified ℓ	Durable probability	Resource cost	CDI
looser	0.50	0.60	0.38	1.000	0.659	0.659
baseline	0.55	0.58	0.56	1.000	0.659	0.659
stricter	0.60	0.55	0.96	0.000	0.659	–

The looser and baseline criteria certify states below the economic withdrawal boundary, so the same optimal policy achieves durable withdrawal with probability essentially one. The stricter criterion first certifies at $\ell = 0.96$, above the policy's withdrawal boundary; durable probability is then zero and CDI is undefined, not zero. The ratio is therefore useful only together with a disclosed durability standard. This is denominator discipline, not a claim that the baseline cutoffs are empirical standards.

G.7 Numerical scale check

Table 17: Positive scale invariance in the numerical model

Scale	$\widehat{\ell}_T$	$\widehat{\ell}_W$	LD*	LD*/scale	Policy mismatches
0.10	0.26	0.64	0.884	8.838	0
0.50	0.26	0.64	4.419	8.838	0
1.00	0.26	0.64	8.838	8.838	0
2.00	0.26	0.64	17.676	8.838	0
10.00	0.26	0.64	88.382	8.838	0

Across factors from 0.1 to 10, both learned and frozen policy arrays are identical to the baseline. Values, Learning Dividends, and the normalized resource ledger scale linearly. This supports the use of normalized cardinal units while leaving relative mission weights and the durability denominator economically substantive.

G.8 Robustness conclusion

The numerical conclusions are narrow and auditable: the code reproduces the disclosed headline solution; the transition kernels satisfy the finite-grid event and product-order diagnostics; the solved values and selected support policy are monotone; the stopping set is one upper interval; the weak single-crossing clause, rather than global decreasing differences, supports the finite-grid taper of the reported minimum selection; negative learning survives bounded no-overshoot stress cases; and the matched fine-exposure comparison changes no qualitative conclusion. None of these statements is an efficacy estimate, a monetary valuation, or a proof that every continuum sufficient condition holds.