Schedules and Prioritization: A Behavioral Foundation for Multi-Armed Bandits and Stopping Problems

Jaden Yang Chen* Can Urgun†

June 8, 2026

Abstract

Bandit models typically begin with arms, states, rewards, and transition rules. This paper instead begins with preferences over stopped local contingent schedules: possible unfoldings of a responsibility, project, experiment, or opportunity in its own local time. Behavioral axioms on single schedules characterize a generalized stopping representation with current utility, local discounting, and a broad continuation aggregator. A common-tail compensation axiom then allows calendar time to be priced across schedules. Imposing a tight elapsed-calendar constraint generates a rested generalized bandit and yields index optimality: the index is the shadow price of advancing a local clock. Expected-utility, learning, robust, rank-dependent, Choquet, and Pandora models arise as special cases.

Keywords: contingent schedules, recursive preferences, stopping, ambiguity, bandits, index policies.

*UNC Chapel Hill. Email: yangch@unc.edu.
†UNC Chapel Hill. Email: curgun@unc.edu.

1 Introduction

A person has many strands of responsibility in her life. There may be a work strand: answer a referee report, revise a draft, wait for feedback, decide whether the next step is a small edit or a larger rewrite. There may be a child-related strand: take a child to school, see how the school day went, check homework, learn from a report card that mathematics needs attention, decide whether to help directly or arrange outside help. There may be a household or caregiving strand: make dinner, schedule a repair, follow up with a family member, learn that a small task has generated a larger one. There may be a relationship strand: plan something thoughtful, respond to how it is received, and decide what kind of further attention is called for. These examples are deliberately ordinary. The point is that a strand of responsibility is not a single dated payoff. It is a contingent sequence of consequences.

The larger problem is to fit many such strands into one calendar. A day has only so many slots. Work, household, family, and relationship responsibilities cannot all be advanced at once. But neither are they consumed and discarded like independent goods. If a work file is put aside while a family matter is handled, the work file remains there. If a household repair is postponed while a deadline is met, the repair does not vanish. Each strand has its own internal time: the first time it is advanced, the second time it is advanced, the next time something is learned, and the point at which that strand is completed or abandoned. Calendar time is the scarce resource that determines which of these local clocks is allowed to move.

The reason to begin with these mundane strands is the same reason consumer theory begins with consumption bundles. Demand is a workhorse object of economics, but a demand function is not usually taken as primitive. One starts with preferences over the right domain, obtains a utility representation from behavioral restrictions, and then derives demand by imposing a budget constraint. The domain and the constraint do different jobs. Bundles are the objects the consumer can rank; the budget is the scarcity constraint that turns those rankings into demand.

This paper develops an analogous foundation for bandit problems. Bandits are used throughout economics to study experimentation, learning, search, attention, project selection, medical trials, and dynamic resource allocation. Yet the standard formulation usually begins one step downstream. The analyst gives the decision maker a collection of arms, specifies states, rewards, transition rules, and an objective, and then asks how the arms should be operated. That is often the right applied model. It is not, however, a behavioral foundation for why such a bandit problem is the right representation of the decision maker's situation.

Our starting point is that, in many economic settings, what the decision maker sees is not a Markov machine called an arm. She sees a responsibility, project, experiment, or opportunity that may unfold in different ways as attention is devoted to it. The primitive alternatives are therefore possible unfoldings of one strand of responsibility. We call these objects local contingent schedules. They are local because they are written in the internal time of the strand, before the strand is placed into a common calendar with other strands. They are contingent because later consequences may depend on what has happened so far. They are stopped because, along every

realized branch, the strand eventually ends.

The first part of the paper studies one strand in isolation. Before asking whether work, family, household, or some other responsibility should receive the next unit of calendar time, we ask how the decision maker evaluates possible continuations of a single responsibility. We impose two groups of behavioral restrictions. The deterministic-stream axioms apply when the strand has no remaining branching. They identify the ordinary intertemporal part of the evaluation: a common scale for current consequences and a common discount factor over local time. The recursive one-step axioms apply when the strand has a current consequence and a contingent folder of future continuations. They require that equivalent continuations can be substituted, that branchwise improvements are respected, and that current consequences and continuation prospects can be traded off in a stable way.

The first main theorem says that these restrictions are equivalent to a recursive stopping representation. This is a strong equivalence, not merely a sufficient condition for a convenient formula. Recursive representations are fairly common in decision theory. The novelty is that, on our stopped-schedule domain, the usual behavioral restrictions generate stopping problems. The representation says that advancing a strand delivers some utility today, future consequences are discounted in local time, and the remaining uncertainty is evaluated by an aggregator over the possible continuation values. That aggregator is deliberately broad: in special cases it becomes expected utility, multiplier or variational preferences, rectangular max–min, rank-dependent or Choquet preferences, and other evaluations of future branches.

This is the first sense in which the paper is foundational. Stopping theory usually begins with a process, a payoff, and a set of admissible stopping rules. The analyst writes down what the state is, how it evolves, and what the decision maker receives upon stopping. Here, the stopped tree of consequences is the primitive alternative. The theorem identifies when preferences over those trees can be treated as if the decision maker were solving a stopping problem. Thus the stopping problem is not imposed at the outset. It is represented from preferences over stopped schedules.

This one-strand representation is the analogue of utility in consumer theory. It tells us how the decision maker evaluates each primitive object. But utility alone is not demand, and a stopping representation alone is not yet a bandit. The second step is to add the relevant scarcity constraint. For consumers, the constraint is a budget. For responsibility strands, the constraint is the calendar: only one local clock can move at a time.

Before calendar time can be priced, however, one more behavioral issue must be addressed. A priority index is a scalar object. It compares the claims that different strands have on the next unit of calendar time. For this comparison to be meaningful, scalar charges and compensations must pass through the evaluation of continuation folders without being distorted. This is automatic under ordinary expected utility, but it is not automatic when continuations are evaluated nonlinearly. Ambiguity, robustness, or rank dependence can make a common shift in all future branches fail to behave like the same common shift after aggregation.

We capture the needed property with a common-tail compensation axiom. Suppose two ways

of handling the beginning of a strand are exactly balanced. Now append the same downstream continuation behind both of them. Since the appended continuation is common, it should not alter the original compensation. In the responsibility-strand language, if two ways of opening a file are equally good before the same future file is attached, attaching that same future file should not make one opening suddenly better than the other. The axiom is behavioral, but its economic role is to make scalar shadow prices legitimate inside local schedule problems.

This is the bridge from stopping to bandits. Once scalar calendar-time charges can be moved through the continuation evaluation, represented preferences over local schedules can be combined with a tight calendar constraint. The elapsed-calendar constraint says that local clocks cannot in aggregate run faster than calendar time. Each calendar period advances one strand, while the others remain fixed. This is the scarce-resource constraint that turns many local responsibility schedules into one prioritization problem.

The central result is the Bandit Representation and Index Theorem. Its first part says that the elapsed-calendar problem has the form of a generalized rested multi-armed bandit. The state of the decision maker is the collection of current local states, one for each open strand. Advancing one strand changes only that strand's local state; all other strands are frozen. The arms of the bandit are therefore not primitive Markov machines. They are represented responsibility strands placed under a common calendar constraint.

The second part says that the optimal policy is an index policy. The index of a strand at a given local state is the critical shadow price of calendar time at which that strand is just worth putting on hold. This is a priority price, not an additional behavioral primitive. A strand with a high index is one for which the next local step is expensive to postpone, because the decision maker would still be willing to advance it even when calendar time is priced highly. The optimal rule is to advance a strand with the highest current index. Thus index optimality is derived from the domain, the axioms, and the calendar constraint, rather than added as an independent rule.

The paper then shows how familiar models arise from additional behavioral or domain restrictions. Under ordinary independence for continuation schedules, the continuation aggregator is linear. The local schedule is then represented by a subjective transition law, and the index reduces to the classical Gittins–Jones index. A learning model is obtained by treating beliefs as part of the local state. Advancing a strand produces both a current consequence and a signal about what kind of strand it is. Resting the strand leaves its belief fixed, because no local signal arrives while attention is elsewhere.

With variational or multiplier-style preferences, the decision maker is not willing to evaluate a continuation folder under a single forecast. Instead, alternative probabilistic descriptions are considered, with penalties attached to less plausible descriptions. Max–min preferences are the stark limiting case. There, the relevant adverse description is not merely a single bad one-step transition chosen today. It is closer to a series of unfortunate events: an unfavorable rectangular description of how the strand may continue to evolve is selected against the contemplated stopping plan. The resulting index is a robust priority price for advancing the strand under dynamic

ambiguity.

With rank-dependent or Choquet preferences, the decision maker evaluates continuation branches by their rank rather than only by additive probabilities. This is natural in the schedule domain. A work strand may command attention because a very promising response would create a valuable continuation; a school matter may command attention because a bad report would require immediate intervention; a household repair may loom large because one branch turns it into a serious problem. In a static lottery this looks like rank dependence over prizes. In a continuation problem it looks closer to fear of missing out: the cost of putting a strand on hold includes the possibility of failing to reach a salient future branch.

Pandora's problem arises in a different way. It is not primarily a different attitude toward continuation uncertainty. It is a restriction on the schedule domain. A box is a strand whose uncertainty is resolved in one local step. Once the box is opened, the revealed value is absorbing. Under this domain restriction, the general index becomes a reservation value. Thus Pandora's rule is another instance of the same schedule logic: a special kind of responsibility strand generates a special kind of priority price.

The unifying message is that bandit models can be built from the same ingredients that underlie other economic choice problems: a primitive domain, behavioral restrictions that yield a representation, and a scarcity constraint. A transition law, an ambiguity set, a capacity, or a reservation value is not the starting point. Each appears after adding further restrictions to preferences or to the schedule domain. Contingent schedules therefore play for bandits the role that consumption bundles play for demand: they are the primitive domain on which behavioral restrictions are imposed before scarcity is introduced.

The paper is organized as follows. The next subsection briefly reviews the related literature. Section 2 defines local contingent schedules and the one-step composition operation. Section 3 states the deterministic-stream and recursive one-step axioms and proves the recursive stopping representation. Section 4 introduces common-tail compensation and its scalar-charge implication. Section 5 establishes the Bandit Representation and Index Theorem. Section 6 develops the subjective expected-utility, learning, variational, max–min, rank-dependent, Choquet, and Pandora special cases. Section 7 concludes. All proofs are relegated to the appendix.

1.1 Literature

Bandit models are one of the standard ways economists represent dynamic allocation under uncertainty. They appear in experimentation, learning, search, attention, project choice, medical trials, and other settings in which choosing an alternative today both produces a payoff and changes what can be learned or done tomorrow. Early economic applications include pricing under demand uncertainty and strategic experimentation [Rothschild, 1974, Bolton and Harris, 1999, Keller, Rady, and Cripps, 2005]; Bergemann and Välimäki [2008] surveys the experimentation literature. The classical Gittins–Jones theory gives the main solution result once the bandit problem has been specified: a discounted allocation problem with projects that evolve only

when selected can be reduced to single-project stopping problems, and an optimal policy selects a project with maximal index [Gittins and Jones, 1974, Gittins, 1979, Whittle, 1980, Berry and Fristedt, 1985, Katehakis and Veinott, 1987, Weber, 1992, Gittins, Glazebrook, and Weber, 2011]. These papers solve bandits once the relevant objects are already in place. The decision maker is handed arms governed by known stochastic processes, and in learning applications she is handed an information structure that determines which observations arrive when an alternative is tried and how beliefs update. Our question is prior to that specification. In the same way that revealed-preference theory disciplines demand by starting from preferences over a domain of objects rather than from an arbitrary demand curve [Samuelson, 1938, Afriat, 1967, Varian, 1982], we ask what primitive domain and what refutable behavioral restrictions make a dynamic choice problem representable as a bandit. In our model, the primitive objects are local stopped schedules; the bandit appears only after those schedules have a stopping representation and are placed under one elapsed-calendar constraint.

The same issue appears already at the level of stopping. Optimal stopping problems are ubiquitous in economics and finance: agents decide when to invest, exercise, liquidate, abandon, inspect, or accept an alternative. Classical stopping theory begins with a stochastic process and asks when it should be stopped, with Snell-envelope and martingale methods providing the standard tools [Chow, Robbins, and Siegmund, 1971, Shiryaev, 1978, Peskir and Shiryaev, 2006]; real-options models use the same structure to study investment timing under uncertainty [Dixit and Pindyck, 1994]. Search models in the Pandora tradition combine ordering and stopping: Weitzman [1979] gives the classical reservation-value rule, and subsequent work studies how far Pandora-style rules extend beyond the original environment [Olszewski and Weber, 2015, Doval, 2018]. A particularly relevant recent contribution is Auster and Che [2025], who revisit Pandora's problem under robust search and connect robust search behavior to choice overload. These papers begin with processes, boxes, inspection costs, or reservation equations. Our single-strand theorem begins instead with a stopped object in the decision maker's domain and identifies when preferences over such objects have a stopping representation. Pandora's problem then appears as a special schedule domain, in which opening a box resolves uncertainty in one local step and the revealed value is absorbing.

A further literature takes the bandit object as given and changes the way uncertainty is evaluated. Anderson [2012] shows how ambiguity aversion can reduce the value of experimentation. Li [2019] develops a multiple-priors bandit model and obtains Gittins–Jones-style results. Cohen and Treetanthiploet [2022] prove a Gittins theorem under nonlinear expectations. Robust and risk-sensitive bandit models modify the objective, reward process, or uncertainty set, and identify conditions under which index-like rules survive or fail [Kim and Lim, 2016, Caro and Das Gupta, 2022, Denardo, Park, and Rothblum, 2007, Malekipirbazari and Cavus, 2024]. This literature is useful for us because it makes clear that the location of uncertainty is substantive. If ambiguity is placed on the full multi-project environment, the adverse model may choose a joint description of all projects. Projects that are technologically separate in the classical bandit can then become

linked through the worst-case joint law. In our setting, this is not the natural interpretation of the underlying economic problem. A work obligation, a household repair, a medical follow-up, and a family responsibility may all compete for the same calendar, but uncertainty about one strand need not be tied to uncertainty about the others. We therefore allow rich ambiguity, robustness, and rank dependence within each local continuation problem, while keeping the cross-strand interaction exactly where the schedule interpretation puts it: in the scarce calendar. The product structure is not an independence assumption imposed on an already-formed multi-arm Markov process; it is the consequence of representing distinct local schedules separately before they are woven into one calendar.

Finally, the axiomatic part of the paper belongs to the tradition of dynamic decision theory and recursive preferences. Koopmans's work on stationary ordinal utility and impatience, and the Koopmans–Diamond–Williamson extension, provide classic foundations for stationary intertemporal utility and geometric discounting [Koopmans, 1960, Koopmans, Diamond, and Williamson, 1964]. Kreps and Porteus study temporal lotteries and the timing of uncertainty resolution, while Epstein and Zin develop recursive, not-necessarily expected-utility preferences over intertemporal consumption lotteries [Kreps and Porteus, 1978, Epstein and Zin, 1989]. Chew and Epstein axiomatize recursive utility under uncertainty and integrate atemporal non-expected-utility models into a dynamic framework [Chew and Epstein, 1991]. Related work shows how learning, information, updating, and ambiguity primitives can themselves be disciplined by choice: Dillenberger, Lleras, Sadowski, and Takeoka [2014] study subjective learning when information arrival is unobserved by the analyst, Dillenberger, Krishna, and Sadowski [2023] model subjective constraints on information acquisition and absorption, Denti [2022] gives testable restrictions for posterior-separable information costs, and Cheng [2022] axiomatizes relative maximum-likelihood updating for ambiguous beliefs. For ambiguity and nonadditive beliefs, the relevant antecedents also include max–min expected utility, Choquet expected utility, variational preferences, recursive multiple priors, smooth ambiguity, multiplier preferences, dynamic mixture aversion, adaptive preferences, dynamic α-MEU, and dual-self representations of ambiguity [Anscombe and Aumann, 1963, Gilboa and Schmeidler, 1989, Schmeidler, 1989, Maccheroni, Marinacci, and Rustichini, 2006a,b, Epstein and Schneider, 2003, Klibanoff, Marinacci, and Mukerji, 2005, Hansen and Sargent, 2001, 2008, Denti and Pomatto, 2022, Sarver, 2018, Sadowski and Sarver, 2024, Frick, Iijima, and Le Yaouanq, 2022, Chandrasekher, Frick, Iijima, and Le Yaouanq, 2022]. Our first representation theorem is deliberately in this recursive-preference tradition: a one-step schedule is evaluated by combining a current consequence with an aggregator of continuation values. The novelty is not recursion per se. The difference is the domain and the use of the representation. Standard recursive preferences are usually defined over global temporal consumption lotteries, streams, or acts; our primitives are stopped local schedules, or responsibility strands, whose clocks advance only when selected. This local-time domain is what makes a bandit interpretation possible. The common-tail axiom then supplies the central additional step: it implies cash additivity of continuation values, which allows elapsed calendar time to be priced by a scalar

charge. When many local schedules share one calendar clock, these scalar charges become indices, and advancing a maximal-index strand is optimal.

2 A Model of Contingent Schedules

We start by introducing a local contingent schedule, which broadly corresponds to an ordered flow of consequences generated by a strand of responsibility. We use one notational convention throughout. Local states are treated as consequence-bearing states. Formally, a schedule is a map from local histories into one-period consequences, so the state and the assigned consequence could be kept as separate objects. For simplicity we suppress that extra projection. Thus, a local state records the payoff-relevant situation of the strand, including the current one-period consequence generated at that local situation. When later formulas write $u(x)$ or $u(x_i)$, this means the utility of the current consequence attached to that local state. This convention is only a relabeling.[1]

For a single strand, let $S = \{1, \dots, n\}$ be a finite local state space. For an initial state $x \in S$, define

$$H_t(x) := \{(x_0, \dots, x_t) \in S^{t+1} : x_0 = x\}, \qquad H(x) := \bigcup_{t \geq 0} H_t(x).$$

The set $H_t(x)$ is the set of possible histories after t local steps of the strand. For an infinite local history $\omega = (x_0, x_1, \dots)$ with $x_0 = x$, write $\omega^t = (x_0, \dots, x_t)$. Let $\mathcal{K}$ be the space of one-period consequences, and let $\perp$ denote the terminal symbol indicating that the schedule has ended.

Definition 1 (Local contingent schedule)**.** *A local contingent schedule at x is a map*

$$s : H(x) \to \mathcal{K} \cup \{\perp\}$$

satisfying two restrictions.

(i) Absorbing nullity. *If $s(h_t) = \perp$, then $s(h_s) = \perp$ for every extension $h_s \supset h_t$.*

(ii) Finite termination. *For every infinite history ω starting at x, there exists a finite date $T_s(\omega)$ such that $s(\omega^t) \in \mathcal{K}$ for $t < T_s(\omega)$ and $s(\omega^t) = \perp$ for $t \geq T_s(\omega)$.*

Let $\mathcal{A}_x$ denote the set of schedules at x. A schedule specifies what consequence is received after each possible local history, until that branch ends. Absorbing nullity says that an ended branch stays ended. Finite termination says that, along each branch, only finitely many nonterminal consequences are generated. This is the domain on which the primitive preference relation $\succeq_x$ is imposed.

The finite-termination convention is behavioral rather than ontological. The axioms are imposed on stopped schedules because these are the finite objects over which choices can in principle be observed, compared, and refuted. This is the same discipline behind revealed-preference restrictions such as WARP: the primitive restrictions are placed on finite choice objects, even if the representation later assigns values on a larger mathematical domain. Thus the restriction

[1] D8 below fixes the common affine scale on which current consequences are compared; the objects that may remain state- and arm-dependent are the continuation operators.

to stopped schedules should not be read as saying that the decision maker cannot contemplate infinite streams or that the analyst cannot represent their values. Once the representation is obtained, finite schedules determine a natural extension to infinite deterministic streams and infinite-horizon recursive problems. That extension is generated by discounted limits and by concatenating continuation schedules at finite tails.

The point of writing the domain this way is that the uncertainty of a strand is not imposed on the decision maker as a fully specified Markov process or signal structure. What is primitive is the comparison between possible stopped trees of consequences. A work strand, for example, may branch depending on feedback; a child-related strand may branch depending on what is learned from school; a household strand may branch depending on whether a repair turns out to be minor or serious. The representation result asks when these comparisons can be summarized by familiar ingredients: a utility scale for current consequences, discounting across local time, and a continuation rule that captures both the possible next local situations and the decision maker's attitude toward the uncertainty in those continuations. In the expected-utility special case, that continuation rule is a subjective transition kernel; in the more general cases, it may be a nonlinear aggregator reflecting ambiguity, robustness, or risk-sensitive evaluation. Thus the usual stochastic process, transition kernel, or learning rule appears only after additional structure is imposed on schedule preferences, rather than being part of the primitive description [Rothschild, 1974, Bolton and Harris, 1999, Keller et al., 2005, Bergemann and Välimäki, 2008].

The one-period consequence space can be specialized according to the source of uncertainty. In the baseline formulation below, $\mathcal{K} = \mathcal{L} := \Delta(Z)$, the space of objective lotteries over prizes, and $0 \in \mathcal{L}$ is a null lottery. If ambiguity also concerns one-period rewards, $\mathcal{K}$ can instead be taken to be a space of horse lotteries or subjective one-period acts. The recursive representation is stated on the lottery-flow specialization.

A deterministic finite stream is a tuple $c = (\ell_0, \dots, \ell_{T-1}) \in \mathcal{L}^T$. Let

$$\mathcal{C} := \bigcup_{T \in \mathbb{N}_0} \mathcal{L}^T.$$

The deterministic schedule $d_x(c)$ assigns ℓ_t to every history-date node at local date $t < T$ and assigns $\perp$ at all later local dates. Thus a deterministic schedule is the same sequence of consequences on every branch of the local tree. Let

$$\mathcal{D}_x := \{d_x(c) : c \in \mathcal{C}\},$$

and write $d_x(())$ for immediate stopping. We normalize $d_x(()) \sim_x d_x((0))$ for every x.

For each $x \in S$, preferences over schedules starting at x are denoted by $\succeq_x$, with symmetric and asymmetric parts $\sim_x$ and $\succ_x$. The subscript records the local state at which the strand is evaluated. The same formal continuation may be evaluated differently when it is reached from different local states, because the continuation is attached to a different local situation.

2.1 One-step composition

The recursive structure is built from a simple operation: isolate the current consequence of a strand from the continuation that follows the next local state. Given a current lottery $\ell \in \mathcal{L}$ and continuation schedules $a^y \in \mathcal{A}_y$ for all next states $y \in S$, define

$$\Gamma_x(\ell, (a^y)_{y \in S}) \in \mathcal{A}_x$$

by assigning ℓ at the initial history (x) and, after a first transition to y, following a^y from that point onward:

$$\Gamma_x(\ell, (a^y)_y)(x, x_1, \ldots, x_t) = a^{x_1}(x_1, \ldots, x_t), \qquad t \geq 1.$$

A one-step schedule therefore has two parts: a current consequence and a contingent collection of future schedules, one for each possible next local state. Repeated applications of Γ_x build finite local trees and attach continuation schedules at their terminal histories. This is the local object from which the recursive representation is built.

The first result below asks what restrictions on preferences over such schedules make this recursive description possible. The answer will be that a single strand can be evaluated as a stopping problem. Only after that one-strand problem is understood do we return to the larger calendaring problem: several strands of responsibility, each with its own local time, must be woven into a single elapsed calendar.

3 Recursive Stopping Preferences

We now put a behavioral/axiomatic structure on preferences over a single strand. At this point there is still no calendar problem. The decision maker is not yet choosing between work, household, family, or other strands. She is evaluating possible ways in which one strand may unfold: different stopped trees of consequences, all written in that strand's own local time.

The first group of axioms concerns deterministic strands. These are schedules with no branching: the same finite sequence of consequences is obtained along every history. They are useful because they pin down the ordinary intertemporal tradeoffs that are already present before uncertainty over the continuation tree enters, in the spirit of Koopmans-style foundations for discounted utility [Koopmans, 1960, Koopmans et al., 1964]. The second group of axioms concerns one-step contingent schedules. These ask when the decision maker's comparison between "one consequence now" and "a contingent continuation from tomorrow onward" can be summarized recursively, as in the recursive-preference tradition [Kreps and Porteus, 1978, Epstein and Zin, 1989, Chew and Epstein, 1991].

Assumption 1 (Deterministic-stream axioms D)**.** *For each $x \in S$, the restriction of $\succeq_x$ to $\mathcal{D}_x$ satisfies the following conditions.*

D1. Weak order. *The relation is complete and transitive.*

D2. Mixture independence on common horizons. *For fixed T, streams $c, d, e \in \mathcal{L}^T$, and $\alpha \in (0,1)$,*

$$d_x(c) \succeq_x d_x(d) \quad \Longrightarrow \quad d_x\Big(\alpha c + (1-\alpha)e\Big) \succeq_x d_x\Big(\alpha d + (1-\alpha)e\Big),$$

where mixtures are coordinatewise.

D3. Same-length tail independence. *For fixed T, streams $c, d \in \mathcal{L}^T$, and tails $h, k \in \mathcal{C}$,*

$$d_x(c \oplus h) \succeq_x d_x(d \oplus h) \quad \Longleftrightarrow \quad d_x(c \oplus k) \succeq_x d_x(d \oplus k).$$

D4. Stationarity. *For all $c, d \in \mathcal{C}$,*

$$d_x(0 \oplus c) \succeq_x d_x(0 \oplus d) \quad \Longleftrightarrow \quad d_x(c) \succeq_x d_x(d).$$

D5. Zero-tail irrelevance. *For every $c \in \mathcal{C}$, $d_x(c \oplus (0)) \sim_x d_x(c)$.*

D6. Impatience. *There exist lotteries $\ell, m \in \mathcal{L}$ such that $d_x((\ell)) \succ_x d_x((m))$ and $d_x((\ell, m)) \succ_x d_x((m, \ell))$.*

D7. Continuity. *For each T, the restriction to $\mathcal{L}^T$ is continuous in the product topology.*

D8. State-independence on deterministic streams. *For all $x, y \in S$ and $c, d \in \mathcal{C}$,*

$$d_x(c) \succeq_x d_x(d) \quad \Longleftrightarrow \quad d_y(c) \succeq_y d_y(d).$$

These assumptions say that when a strand has no remaining uncertainty, the decision maker evaluates its finite flow of consequences in the usual discounted-utility way. D1 gives coherent rankings. D2 is the source of affine utility on fixed-horizon lottery streams. D3 says that if two front pieces are compared before a common tail, the particular common tail should not change the ranking of those front pieces. D4 says that adding a common blank local period at the beginning does not change a comparison, which is what delivers stationarity across local dates. D5 says that appending a blank period after a schedule has ended is not a new consequence. D6 rules out indifference to timing and pins the discount factor below one. D7 is continuity. D8 is the cross-state normalization: deterministic consequence streams are evaluated on the same one-period scale no matter which local state the strand is written from.

Proposition 1 (Discounted utility on deterministic finite streams)**.** *Under Assumption 1, there exist an affine vNM index $u : \mathcal{L} \to \mathbb{R}$ and a discount factor $\beta \in (0,1)$ such that deterministic schedules are represented by*

$$U_x(d_x(\ell_0, \ldots, \ell_{T-1})) = \sum_{t=0}^{T-1} \beta^t u(\ell_t) \qquad \text{for every } x \in S. \tag{3.1}$$

The pair (u, β) is unique up to positive affine rescaling of u.

The proposition pins down the one-period utility scale and the local discount factor on branchless schedules. We next turn to one-step schedules, where the current consequence is followed by a contingent folder of continuation schedules. Let

$$Q := \prod_{y \in S} (\mathcal{A}_y / \sim_y)$$

be the product of continuation equivalence classes. For $[a] = ([a^y])_{y\in S} \in Q$, define the induced one-step preference $\succeq_x^1$ on $\mathcal{L}\times Q$ by

$$(\ell,[a]) \succeq_x^1 (m,[b]) \quad \Longleftrightarrow \quad \Gamma_x(\ell,(a^y)_{y\in S}) \succeq_x \Gamma_x(m,(b^y)_{y\in S}).$$

The quotient notation emphasizes the idea that once the first local step is taken, a continuation schedule matters through how it is evaluated from the next local state onward.

Assumption 2 (Recursive one-step axioms R)**.** *For each $x \in S$, one-step comparisons satisfy the following conditions.*

R1. Conditional substitution of equivalents. *If $a^y \sim_y \tilde{a}^y$ for every y, then*

$$\Gamma_x(\ell,(a^y)_y) \sim_x \Gamma_x(\ell,(\tilde{a}^y)_y) \qquad \text{for every } \ell \in \mathcal{L}.$$

R2. Branchwise monotonicity. *If $a^y \succeq_y b^y$ for every y, then*

$$\Gamma_x(\ell,(a^y)_y) \succeq_x \Gamma_x(\ell,(b^y)_y) \qquad \text{for every } \ell \in \mathcal{L}.$$

R3. Weak order and continuity on one-step schedules. *The induced relation $\succeq_x^1$ is complete, transitive, and continuous on $\mathcal{L}\times Q$.*

R4. Coordinate independence. *Current lotteries and continuation profiles are separately independent:*

$$(\ell,q) \succeq_x^1 (m,q) \Longleftrightarrow (\ell,q') \succeq_x^1 (m,q') \qquad \forall \ell,m \in \mathcal{L},\ \forall q,q' \in Q,$$
$$(\ell,q) \succeq_x^1 (\ell,q') \Longleftrightarrow (m,q) \succeq_x^1 (m,q') \qquad \forall \ell,m \in \mathcal{L},\ \forall q,q' \in Q.$$

R5. Double cancellation. *If*

$$(\ell_1,q_1) \succeq_x^1 (\ell_2,q_2) \quad \text{and} \quad (\ell_2,q_3) \succeq_x^1 (\ell_3,q_1),$$

then

$$(\ell_1,q_3) \succeq_x^1 (\ell_3,q_2).$$

R6. Solvability. *Intermediate one-step values can be matched either by changing the current lottery while holding a continuation fixed or by changing the continuation while holding a current lottery fixed: if*

$$(\ell_1,q) \succeq_x^1 (\ell_2,q') \succeq_x^1 (\ell_3,q),$$

then $(\tilde{\ell},q) \sim_x^1 (\ell_2,q')$ for some $\tilde{\ell} \in \mathcal{L}$; and if

$$(\ell,q_1) \succeq_x^1 (m,q_2) \succeq_x^1 (\ell,q_3),$$

then $(\ell,\tilde{q}) \sim_x^1 (m,q_2)$ for some $\tilde{q} \in Q$.

The R axioms say that the decision maker's treatment of one local step is internally consis-

tent. R1 says that continuation schedules that are already equivalent from the relevant next state can be substituted branch by branch. R2 says that improving every possible continuation branch improves the one-step schedule. R3 gives coherent and continuous rankings on the one-step domain. R4 separates the ranking of current consequences from the ranking of continuation profiles: changing the common continuation should not reverse the ranking of two current consequences, and changing the common current consequence should not reverse the ranking of two continuation profiles. R5 is the corresponding consistency of compensations between current and future parts of the strand. R6 is the richness condition that lets those compensations be found inside the primitive domain.

The result is not that the strand was secretly an expected-utility Markov process all along. The result is that, under D and R, the decision maker behaves as if a single strand can be evaluated recursively: current utility is evaluated on the scale identified above, and the future tree is summarized by a continuation operator. That continuation operator may later reduce to a subjective transition law, a set of priors, a variational penalty, or another ambiguity-sensitive object; at this level it is simply the way the decision maker evaluates the possible continuations of the strand.

Theorem 1 (Recursive stopping representation)**.** *Suppose the interval-richness required for solvability holds on the relevant attainable sets. Then preferences satisfy Assumptions 1 and 2 if and only if there exist numerical representations $U_x : \mathcal{A}_x \to \mathbb{R}$, a common affine one-period utility index $u : \mathcal{L} \to \mathbb{R}$, a common discount factor $\beta \in (0,1)$, and continuous increasing continuation operators*

$$\rho_x : \prod_{y \in S} U_y(\mathcal{A}_y) \to \mathbb{R}$$

such that deterministic schedules satisfy (3.1) and every one-step schedule satisfies

$$U_x\Big(\Gamma_x(\ell, (a^y)_{y \in S})\Big) = u(\ell) + \beta \rho_x\Big(U_1(a^1), \ldots, U_n(a^n)\Big), \tag{3.2}$$

where the one-period index is the same affine scale as in (3.1).

The theorem is the first main reduction. Primitive preferences over trees of consequences are equivalent to a recursive evaluation of a single schedule. The current consequence enters through the common utility index u. The continuation enters only through the vector of continuation values and the state-dependent aggregator ρ_x. Thus the theorem separates what is common across different strands of responsibility and states, namely the utility scale for current consequences, from what may vary with the local schedule corresponding to a strand of responsibility, namely the evaluation of continuation uncertainty.

The theorem can be read as a schedule-accounting result. The proof first fixes the ruler for a single strand when nothing is uncertain. For a branchless local file, the D axioms say that the value cannot depend on irrelevant descriptions of the same stopped schedule, only on which consequences occur at which local dates. They therefore calibrate every deterministic local schedule on the common discounted scale in Proposition 1. This is the calibration that must also be used once the strand becomes uncertain.

Now consider a one-step uncertain schedule. In the schedule language, it is a card with two fields: the consequence produced when the file is advanced now, and the contingent folder of schedules left for the next local state. Rather than adding substantive assumptions about probabilities, the R axioms impose a bookkeeping discipline on this card. They ensure that equivalent continuation folders can be substituted, that branchwise improvements are respected, and that current consequences and continuation folders have stable compensating tradeoffs. In other words, the current field and the continuation field can be separated without changing the underlying ordering of schedules. That separation is exactly what is needed for an additive representation of a one-step card:

$$A_x(\ell) + B_x\Big((a^y)_y\Big).$$

The deterministic representation then identifies the first coordinate. If all continuations are deterministic one-date schedules, the representation must agree with the already-calibrated value of a two-date deterministic schedule; therefore A_x is the same affine consequence index u, and the continuation coordinate is on the β-discounted scale. What remains unrestricted is how local state x summarizes the vector of possible continuation values. Calling that summary ρ_x gives

$$U_x\Big(\Gamma_x(\ell,(a^y)_y)\Big) = u(\ell) + \beta\rho_x\Big((U_y(a^y))_{y\in S}\Big).$$

Monotonicity and continuity of ρ_x are inherited from the corresponding schedule axioms. Thus the axioms turn a primitive preference over local schedules into a recursive local accounting rule: score what the strand delivers now, then add a discounted assessment of the file that remains.

The representation also explains why the word "stopping" is the right one, but it should be read locally. Classical stopping theory typically begins with a stochastic process and an admissible class of stopping times, and then asks when the process should be stopped [Chow et al., 1971, Shiryaev, 1978, Peskir and Shiryaev, 2006]. Here the order is reversed. The primitive object is a local schedule, and the axioms D and R are *equivalent* to a generalized stopping representation of preferences over such schedules. Stopping does not mean that the person has literally abandoned the real-world responsibility, or that no other strand can be attended to. It means that, in the isolated one-strand problem, the local schedule is no longer advanced and no further consequences from that schedule are counted. Later, in the multi-schedule problem, not advancing a strand at a calendar date is different from stopping it: the calendar may simply be used on another strand.

For a finite history $h = (x_0,\dots,x_t)$, write $\bar{x}(h) := x_t$ for its current local state. If $g = (x_t, y_1,\dots,y_m) \in H_m(x_t)$ is a continuation history starting from $\bar{x}(h)$, write $h \circ g := (x_0,\dots,x_t,y_1,\dots,y_m)$ for the history obtained by grafting g after h. In particular, for $y \in S$, write $hy := (x_0,\dots,x_t,y)$. Let $H(x) := \bigcup_{t\geq 0} H_t(x)$ be the set of finite local histories starting from x.

Corollary 1 (Induced stopping representation)**.** *Assume the hypotheses of Theorem 1, and use the normalization $u(0) = 0$. Under the consequence-bearing-state convention, the completed one-strand stopping value starting from x is represented by the minimal bounded nonnegative function $V_x : H(x) \to$*

$\mathbb{R}_+$ *satisfying*

$$V_x(h) = \max\left\{0,\, u(\bar{x}(h)) + \beta\rho_{\bar{x}(h)}\left(\left(V_x(hy)\right)_{y\in S}\right)\right\}. \tag{3.3}$$

Equivalently,

$$V_x((x)) = \sup_{\tau\leq\infty} \mathcal{U}_x(\tau), \tag{3.4}$$

where τ ranges over local stopping rules and $\mathcal{U}_x(\tau)$ is the completed recursive value of the schedule that advances the local history before τ and assigns $\perp$ thereafter.

Thus Theorem 1 turns the local evaluation of a strand into a generalized Snell envelope. At each finite local history, the decision maker compares ending the schedule, with normalized value zero, to advancing the schedule one more local step, receiving the consequence carried by the current local state, and recursively evaluating the continuation folder indexed by the next local state. The statement uses no exogenous probability space: the continuation folder is evaluated by the represented operator $\rho_{\bar{x}(h)}$.

4 Common Tails and Scalar Charges

The preceding section assigns a value to the schedule associated with a single strand of responsibility. This is already useful locally: it says how the decision maker evaluates one stopped tree of consequences when that schedule is considered on its own. The calendaring problem, however, will require something more. Several schedules will compete for the same unit of elapsed calendar time, and the value of moving one schedule forward must be comparable to the value of moving another schedule forward.

D8 gives a common scale for current consequences. It does not by itself say that scalar changes in continuation value remain scalar changes after the continuation tree is evaluated. This distinction matters because the representation in Theorem 1 permits nonlinear continuation operators. A continuation tree may be evaluated through expected utility, max–min preferences, variational preferences, smooth ambiguity, or some other uncertainty-sensitive rule. In such environments, a number added to every possible continuation branch need not automatically remain that same number after the continuation operator is applied.

For the index argument, the needed invariance is exactly this scalar portability. If one unit of calendar time is eventually priced by a scalar charge, then that scalar must enter a one-schedule stopping problem as a scalar. It cannot be reshaped by the continuation operator or depend on the rest of the calendar. The next axiom gives the primitive content of this requirement.

Assumption 3 (Common-tail compensation invariance, R7)**.** *For every $x \in S$, lotteries $\ell, m, \bar{\ell}, \bar{m} \in \mathcal{L}$, and continuation profiles*

$$r_y = (r_y^z)_{z\in S} \in \prod_{z\in S} \mathcal{A}_z, \qquad y \in S,$$

if

$$d_x((\ell, \bar{m})) \sim_x d_x((m, \bar{\ell})), \tag{4.1}$$

then

$$\Gamma_x\Big(\ell, (\Gamma_y(\bar{m}, r_y))_{y\in S}\Big) \sim_x \Gamma_x\Big(m, (\Gamma_y(\bar{\ell}, r_y))_{y\in S}\Big). \tag{4.2}$$

The axiom is easiest to read as a compensation test. In the two-date comparison (4.1), the decision maker is exactly indifferent between two ways of distributing consequences across today and tomorrow. One schedule gives ℓ today and $\bar{m}$ tomorrow; the other gives m today and $\bar{\ell}$ tomorrow. Thus the move from $\bar{m}$ to $\bar{\ell}$ is precisely the tomorrow-compensation for the move from ℓ to m today.

R7 says that this compensation remains exact when the same contingent continuation is attached behind tomorrow. The common continuation may be uncertain. It may branch in many ways. It may be evaluated nonlinearly. But it is the same continuation in both schedules, and so it should not change the exchange rate between current utility and a common one-period-later compensation. In the responsibility-strand language, if two ways of handling the first two local steps of a schedule are exactly balanced, then appending the same downstream continuation of the strand should not disturb that balance.

Proposition 2 (Common tails and cash additivity)**.** *Assume the recursive representation in Theorem 1. Fix $x \in S$. R7 is equivalent to cash additivity of ρ_x on all behaviorally attainable common scalar changes. In particular, whenever (4.1) holds and*

$$a^y := \Gamma_y(\bar{m}, r_y), \qquad b^y := \Gamma_y(\bar{\ell}, r_y), \qquad v_y := U_y(a^y), \qquad c := u(\bar{\ell}) - u(\bar{m}),$$

one has

$$\rho_x(v + c\mathbf{1}) = \rho_x(v) + c. \tag{4.3}$$

Consequently, if the attainable continuation domain is closed under common scalar translations, then R7 is equivalent to full cash additivity on that domain:

$$\rho_x(v + c\mathbf{1}) = \rho_x(v) + c, \tag{4.4}$$

whenever both v and $v + c\mathbf{1}$ lie in the domain of ρ_x.

R7 is the axiom that makes schedule values behave like scalar amounts when a common tail is attached. First ignore branching. The premise (4.1) says that two two-step schedules are exactly balanced: doing ℓ now and $\bar{m}$ next is just as good as doing m now and $\bar{\ell}$ next. Since deterministic schedules have already been put on the common discounted ruler, this balance means

$$u(\ell) + \beta u(\bar{m}) = u(m) + \beta u(\bar{\ell}).$$

If $c := u(\bar{\ell}) - u(\bar{m})$, then replacing $\bar{m}$ by $\bar{\ell}$ at the next local step is a c-unit improvement on the

continuation-value scale, and the current consequence ℓ is exactly the discounted compensation for using m instead.

Now put uncertainty back into the schedule by attaching the same downstream file behind that second local step. In every possible next local state y, compare the continuation schedule that begins with $\bar{m}$ and then follows r_y with the one that begins with $\bar{\ell}$ and then follows the same r_y. The only difference is the first consequence of that continuation file; the rest of the file is identical. Therefore the whole future folder has been shifted up by c in every branch, without changing its downstream structure.

R7 says that this common downstream structure cannot disturb the original compensation. Substituting the recursive representation into the conclusion of R7 gives indifference exactly when

$$u(m) + \beta c + \beta \rho_x(v) = u(m) + \beta \rho_x(v + c\mathbf{1}),$$

or equivalently

$$\rho_x(v + c\mathbf{1}) = \rho_x(v) + c.$$

Thus the common-tail axiom is precisely cash additivity on the scalar shifts that can be generated by primitive schedules. The converse is the same schedule story read backward: if a common c-shift passes through ρ_x unchanged, then the original two-step compensation remains exact after the same continuation file is appended. If the continuation domain contains all common scalar translations, the attainable version becomes full cash additivity.

This proposition is the bridge from one-schedule stopping values to calendar-time charges. Theorem 1 and Corollary 1 say that a schedule associated with one strand has a scalar stopping value. R7 says that scalar shifts of such values pass through the recursive continuation operator without distortion. That is the property needed when many local schedules are later placed into one calendar and elapsed calendar time is priced by a scalar charge.

In the classical index literature, especially Weber's charge formulation and Whittle's related subsidy formulation, the charge is not a new preference primitive; it is a scalar price attached to the local decision to continue or to leave a schedule inactive [Gittins and Jones, 1974, Whittle, 1980, Weber, 1992, Gittins et al., 2011]. Cash additivity is what lets that scalar remain a scalar inside each one-schedule stopping problem. Without it, local stopping representations may still exist, but a one-dimensional index need not summarize how different schedules should compete for calendar time.

5 Bandit Representation and the Index

We now return to the larger calendaring problem. The previous sections studied the schedule associated with a single strand of responsibility. They asked how one local schedule is evaluated, when it is worth advancing in its own local time, and when scalar changes in its continuation value remain scalar. The real allocation problem begins when several such schedules are open at

once. Work, household, family, and other strands may each have their own local state and their own local clock, but they must be fit into one elapsed calendar.

The passage from one schedule to many schedules adds a feasibility constraint, not a new preference axiom. Each local schedule is still evaluated by the one-schedule representation above. What changes is that local time is now scarce across schedules. A unit of calendar time can advance one local schedule; it cannot advance all of them at once. Thus the role played by a budget in a consumer problem is played here by an elapsed-calendar-time constraint: it describes which collections of local schedule steps can be afforded by a single calendar.

Fix N local schedules, indexed by $i = 1,\dots,N$. In the bandit terminology below, each local schedule is an arm. Schedule i has finite local state space S_i. The product state space is

$$\mathbf{S} := \prod_{i=1}^{N} S_i, \qquad x = (x_1,\dots,x_N).$$

For $x \in \mathbf{S}$ and $y_i \in S_i$, write

$$(x_{-i}, y_i) := (x_1,\dots,x_{i-1},y_i,x_{i+1},\dots,x_N).$$

The product construction is rested: if schedule i is advanced, then only x_i changes. All other local states remain fixed.

The calendar constraint is written in terms of local times. For each $n \in \mathbb{N}_0$, set

$$G^n := \left\{ a = (a_1,\dots,a_N) \in \mathbb{N}_0^N : \sum_{i=1}^{N} a_i = n \right\}, \qquad G := \bigcup_{n\geq 0} G^n.$$

For $a \in G^n$, let

$$D_a := \{a + \mathbf{e}_i : i = 1,\dots,N\} \subseteq G^{n+1}.$$

Definition 2 (Increasing path and elapsed calendar constraint)**.** *An increasing path is a sequence $\gamma = (\gamma_n)_{n\geq 0}$ with values in G such that*

$$\gamma_0 = 0, \qquad \gamma_{n+1} \in D_{\gamma_n} \quad \textit{for every } n \geq 0.$$

*Writing $\gamma_n = (T_1(n),\dots,T_N(n))$, an increasing path satisfies the **elapsed calendar time constraint** if*

$$T_i(0) = 0, \qquad T_i(n+1) - T_i(n) \in \{0,1\}, \qquad \sum_{i=1}^{N} \Big(T_i(n+1) - T_i(n) \Big) = 1. \tag{5.1}$$

Here $T_i(n)$ is the amount of local time assigned to schedule i by calendar date n. Equation (5.1) says that each calendar period advances exactly one local schedule. Equivalently,

$$\sum_{i=1}^{N} T_i(n) = n \qquad \text{for every calendar date } n.$$

This is the calendar-budget identity. The local schedules may have different local lengths and may

end at different histories, but their aggregate use of local time cannot exceed the calendar time available to the decision maker.

Definition 3 (Contingent multi-schedule)**.** *A contingent multi-schedule based at $x = (x_1, \dots, x_N)$ consists of one local contingent schedule for each schedule i, written in schedule i's own local time, and an increasing path γ satisfying the elapsed calendar constraint* (5.1)*. At each calendar date, the unique schedule whose local clock advances takes one step in its local schedule; all other local schedules remain at their current local states.*

This is the feasible set of the multi-schedule problem. The primitive alternatives remain local schedules. The increasing path is the feasible allocation of elapsed calendar time across those schedules. In this sense, the analogy with consumer choice is exact at the level of structure: preferences over primitive objects are one input, and a feasibility constraint is the other. The theorem below says that, once the primitive objects are local schedules and the feasibility constraint is elapsed calendar time, the resulting problem is a generalized rested bandit. The connection with the classical Gittins–Jones theorem is therefore through representation rather than through primitives: the object to which the usual single-project stopping logic applies is generated from schedule preferences and the calendar constraint, rather than assumed at the outset [Gittins and Jones, 1974, Gittins, 1979, Weber, 1992].

Theorem 2 (Bandit representation and index)**.** *For each schedule i, suppose the local schedule preferences satisfy D, R, and R7, and suppose the attainable continuation domain is closed under common scalar translations. Write the resulting schedule-i recursive representation on a common utility scale and common discount factor $\beta \in (0,1)$ as follows: advancing schedule i in local state x_i gives current utility $u(x_i)$, and the continuation vector over next local states is evaluated by an increasing cash-additive operator $\rho_i^{x_i}$ satisfying*

$$\rho_i^{x_i}(v + c\mathbf{1}) = \rho_i^{x_i}(v) + c \qquad \text{for every } v \in \mathbb{R}^{S_i},\ c \in \mathbb{R}.$$

Then the elapsed-calendar constrained contingent multi-schedule problem is a rested generalized multi-armed bandit. Its value $V : \mathbf{S} \to \mathbb{R}$ is characterized by the Bellman equation

$$V(x) = \max_{1 \le i \le N} \left\{ u(x_i) + \beta \rho_i^{x_i} \left(\left(V(x_{-i}, y_i) \right)_{y_i \in S_i} \right) \right\}. \tag{5.2}$$

For each scalar shadow price $\lambda \in \mathbb{R}$ of elapsed calendar time, define the one-schedule value $V_i^\lambda : S_i \to \mathbb{R}$ by

$$V_i^\lambda(x_i) = \max \left\{ \frac{\lambda}{1-\beta},\ u(x_i) + \beta \rho_i^{x_i} \left(\left(V_i^\lambda(y_i) \right)_{y_i \in S_i} \right) \right\}. \tag{5.3}$$

Let

$$D_i(\lambda) := \left\{ x_i \in S_i : V_i^\lambda(x_i) = \frac{\lambda}{1-\beta} \right\}$$

be the set of local states in which schedule i is put on hold at shadow price λ, and define the index

$$\Lambda_i(x_i) := \inf\{\lambda \in \mathbb{R} : x_i \in D_i(\lambda)\}. \tag{5.4}$$

The on-hold sets are nested:

$$\lambda' \geq \lambda \quad \Longrightarrow \quad D_i(\lambda) \subseteq D_i(\lambda') \qquad \text{for every } i.$$

Hence $\Lambda_i(x_i)$ is well defined. An optimal policy for the elapsed-calendar constrained contingent multi-schedule problem is to advance, at each product state x, any schedule in

$$\arg\max_{1\leq i\leq N} \Lambda_i(x_i). \tag{5.5}$$

The first part of the theorem is the bandit representation. At a product state x, the only calendar decision is which local clock to advance. If schedule i is advanced, only the ith coordinate moves. The continuation values relevant to that local schedule are therefore

$$\Big(V(x_{-i}, y_i)\Big)_{y_i \in S_i},$$

which is exactly the vector appearing inside $\rho_i^{x_i}$ in (5.2). The elapsed-calendar constraint therefore turns a collection of local stopping problems into a rested product problem.

The second part identifies the index. The index is not an additional priority ranking imposed from outside the model. It is computed from the one-schedule problem. At shadow price λ, putting schedule i on hold indefinitely has value $\lambda/(1-\beta)$. The set $D_i(\lambda)$ consists of the local states at which that on-hold value is at least as good as advancing the schedule one more local step. The index $\Lambda_i(x_i)$ is the critical shadow price at which putting the schedule on hold first becomes optimal in this isolated one-schedule comparison.

This language should be read as part of the one-schedule shadow-price calculation, not as literal abandonment of a real-world responsibility. A schedule that is not advanced at a given calendar date has not disappeared. In the constrained multi-schedule problem, it remains at its current local state while another schedule's local clock advances. The one-schedule calculation asks a sharper question: how high must the shadow price of calendar time be before advancing this schedule is no longer worthwhile? The answer is the index.

This is the shadow-price interpretation of the calendar-budget analogy. In a consumer problem, a Lagrange multiplier prices the scarce resource in the budget constraint. Here, λ prices one unit of elapsed calendar time. The one-schedule problem asks whether schedule i, at local state x_i, is still worth advancing when calendar time has shadow price λ. The max-index rule advances the schedule with the highest current critical shadow price.

Cash additivity is what makes this scalar shadow price legitimate. To see this, subtract the on-hold value from the one-schedule problem. Define

$$F_i^{\lambda}(x_i) := V_i^{\lambda}(x_i) - \frac{\lambda}{1-\beta}.$$

Using cash additivity of $\rho_i^{x_i}$, (5.3) is equivalent to

$$F_i^\lambda(x_i) = \max\left\{0,\, u(x_i) - \lambda + \beta\rho_i^{x_i}\left(\left(F_i^\lambda(y_i)\right)_{y_i \in S_i}\right)\right\}. \tag{5.6}$$

In this excess-value form, λ is the shadow price of one unit of elapsed calendar time. The on-hold branch is normalized to zero. Increasing λ lowers the active branch by the same scalar in every local state. Since the on-hold branch is fixed, the set of states in which holding the schedule fixed is optimal expands monotonically with λ. This is the indexability statement in the theorem.

The connection with the usual Weber and Whittle arguments is direct. Weber's charge formulation prices continuation directly: after subtracting the common on-hold value, the scalar appears as the price of advancing the schedule. Whittle's passive-subsidy formulation uses the opposite side of the same comparison: a rested inactive schedule receives the per-period subsidy λ, which is equivalent to the one-time value $\lambda/(1-\beta)$. In either sign convention, the scalar is the Lagrange multiplier, or shadow price, associated with elapsed calendar time. R7 is the primitive condition ensuring that this scalar can be moved through the recursive continuation aggregator without being distorted.

To get an idea about how the proof works, notice that the product-state Bellman equation follows from the elapsed-calendar constraint: a feasible calendar step advances exactly one local clock. Maximizing over the schedule to be advanced gives (5.2). Monotonicity and cash additivity imply that each $\rho_i^{x_i}$ is nonexpansive in the sup norm, so the discounted Bellman operator is a contraction.

For the index part, introduce a Lagrangian relaxation of the calendar constraint. In the relaxed problem, schedules choose active/passive independently rather than being required path by path to have exactly one active schedule each date; the scarcity of calendar time is priced by a scalar multiplier λ. Cash additivity is what lets this scalar price enter each one-schedule problem as in (5.6), so the relaxed problem separates schedule by schedule. The key point is that a relaxed optimum is not generally feasible for the original problem: it may advance several schedules, or none. The index argument shows why the relaxation is tight here. A suitable tie-breaking rule selects a relaxed optimizer that advances exactly one maximal-index schedule and holds all nonmaximal schedules fixed. This selected relaxed optimizer satisfies the original elapsed-calendar constraint. Since the relaxation upper-bounds the original problem, and the max-index implementation attains that relaxed bound while satisfying the original constraint, the max-index policy is optimal.

The result is deliberately one-directional. The primitive schedule axioms, common-tail invariance, and rested product construction deliver a generalized bandit representation and its index policy. The theorem does not claim that every abstract bandit representation has a primitive schedule foundation of this form. In the special cases below, stronger equivalence statements are available because additional axioms pin down the continuation aggregator more sharply.

6 Extensions and Special Cases

Now we show familiar models arise through additional restrictions/axioms on our schedule domain. The first is the subjective expected-utility case, where the abstract continuation aggregator becomes a subjective transition kernel and the index reduces to the classical Gittins–Jones ratio. The second is variational preferences, which nest max-min as a further special case. The third is the rank-dependent, or Choquet, case, where continuation branches are evaluated by their rank rather than by additive probabilities. The final one is Pandora's problem, which is not a different continuation aggregator but a restriction on the local schedule domain: one local step opens a box and resolves all uncertainty.

For each state x, define the induced continuation preference $\succeq_x^C$ on continuation families by

$$(a^y)_{y\in S} \succeq_x^C (b^y)_{y\in S} \quad \Longleftrightarrow \quad \Gamma_x(0,(a^y)_{y\in S}) \succeq_x \Gamma_x(0,(b^y)_{y\in S}). \tag{6.1}$$

Thus the current consequence is fixed at the null lottery, and only the contingent continuation tree varies. By R1, this comparison depends only on the continuation equivalence classes. Under D+R, it is represented by

$$(a^y)_y \mapsto \rho_x\Big((U_y(a^y))_{y\in S}\Big).$$

In this section, the continuation domain is assumed to contain the mixtures and common shifts used in the stated axioms. For notational purposes, continuation families can be identified with their continuation-value vectors

$$U(a) := (U_y(a^y))_{y\in S} \in \mathbb{R}^S.$$

Let $\mathcal{T}_x^+$ denote the finite stopping times $\tau \geq 1$ for the one-schedule problem starting from x. Under the consequence-bearing-state convention, the cash-additive shadow-price equation is

$$F^\lambda(x) = \max\Big\{0,\, u(x) - \lambda + \beta\rho_x(F^\lambda)\Big\}. \tag{6.2}$$

The associated per-period index is

$$I(x) := \inf\{\lambda \in \mathbb{R} : F^\lambda(x) = 0\}. \tag{6.3}$$

The role of each special case is to identify the relevant continuation aggregator ρ_x, and then substitute it into (6.2).

6.1 Subjective expected utility

The subjective expected-utility case is the first special case we consider. Many familiar economic bandit models, including Bayesian experimentation and learning models, fall into this framework [Rothschild, 1974, Bolton and Harris, 1999, Keller et al., 2005, Bergemann and Välimäki, 2008]. The restriction is not imposed directly on a transition law. It is imposed on continuation schedules. Once the current consequence is fixed, the decision maker satisfies ordinary independence over

objective mixtures of continuation trees. Under that restriction, the continuation aggregator is linear, so uncertainty in a schedule is summarized by a subjective transition kernel.

This case is sharper than the general bandit representation above. In the general result, D, R, and R7 are enough to obtain a recursive stopping problem, scalar shadow prices, and an index rule. Under subjective expected utility, an additional continuation-independence axiom pins down the continuation aggregator as a linear expectation. Thus the primitive schedule restrictions become *equivalent* to the standard expected-utility bandit representation.

Assumption 4 (EU: one-step continuation independence)**.** *For every $x \in S$, the continuation domain is closed under objective statewise mixtures.*[2] *For all continuation families a, b, c and all $\alpha \in (0,1)$,*

$$a \succeq_x^C b \quad \Longleftrightarrow \quad \alpha a + (1-\alpha)c \succeq_x^C \alpha b + (1-\alpha)c. \tag{6.4}$$

The axiom is the Anscombe–Aumann independence condition applied to continuation schedules. Once the current consequence is fixed, mixing both continuation trees with the same background continuation tree cannot change their ranking. The representation therefore recovers a subjective law over next local states.

Proposition 3 (Subjective expected utility and the Gittins–Jones index)**.** *Assume D, R and assumption 4. Then the following are equivalent.*

(i) EU holds for every $x \in S$.

(ii) There exists a Markov kernel $p = (p_x)_{x \in S}$, with $p_x \in \Delta(S)$, such that

$$\rho_x(v) = \sum_{y \in S} p_x(y) v_y \qquad \text{for every } x \in S. \tag{6.5}$$

Equivalently,

$$U_x\Big(\Gamma_x(\ell, (a^y)_{y \in S})\Big) = u(\ell) + \beta \sum_{y \in S} p_x(y) U_y(a^y). \tag{6.6}$$

Under either condition, R7 holds. The shadow-price equation is

$$F^\lambda(x) = \max\left\{0,\, u(x) - \lambda + \beta \sum_{y \in S} p_x(y) F^\lambda(y)\right\}, \tag{6.7}$$

and the resulting index is the Gittins–Jones ratio

$$I^{EU}(x) = \sup_{\tau \in \mathfrak{T}_x^+} \frac{\mathbf{E}_x^p\left[\sum_{t=0}^{\tau-1} \beta^t u(X_t)\right]}{\mathbf{E}_x^p\left[\sum_{t=0}^{\tau-1} \beta^t\right]}. \tag{6.8}$$

[2] Mixtures are formed componentwise across next states. If one component schedule has already stopped, it contributes the null lottery 0 at all remaining local dates until the other component also stops; if both have stopped, the mixed component is stopped.

6.1.1 Finite good-news learning

Learning is not a new continuation aggregator. It is a special case of the subjective expected-utility representation in which the local state of a strand records the progress of information. The primitive object is still a schedule whose local states carry consequences. The additional structure is a finite schedule subdomain and preference restrictions that make the represented SEU kernel look like a familiar learning process.

We spell this out for a stationary good-news model. This is only one example. The same logic can accommodate other finite learning structures by changing the finite schedule subdomain and the corresponding consistency restrictions. The point of using integer states below is deliberate: the state space need not be written as a set of posterior probabilities or likelihood ratios. For any fixed finite horizon, the reachable learning states are finite. We can label them by integers, or by any other finite ordered set. The posterior and log-likelihood scale are recovered after the representation.

For a finite horizon T, consider the finite learning subdomain

$$S_T^G = \{\gamma, 0, 1, \dots, T, \rho\} \subseteq S.$$

The integer states $0, 1, \dots, T$ are unresolved learning states. The state ρ is the resolved-good state, and γ is a lower benchmark state. The active unresolved states are $1, \dots, T$. State 0 is the terminal unresolved state for this finite truncation. The primitive no-news successor of an active state n is $n-1$. These are only labels. In particular, the integer distance between n and $n-1$ is not assumed to be a unit of log likelihood.

For an active integer state n, write $[z]_n$ for the constant continuation family at n that delivers the consequence carried by state z at every next-state coordinate:

$$[z]_n^y = d_y((z)), \qquad y \in S_T^G.$$

Here $d_y((z))$ denotes the one-period deterministic schedule from state y that delivers the consequence carried by state z. For an event $A \subseteq S_T^G$, write

$$[\rho \text{ on } A; \gamma \text{ off } A]_n$$

for the continuation family at n that delivers the consequence carried by ρ on coordinates in A, and the consequence carried by γ on coordinates outside A. More generally,

$$[\rho \text{ on } \{\rho\}; n-1 \text{ on } \{n-1\}; \gamma \text{ off } \{\rho, n-1\}]_n$$

denotes the continuation family that delivers ρ on the ρ-coordinate, $n-1$ on the $n-1$-coordinate, and γ on every other coordinate. Mixtures of such continuation families are the objective statewise mixtures used in Assumption 4.

Assumption 5 (GNL: finite good-news learning)**.** *On the finite learning subdomain*

$$S_T^G = \{\gamma, 0, 1, \dots, T, \rho\},$$

the following restrictions hold.

(i) State ordering. *For every active integer state $n = 1, \ldots, T$ and every unresolved state $j = 0, \ldots, T$,*

$$[\rho]_n \succ_n^C [j]_n \succ_n^C [\gamma]_n.$$

Moreover, for every active integer state $n = 1, \ldots, T$,

$$[n]_n \succ_n^C [\rho \text{ on } \{\rho\}; \gamma \text{ off } \{\rho\}]_n \succ_n^C [\gamma]_n.$$

(ii) Good-news/no-news exhaustiveness. *For every active integer state $n = 1, \ldots, T$,*

$$[\rho \text{ on } \{\rho, n-1\}; \gamma \text{ off } \{\rho, n-1\}]_n \sim_n^C [\rho]_n. \tag{6.9}$$

(iii) Stationary relative value of good news. *For every pair of active integer states $m, n = 1, \ldots, T$ and every $\alpha \in (0,1)$,*

$$[\rho \text{ on } \{\rho\}; \gamma \text{ off } \{\rho\}]_n \succeq_n^C \alpha[n]_n + (1-\alpha)[\gamma]_n \tag{6.10}$$

if and only if

$$[\rho \text{ on } \{\rho\}; \gamma \text{ off } \{\rho\}]_m \succeq_m^C \alpha[m]_m + (1-\alpha)[\gamma]_m.$$

(iv) Successor-consequence consistency. *For every active integer state $n = 1, \ldots, T$,*

$$[n]_n \sim_n^C [\rho \text{ on } \{\rho\}; n-1 \text{ on } \{n-1\}; \gamma \text{ off } \{\rho, n-1\}]_n. \tag{6.11}$$

The first clause gives the ordinal role of the special states. The resolved state ρ is better than every unresolved integer state, and every unresolved integer state is better than the lower benchmark γ. The additional strict comparison says that the good-news branch is valuable, but not already as valuable as the current unresolved state. This rules out degenerate cases in which good news is either worthless or already certain.

The second clause is an exhaustiveness test. At active state n, a continuation folder that pays ρ on the event $\{\rho, n-1\}$ and γ otherwise is equivalent to receiving ρ for sure. Under the SEU representation this will mean that, from n, all represented one-step weight lies on the two coordinates ρ and $n-1$. The axiom itself is only a preference comparison between continuation folders.

The third clause is the stationarity restriction. It does not say that a primitive probability is constant. It says that the branch bet on the resolved-good coordinate has the same relative value at every active integer state. The comparison is measured against mixtures of the current state n and the lower benchmark γ, not against an exogenous posterior label. The common relative value will become the good-news intensity only after the SEU representation is applied.

The fourth clause connects the integer labels to learning. At state n, the current unresolved consequence is equivalent to the one-step successor-consequence folder: receive ρ on the resolved-good coordinate, receive $n-1$ on the no-news coordinate, and receive γ elsewhere. This is the schedule version of a posterior martingale restriction, but it is stated only as an indifference between continuation folders.

Lemma 1 (Finite good-news representation)**.** *Assume D, R, EU, and GNL on*

$$S_T^G = \{\gamma, 0, 1, \dots, T, \rho\}.$$

Define the normalized revealed value of an unresolved integer state by

$$\mu_n := \frac{u(n) - u(\gamma)}{u(\rho) - u(\gamma)}, \qquad n = 0, \dots, T.$$

Then $\mu_n \in (0,1)$. For every active integer state $n = 1, \dots, T$, let $q_n \in \Delta(S_T^G)$ be the represented one-step continuation law at n. Then

$$q_n(\{\rho, n-1\}) = 1.$$

Moreover, there exists a constant $\theta \in (0,1)$, independent of n, such that

$$q_n(\rho) = \theta\mu_n, \qquad q_n(n-1) = 1 - \theta\mu_n.$$

The normalized revealed values satisfy

$$\mu_{n-1} = \frac{(1-\theta)\mu_n}{1 - \theta\mu_n}. \tag{6.12}$$

Equivalently,

$$\log \frac{\mu_{n-1}}{1-\mu_{n-1}} = \log \frac{\mu_n}{1-\mu_n} + \log(1-\theta). \tag{6.13}$$

Hence the integer states are identified with an affine log-likelihood scale:

$$r_n = a + n\delta, \qquad \delta := -\log(1-\theta).$$

The represented process is the finite stationary good-news process: good news arrives with probability θ conditional on the good type and never arrives conditional on the bad type.

The lemma shows how the finite integer schedule can be read as good-news learning. Before the representation, the states are just

$$\gamma, 0, 1, \dots, T, \rho.$$

After D, R, and EU, each active integer state has a represented one-step law over next-state coordinates. GNL disciplines those represented laws. Exhaustiveness puts all weight on the resolved-good coordinate ρ and the one-step no-news coordinate $n-1$. Stationary relative valuation makes the represented value of the ρ-branch proportional to the current revealed value μ_n, with the same factor θ at every active integer state. Successor-consequence consistency then forces the no-news successor values to satisfy

$$\frac{\mu_{n-1}}{1-\mu_{n-1}} = (1-\theta)\frac{\mu_n}{1-\mu_n}.$$

Thus the integer labels are not assumed to be log-likelihoods. They become log-likelihood states

in the representation, with one step down the integer chain corresponding to the constant no-news decrement

$$-\log(1-\theta).$$

This construction illustrates the general point. A learning model can be encoded by a finite schedule subdomain whose states are ordinary consequence-bearing states. The posterior, likelihood ratio, and signal probabilities are then features of the represented SEU kernel, not primitives of the domain. We used integers because they make this separation transparent, but the same finite chain could be relabeled by rationals, names, or any other finite set. What matters is the preference structure that makes the represented process isomorphic to the standard good-news experiment.

No additional index theorem is required. Once the represented SEU kernel is obtained, the learning strand is governed by the same Gittins–Jones index in Proposition 3. The learning restriction only explains when the kernel behind that index is the one generated by good-news-only Bayesian updating. Extending to infinite horizons also follows by simply taking limits.

6.2 Variational preferences

Expected utility treats a continuation folder as if there were a single subjective law for the next local state. Many responsibility strands are not naturally like that. A work file may depend on how an editor or referee will react; a school matter may depend on how a teacher interprets a child's performance; a household repair may turn out to be minor, serious, or entangled with another task. The person may rank these continuation folders without being willing to name one transition law as the right description of the strand.

Variational preferences, with max-min expected utility as a special case, capture this kind of ambiguous scheduling. The decision maker evaluates a continuation folder against possible one-step descriptions of the next local state. A law p may be used to summarize the branch that will be reached, but using that law carries a penalty $\alpha_x(p)$. Low-penalty laws are treated as plausible descriptions of the strand; high-penalty laws remain possible but are costly to use in the evaluation. This differs from robust bandit models that place ambiguity on the full multi-project environment, where an adverse model may choose a joint description of all projects and thereby link alternatives that would otherwise be separate [Kim and Lim, 2016, Caro and Das Gupta, 2022]. In the schedule interpretation, a work obligation, a household repair, a medical follow-up, and a family responsibility may all compete for the same calendar, but uncertainty about one strand need not be tied to uncertainty about the others. We therefore allow rich ambiguity and robustness within each local continuation problem, while keeping the cross-strand interaction exactly where the model puts it: in the scarce calendar.

Assumption 6 (VA: one-step continuation uncertainty aversion)**.** *For every $x \in S$, the continuation*

domain is closed under objective statewise mixtures, and for all continuation families a, b,

$$a \sim_x^C b \quad \Longrightarrow \quad \alpha a + (1-\alpha) b \succeq_x^C a \qquad \forall \alpha \in (0,1).$$

The axiom is the continuation-schedule version of uncertainty aversion. If two continuation folders are equally good, then objective randomization between them cannot make the decision maker worse off; mixing smooths the exposure to any single ambiguous continuation folder. Mixtures are formed state by state, using the same stopped-schedule convention as in Assumption 4.

For the dynamic formulas below, a path law is a complete scenario book for the future local evolution of the strand. Given a finite history $h_t = (x_0, \ldots, x_t)$, write $p_t^{\mathbb{P}}(\cdot \mid h_t)$ for the one-step law selected by $\mathbb{P}$ after that history. Let

$$\begin{aligned} \mathcal{P}_x^{\Delta} := \Big\{ & \mathbb{P} : \mathbb{P}(X_0 = x) = 1, \text{ and there are kernels } p_t^{\mathbb{P}}(\cdot \mid h_t) \in \Delta(S) \\ & \qquad \text{such that } \mathbb{P}(X_{t+1} \in \cdot \mid H_t = h_t) = p_t^{\mathbb{P}}(\cdot \mid h_t) \text{ for every } t, h_t \Big\}. \end{aligned} \tag{6.14}$$

where $H_t = (X_0, \ldots, X_t)$. Thus the infimum over $\mathbb{P}$ below is an infimum over rectangular, history-dependent one-step descriptions of how the responsibility strand may branch. Just like subjective utility, assumption 6 disciplines the aggregator so that the representation is once again equivalent.

Proposition 4 (Variational continuation and the penalized index)**.** *Assume D, R and assumption 6. Then the following are equivalent.*

(i) R7 and VA hold for every $x \in S$.

(ii) For each x, there exists a convex lower-semicontinuous penalty $\alpha_x : \Delta(S) \to [0, \infty]$, normalized by $\min_p \alpha_x(p) = 0$, such that

$$\rho_x(v) = \min_{p \in \Delta(S)} \{ p \cdot v + \alpha_x(p) \}. \tag{6.15}$$

Under either condition, the charged stopping equation is

$$F^W(x) = \max \left\{ 0,\, u(x) - W + \beta \min_{p \in \Delta(S)} \left[p \cdot F^W + \alpha_x(p) \right] \right\}. \tag{6.16}$$

The resulting variational index is

$$I^{VAR}(x) = \sup_{\tau \in \mathfrak{T}_x^+} \inf_{\mathbb{P}} \frac{\mathbf{E}_{\mathbb{P}} \left[\sum_{t=0}^{\tau-1} \beta^t u(X_t) + \sum_{t=0}^{\tau-1} \beta^{t+1} \alpha_{X_t}(p_t) \right]}{\mathbf{E}_{\mathbb{P}} \left[\sum_{t=0}^{\tau-1} \beta^t \right]}, \tag{6.17}$$

where p_t is the one-step transition law selected at history t.

Max–min is the stark robust version of the variational story. At each local state x, the decision maker keeps a set $P(x)$ of plausible one-step stories and evaluates the continuation folder by the least favorable one. Laws inside $P(x)$ receive zero penalty; laws outside $P(x)$ are ruled out by an infinite penalty. The rectangularity in the index means that after each future local history the relevant ambiguity set is determined by the current local state of the same strand.

Given a correspondence $P : S \rightrightarrows \Delta(S)$, define the associated rectangular path-law family by

$$\mathcal{P}_x^P := \left\{ \mathbb{P} \in \mathcal{P}_x^\Delta : p_t^{\mathbb{P}}(\cdot \mid H_t) \in P(X_t) \text{ for every } t, \mathbb{P}\text{-a.s.} \right\}. \tag{6.18}$$

Corollary 2 (Rectangular max–min as a special case)**.** *Assume D and R, and let $P(x) \subseteq \Delta(S)$ be nonempty, closed, and convex for every $x \in S$. In Proposition 4, set*

$$\alpha_x^P(p) = \begin{cases} 0, & p \in P(x), \\ +\infty, & p \notin P(x). \end{cases} \tag{6.19}$$

Then

$$\rho_x(v) = \min_{p \in P(x)} p \cdot v. \tag{6.20}$$

The charged stopping equation is

$$F^W(x) = \max \left\{ 0,\, u(x) - W + \beta \min_{p \in P(x)} p \cdot F^W \right\}. \tag{6.21}$$

The resulting robust index is

$$I^{MM}(x) = \sup_{\tau \in \Upsilon_x^+} \inf_{\mathbb{P} \in \mathcal{P}_x^P} \frac{\mathbf{E}_{\mathbb{P}}\left[\sum_{t=0}^{\tau-1} \beta^t u(X_t)\right]}{\mathbf{E}_{\mathbb{P}}\left[\sum_{t=0}^{\tau-1} \beta^t\right]}. \tag{6.22}$$

6.3 Rank-dependent and Choquet preferences

Expected utility treats continuation branches additively. Rank-dependent preferences allow the decision maker to evaluate the continuation tree by the rank of its branches. This is useful in the schedule domain because a strand of responsibility often has salient upside or downside branches. A work strand may become important if the feedback is promising; a child-related strand may become urgent if a report card reveals a problem; a household strand may become costly if a repair turns out to be serious. The decision maker need not weight these branches as if only their probabilities matter.

In this sense, what looks like loss aversion in a static lottery becomes something closer to Fear of Missing Out (FOMO) in a continuation problem. The salient object is not merely a low prize relative to a reference point, but the possibility of missing a valuable continuation branch by putting the schedule on hold. A rank-dependent continuation aggregator can make the best or worst branches loom larger than they would under additive probabilities.

For value vectors $v, w \in \mathbb{R}^S$, call v and w comonotonic if there are no states y, z such that

$$(v_y - v_z)(w_y - w_z) < 0.$$

Continuation families are comonotonic when their induced value vectors are comonotonic.

Assumption 7 (CEU: one-step comonotonic continuation independence)**.** *For every $x \in S$, the continuation domain is closed under objective statewise mixtures, and for all pairwise comonotonic contin-*

uation families a, b, c and all $\alpha \in (0,1)$,

$$a \succeq_x^C b \quad \Longleftrightarrow \quad \alpha a + (1-\alpha)c \succeq_x^C \alpha b + (1-\alpha)c. \tag{6.23}$$

CEU keeps the independence logic only when the continuation trees agree on the ranking of states. This is the primitive schedule restriction behind rank-dependent evaluation of continuation uncertainty. Once again we see that the discipline imposed by the assumption is enough to strengthen the bandit representation to equivalence.

Proposition 5 (Choquet continuation and the rank-dependent index)**.** *Assume D, R and assumption 7. Then the following are equivalent.*

(i) CEU holds for every $x \in S$.

(ii) For each x, there exists a normalized capacity $\nu_x : 2^S \to [0,1]$ such that

$$\nu_x(\varnothing) = 0, \qquad \nu_x(S) = 1, \qquad A \subseteq B \Rightarrow \nu_x(A) \leq \nu_x(B),$$

and

$$\rho_x(v) = \int v \, d\nu_x, \tag{6.24}$$

where the integral is the finite-state Choquet integral.

Under either condition, R7 holds. The shadow-price equation is

$$F^\lambda(x) = \max\left\{0, u(x) - \lambda + \beta \int F^\lambda \, d\nu_x\right\}. \tag{6.25}$$

The resulting rank-dependent index is

$$I^{CH}(x) := \inf\{\lambda : F^\lambda(x) = 0\}. \tag{6.26}$$

If ν_x is additive, this reduces to the subjective expected-utility case.

6.4 Pandora boxes

Pandora boxes are different from the preceding cases. Expected utility, variational preferences, and Choquet preferences restrict how a continuation tree is evaluated. Pandora's problem restricts what the local schedule can look like. Each box has an unopened state. One local step opens the box and resolves all uncertainty. After that, the revealed state is absorbing. This is the sense in which the framework is complementary to the Pandora search literature: reservation values arise here from a domain restriction on local schedules, rather than from taking boxes, inspection costs, and reservation equations as primitive [Weitzman, 1979, Olszewski and Weber, 2015, Doval, 2018, Auster and Che, 2025].

Assumption 8 (PB: absorbing one-step revelation)**.** *For each box i, the local state space is*

$$S_i^P := \{q_i\} \cup Z_i, \tag{6.27}$$

where q_i is the unopened state and Z_i is the finite set of revealed states. The local histories satisfy

$$\Omega^P_{i,q_i} = \{(q_i, z, z, \ldots) : z \in Z_i\}, \qquad \Omega^P_{i,z} = \{(z, z, \ldots)\}.$$

Opening box i has one-period utility cost $c_i \geq 0$. Each revealed state $z \in Z_i$ is absorbing and has one-period utility $u(z)$. Its absorbing value is

$$g_i(z) := \frac{u(z)}{1-\beta}. \tag{6.28}$$

The primitive content is that all uncertainty in a box is resolved by the first local step. Once the box is revealed, the associated schedule has no further branching.

Proposition 6 (Pandora reservation indices)**.** *Assume D, R, and R7 on the restricted Pandora local domain, and impose PB. For each unopened box i, there is a unique reservation value π_i satisfying*

$$\pi_i = -c_i + \beta\rho_i\Big((\pi_i \vee g_i(z))_{z\in Z_i}\Big). \tag{6.29}$$

Equivalently, by cash additivity,

$$(1-\beta)\pi_i + c_i = \beta\rho_i\Big(((g_i(z) - \pi_i)^+)_{z\in Z_i}\Big). \tag{6.30}$$

The per-period indices on the Pandora local domain are

$$I^P_i(q_i) = (1-\beta)\pi_i, \qquad I^P_i(z) = u(z) = (1-\beta)g_i(z). \tag{6.31}$$

Equivalently, in one-time value units, the max-index rule opens a box with maximal reservation value π_i whenever that value exceeds the best currently revealed absorbing value, and otherwise commits to the best revealed absorbing state.

In the subjective expected-utility specialization

$$\rho_i(v) = \sum_{z\in Z_i} p_i(z)v(z),$$

(6.30) *becomes*

$$(1-\beta)\pi_i + c_i = \beta \sum_{z\in Z_i} p_i(z)(g_i(z) - \pi_i)^+. \tag{6.32}$$

Under the classical normalization $g_i(z) = z$, the formal undiscounted boundary case $\beta = 1$ gives Weitzman's reservation equation

$$c_i = \sum_{z\in Z_i} p_i(z)(z - \pi_i)^+.$$

7 Conclusion

Bandit models are usually introduced as collections of arms. Each arm has a state, an evolution rule, and a reward process; the decision maker then chooses which arm to operate. This paper has

taken a different starting point. The primitive object is not an arm, but a stopped local contingent schedule: a possible unfolding of one responsibility, project, experiment, or opportunity in its own local time. The economic problem is then to place many such local schedules into one scarce calendar.

The change in domain is the main point of the paper. In consumer theory, one does not begin with demand; one begins with preferences over consumption bundles and derives demand by adding a budget constraint. Here, one does not begin with a bandit; one begins with preferences over stopped schedules and derives a bandit by adding an elapsed-calendar constraint. The analogy is substantive. Bundles are the objects that can be purchased or forgone. Local contingent schedules are the objects that can be advanced or postponed without disappearing.

The first representation theorem shows when preferences over one stopped schedule can be written as a generalized stopping problem. Advancing a strand gives a current consequence, future consequences are discounted in local time, and uncertainty about continuation branches is evaluated by an aggregator. Recursive representations are familiar in decision theory, but the stopped-schedule domain gives them a different role. On this domain, the usual behavioral restrictions do not merely justify a convenient dynamic formula; they characterize when preferences over possible unfoldings of a strand have the structure of a stopping problem.

The move from stopping to bandits requires one further behavioral restriction. Calendar time must be priced in units that remain meaningful across strands. The common-tail compensation axiom supplies this property. If two ways of beginning a strand are exactly balanced, then attaching the same downstream continuation behind both of them should not break the tie. This is what allows a scalar shadow price of calendar time to pass through the local continuation evaluation. Once this is true, the elapsed-calendar constraint turns the collection of stopped schedules into a rested bandit.

The index is therefore not an additional ranking rule imposed by the analyst. It is the critical shadow price at which a strand is just worth putting on hold. A high-index strand is one whose next local step remains worthwhile even when calendar time is expensive. The max-index rule is optimal because the relaxed calendar problem separates into one-strand stopping problems, and the policy selected by those separated problems already satisfies the original calendar constraint.

The special cases show how familiar models fit inside this foundation. Under expected utility, the continuation aggregator becomes a subjective transition law and the classical Gittins–Jones index is recovered. In learning problems, beliefs are local states: they move when a strand is advanced and remain fixed when it is rested. With variational, multiplier, or max–min preferences, the index becomes a robust priority price. With rank-dependent or Choquet preferences, salient continuation branches can loom large, producing a dynamic version of fear of missing out. Pandora's problem comes from a different source: a restriction on the schedule domain in which opening a box resolves all uncertainty in one local step.

The framework also clarifies what is not primitive. A transition law, ambiguity set, capacity, or reservation value is not the starting point. Each appears only after adding restrictions to

preferences, consequences, or the schedule domain. Ambiguity attitudes, for example, can enter through the continuation aggregator, but they can also be introduced by enriching consequences themselves, as with horse lotteries. The point is not to privilege one formal treatment of uncertainty. It is to identify the schedule structure that makes stopping and prioritization meaningful before any particular representation of uncertainty is chosen.

Several extensions are natural. One is to allow unattended strands to evolve, depreciate, or expire, producing a behavioral foundation for restless bandits. Another is to replace the single calendar with several scarce inputs: attention, expertise, money, or physical capacity. A third is to study richer schedule domains in which strands arrive, merge, split, or create obligations for one another. These extensions would move beyond the rested, single-calendar environment studied here while preserving the same organizing principle: represent the evaluation of one local strand first, then impose the scarcity constraint that makes prioritization necessary.

The paper therefore gives a foundation for bandits as economic choice problems. It identifies the primitive alternatives, the behavioral content of stopping, and the scarcity constraint that turns stopping into prioritization. On this view, an arm is not a primitive machine. It is a represented responsibility strand. A bandit is not just a Markov control problem. It is a demand problem for local advancement under a scarce calendar. The index is the resulting shadow price of moving one strand forward.

A Proofs for the Recursive Stopping Representation

Proof of Proposition 1. The argument is the standard finite-stream representation argument, adapted to the stopped-flow domain.

Fix an initial state x. D1 gives a weak order on deterministic schedules, and D2 gives mixture independence on each fixed-horizon space $\mathcal{L}^T$. Thus, for each T, there is an affine expected-utility representation of the restriction of $\succeq_x$ to $\mathcal{L}^T$. D3 aligns the representation of a front block across common same-length tails. D5 then identifies streams that differ only by an appended terminal null consequence, so the weights cannot depend on an artificial terminal zero. D4 gives stationarity: delaying both streams by a common null consequence rescales the comparison but does not change its sign. Hence, after normalizing the first-period weight to one, the date-t weight is β^t for some $\beta \geq 0$.

D6 rules out both degeneracy and non-impatience. The first part of D6 gives lotteries ℓ, m with $u(\ell) > u(m)$. The second part gives

$$u(\ell) + \beta u(m) > u(m) + \beta u(\ell),$$

and therefore $\beta < 1$. Since the second-period consequence matters in the deterministic representation, $\beta > 0$. Thus $\beta \in (0,1)$. D8 makes the same one-period affine index and the same discount factor valid across initial states. D7 gives the required continuity. This yields

$$U_x(d_x(\ell_0, \ldots, \ell_{T-1})) = \sum_{t=0}^{T-1} \beta^t u(\ell_t)$$

for every x and every finite deterministic stream. The usual vNM uniqueness argument gives uniqueness of u up to positive affine transformations, with the corresponding normalization of the displayed representation. ■

Proof of Theorem 1. By Proposition 1, deterministic schedules are already represented on the common discounted utility scale. It remains to represent the one-step comparison between a current lottery and a vector of continuation schedules.

Fix x. R1 implies that the one-step comparison depends on each continuation schedule only through its equivalence class under the relevant next-state preference. Hence the induced relation $\succeq_x^1$ on

$$\mathcal{L} \times Q, \qquad Q := \prod_{y \in S} (\mathcal{A}_y / \sim_y),$$

is well defined. R3 gives a continuous weak order on this one-step domain. R4 gives coordinate independence between the current lottery coordinate and the continuation-profile coordinate. R5 gives double cancellation, and R6 supplies the solvability needed to use the additive conjoint representation theorem on the attainable part of $\mathcal{L} \times Q$. Therefore there exist numerical functions

A_x and B_x such that

$$(\ell, q) \succeq_x^1 (m, q') \quad \Longleftrightarrow \quad A_x(\ell) + B_x(q) \geq A_x(m) + B_x(q').$$

The restriction of this representation to deterministic two-date schedules pins down the current coordinate. Since the deterministic representation has already identified the affine one-period utility index, A_x is a positive affine transform of u. We choose the common normalization in which $A_x = u$. The remaining coordinate can then be written as

$$B_x(q) = \beta \rho_x(q),$$

where ρ_x is defined on the vector of continuation values represented by the coordinates of q. R2 implies that ρ_x is increasing in the continuation values. Continuity of ρ_x follows from R3 and the continuity of the one-step representation. This gives

$$U_x\Big(\Gamma_x(\ell, (a^y)_{y \in S})\Big) = u(\ell) + \beta \rho_x\Big(U_1(a^1), \ldots, U_n(a^n)\Big),$$

which is (3.2).

Conversely, suppose a representation of the displayed form exists. D follows from the restriction to deterministic schedules. R1 follows because only the continuation values enter the one-step representation. R2 follows from monotonicity of ρ_x. R3 follows from the numerical representation and continuity. R4 and R5 follow from the additive separation between the current coordinate and the continuation coordinate. R6 is exactly the interval-richness and solvability condition imposed on the attainable domain. Hence Assumptions 1 and 2 are equivalent to the recursive stopping representation. ■

Proof of Corollary 1. Fix an initial state x. For a finite history $h = (x_0, \ldots, x_t) \in H(x)$, write $|h| = t$, $\bar{x}(h) = x_t$, and

$$h|_k := (x_0, \ldots, x_k), \qquad k \leq t.$$

For $y \in S$, write

$$h, y := (x_0, \ldots, x_t, y).$$

A local stopping rule is equivalently described by its continuation set $C \subseteq H(x)$, where $h \in C$ means that the local schedule is advanced at history h. The set C must be prehistory closed:

$$h \in C,\ k \leq |h| \quad \Longrightarrow \quad h|_k \in C.$$

Equivalently, once a history is not in C, no extension of that history is in C. This is exactly absorbing nullity for the stopped schedule. Given such a C, the associated stopped schedule assigns the consequence carried by $\bar{x}(h)$ at histories $h \in C$, and assigns $\perp$ at histories $h \notin C$.

Under the consequence-bearing-state convention, the current consequence at history h is $\bar{x}(h)$.

Since S is finite, set

$$\bar{u} := \max_{z \in S} |u(z)|, \qquad B := \frac{\bar{u}}{1-\beta}.$$

On constant continuation vectors in the completed bounded range, the deterministic normalization gives

$$\rho_z(c\mathbf{1}) = c.$$

Define finite-depth values $V_x^n : H(x) \to \mathbb{R}_+$ by

$$V_x^0(h) = 0 \qquad \text{for every } h \in H(x),$$

and, for $n \geq 0$,

$$V_x^{n+1}(h) = \max\left\{0,\, u(\bar{x}(h)) + \beta \rho_{\bar{x}(h)}\left(\left(V_x^n(h,y)\right)_{y \in S}\right)\right\}. \tag{A.1}$$

We interpret $V_x^n(h)$ as the value of the best continuation set in the subtree rooted at h, subject to forced stopping after at most n further local advances.

This interpretation follows by induction on n. For $n = 0$, forced stopping gives value zero. Suppose the claim holds for n. Starting from history h, an $(n+1)$-step rule either stops immediately, giving zero, or advances the schedule at h. If it advances, the current contribution is $u(\bar{x}(h))$, and after next local state y the remaining problem is an n-step problem rooted at h, y. By the induction hypothesis, the branchwise optimal continuation values are $V_x^n(h,y)$. Branchwise monotonicity of $\rho_{\bar{x}(h)}$ then gives (A.1).

The sequence (V_x^n) is increasing. Since $V_x^0 \leq V_x^1$, monotonicity of each ρ_z and of $r \mapsto \max\{0, r\}$ gives

$$V_x^n(h) \leq V_x^{n+1}(h) \qquad \text{for every } n \text{ and every } h \in H(x).$$

It is also bounded above by B. If $0 \leq V_x^n(h) \leq B$ for every h, then

$$0\mathbf{1} \leq \left(V_x^n(h,y)\right)_{y \in S} \leq B\mathbf{1}.$$

By monotonicity and the constant-vector normalization,

$$0 \leq \rho_{\bar{x}(h)}\left(\left(V_x^n(h,y)\right)_{y \in S}\right) \leq B.$$

Therefore

$$0 \leq V_x^{n+1}(h) \leq \bar{u} + \beta B = B.$$

Thus $0 \leq V_x^n(h) \leq B$ for every n and h.

Define

$$V_x(h) := \lim_{n \to \infty} V_x^n(h).$$

The limit exists by monotone convergence. Since S is finite,

$$\left(V_x^n(h,y)\right)_{y\in S} \to \left(V_x(h,y)\right)_{y\in S}.$$

Continuity of $\rho_{\bar{x}(h)}$ then permits passage to the limit in (A.1), yielding

$$V_x(h) = \max\left\{0,\, u(\bar{x}(h)) + \beta\rho_{\bar{x}(h)}\left(\left(V_x(h,y)\right)_{y\in S}\right)\right\}.$$

Minimality follows from the same iteration. Let $W : H(x) \to \mathbb{R}_+$ be any bounded nonnegative solution of (3.3). Since $V_x^0 = 0 \le W$, monotonicity of the recursion implies $V_x^n \le W$ for every n. Taking limits gives $V_x \le W$. Hence V_x is the minimal bounded nonnegative solution.

It remains to identify the stopping supremum. For a continuation set $C \subseteq H(x)$, let

$$C^{(n)} := C \cap \bigcup_{t=0}^{n-1} H_t(x)$$

be the rule that follows C for the first n local dates and then forces stopping. Each $C^{(n)}$ is prehistory closed and finite-depth. Therefore its value is bounded above by $V_x^n((x))$, and hence by $V_x((x))$. Conversely, for each n, $V_x^n((x))$ is the supremum over all finite-depth prehistory closed continuation sets with depth at most n. Hence

$$\sup_{\tau\le\infty} \mathcal{U}_x(\tau) = \sup_{n<\infty} V_x^n((x)) = V_x((x)),$$

where the completed value of an unbounded local stopping rule is understood as the limit of the values of its finite-depth truncations. ■

B Proof of the Common-Tail Characterization

Proof of Proposition 2. Fix x and suppose R7 holds. Let $\ell, m, \bar{\ell}, \bar{m}$ and $(r_y)_{y\in S}$ satisfy (4.1). Define

$$a^y := \Gamma_y(\bar{m}, r_y), \qquad b^y := \Gamma_y(\bar{\ell}, r_y), \qquad v_y := U_y(a^y), \qquad c := u(\bar{\ell}) - u(\bar{m}).$$

By the deterministic representation,

$$d_x((\ell,\bar{m})) \sim_x d_x((m,\bar{\ell})) \quad\Longrightarrow\quad u(\ell) + \beta u(\bar{m}) = u(m) + \beta u(\bar{\ell}).$$

Thus

$$u(\ell) = u(m) + \beta c.$$

Now apply the recursive representation at each next state y. Since a^y and b^y differ only by replacing the one-period consequence $\bar{m}$ with $\bar{\ell}$ before the same continuation profile r_y,

$$U_y(b^y) = u(\bar{\ell}) + \beta\rho_y\left((U_z(r_y^z))_{z\in S}\right) = U_y(a^y) + c = v_y + c.$$

Therefore the two schedules in (4.2) have recursive values

$$u(\ell) + \beta\rho_x(v) \qquad \text{and} \qquad u(m) + \beta\rho_x(v + c\mathbf{1}).$$

R7 says that these values are equal. Substituting $u(\ell) = u(m) + \beta c$ gives

$$u(m) + \beta c + \beta\rho_x(v) = u(m) + \beta\rho_x(v + c\mathbf{1}),$$

and hence

$$\rho_x(v + c\mathbf{1}) = \rho_x(v) + c.$$

This is (4.3).

Conversely, suppose cash additivity holds for every attainable common scalar shift. Whenever (4.1) holds, the same calculation gives $u(\ell) = u(m) + \beta c$, and the two composite schedules in (4.2) have values

$$u(m) + \beta c + \beta\rho_x(v) \qquad \text{and} \qquad u(m) + \beta\rho_x(v + c\mathbf{1}).$$

Cash additivity makes these values equal. Hence the two composite schedules are indifferent, which is R7. If the attainable continuation domain is closed under all common translations $v \mapsto v + c\mathbf{1}$, the attainable-shift identity is exactly full cash additivity on that domain. ■

C Proof of the Bandit Representation and Index Theorem

Proof of Theorem 2. The proof has three steps. The first constructs the product Bellman equation. The second constructs the charged one-schedule index and proves nestedness. The third verifies the max-index rule using the prevailing-charge form of the Weber–Whittle argument.

First, consider the product-state Bellman equation. At a product state

$$x = (x_1, \dots, x_N) \in \mathbf{S},$$

a feasible calendar step chooses exactly one local clock to advance. If schedule i is advanced, the product state can only move to a state of the form

$$(x_{-i}, y_i) = (x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_N), \qquad y_i \in S_i.$$

All other coordinates are frozen. Therefore the continuation vector seen by schedule i is

$$\Big(V(x_{-i}, y_i)\Big)_{y_i \in S_i}.$$

The one-schedule recursive representation gives current utility $u(x_i)$ and evaluates this vector through $\rho_i^{x_i}$. Maximizing over the unique local clock to be advanced gives exactly

$$V(x) = \max_{1 \le i \le N} \left\{ u(x_i) + \beta\rho_i^{x_i}\left(\Big(V(x_{-i}, y_i)\Big)_{y_i \in S_i}\right)\right\}.$$

The same assumptions give well posedness. Since $\rho_i^{x_i}$ is increasing and cash additive, it is nonexpansive in the sup norm. Indeed, if $\|v - w\|_\infty \le a$, then

$$v \le w + a\mathbf{1} \qquad \text{and} \qquad w \le v + a\mathbf{1}.$$

Monotonicity and cash additivity imply

$$\rho_i^{x_i}(v) \le \rho_i^{x_i}(w) + a \qquad \text{and} \qquad \rho_i^{x_i}(w) \le \rho_i^{x_i}(v) + a.$$

Thus

$$|\rho_i^{x_i}(v) - \rho_i^{x_i}(w)| \le \|v - w\|_\infty.$$

Because $\beta \in (0,1)$, the product Bellman operator is a contraction on bounded functions on $\mathbf{S}$. Hence (5.2) has a unique bounded fixed point, which is the value of the elapsed-calendar constrained multi-schedule problem.

Second, fix a schedule i and a scalar charge λ. Define the excess value

$$F_i^\lambda(x_i) := V_i^\lambda(x_i) - \frac{\lambda}{1-\beta}.$$

Using cash additivity in (5.3),

$$\rho_i^{x_i}\left(\left(V_i^\lambda(y_i)\right)_{y_i \in S_i}\right) = \rho_i^{x_i}\left(\left(F_i^\lambda(y_i)\right)_{y_i \in S_i} + \frac{\lambda}{1-\beta}\mathbf{1}\right) = \rho_i^{x_i}\left(\left(F_i^\lambda(y_i)\right)_{y_i \in S_i}\right) + \frac{\lambda}{1-\beta}.$$

Subtracting $\lambda/(1-\beta)$ from both sides of (5.3) gives

$$F_i^\lambda(x_i) = \max\left\{0,\, u(x_i) - \lambda + \beta\rho_i^{x_i}\left(\left(F_i^\lambda(y_i)\right)_{y_i \in S_i}\right)\right\}. \tag{C.1}$$

Thus the charged on-hold problem is equivalently a charged stopping problem with stopping payoff zero and active payoff reduced by the scalar charge λ.

Let $\mathcal{T}_i^\lambda$ denote the operator on the right-hand side of (C.1). The nonexpansiveness argument above implies that $\mathcal{T}_i^\lambda$ is a contraction. Hence F_i^λ is its unique fixed point. If $\lambda' \ge \lambda$, then

$$(\mathcal{T}_i^{\lambda'} f)(x_i) \le (\mathcal{T}_i^\lambda f)(x_i) \qquad \text{for every } f \text{ and every } x_i.$$

By the fixed-point comparison theorem for monotone contractions,

$$F_i^{\lambda'}(x_i) \le F_i^\lambda(x_i) \qquad \text{for every } x_i.$$

Since each F_i^λ is nonnegative by (C.1), the on-hold set can be written as

$$D_i(\lambda) = \{x_i : F_i^\lambda(x_i) = 0\}.$$

Therefore,

$$\lambda' \ge \lambda \quad \Longrightarrow \quad D_i(\lambda) \subseteq D_i(\lambda'),$$

which is the nestedness assertion in the theorem.

The endpoint and boundary properties follow from the same contraction argument. Because S_i is finite and the fixed point of a contraction depends continuously on parameters, $F_i^\lambda(x_i)$ is continuous in λ for each x_i. For λ sufficiently high, $F \equiv 0$ is the fixed point of (C.1), so $D_i(\lambda) = S_i$. For λ sufficiently low, one has $F_i^\lambda(x_i) > 0$ for every x_i, so $D_i(\lambda) = \varnothing$. Hence

$$\Lambda_i(x_i) = \inf\{\lambda : x_i \in D_i(\lambda)\}$$

is finite and well defined. Moreover, at $\lambda = \Lambda_i(x_i)$, the active and on-hold branches in (C.1) are both optimal. Indeed, continuity gives $F_i^{\Lambda_i(x_i)}(x_i) = 0$; if the active branch were strictly below zero at this charge, it would remain below zero for nearby lower charges, contradicting the definition of $\Lambda_i(x_i)$.

Third, we verify the max-index rule. First record the equivalent passive-subsidy form of the same one-schedule problem. For each i, write

$$A_i^\lambda(x_i) := u(x_i) - \lambda + \beta \rho_i^{x_i}\left(\left(F_i^\lambda(y_i)\right)_{y_i \in S_i}\right).$$

Equation (C.1) says

$$F_i^\lambda(x_i) = \max\{0, A_i^\lambda(x_i)\}.$$

Since $F_i^\lambda \geq 0$, this is equivalent to

$$F_i^\lambda(x_i) = \max\left\{\beta F_i^\lambda(x_i),\, u(x_i) - \lambda + \beta \rho_i^{x_i}\left(\left(F_i^\lambda(y_i)\right)_{y_i \in S_i}\right)\right\}. \tag{C.2}$$

Adding $\lambda/(1-\beta)$ to (C.2) gives

$$V_i^\lambda(x_i) = \max\left\{\lambda + \beta V_i^\lambda(x_i),\, u(x_i) + \beta \rho_i^{x_i}\left(\left(V_i^\lambda(y_i)\right)_{y_i \in S_i}\right)\right\}. \tag{C.3}$$

Thus V_i^λ is also the one-schedule value when passivity pays the scalar subsidy λ and leaves the rested state fixed. The index $\Lambda_i(x_i)$ is therefore the critical subsidy at which schedule i, in state x_i, first becomes willing to remain passive.

For any fixed charge λ, define the separated charged upper bound

$$B^\lambda(x) := \frac{\lambda}{1-\beta} + \sum_{i=1}^N F_i^\lambda(x_i). \tag{C.4}$$

This is the value obtained after pricing one unit of elapsed local time at λ and then allowing the one-schedule charged stopping problems to separate. It is an upper bound on the original elapsed-calendar value. To see this directly, fix any product state $z \in \mathbf{S}$ and any schedule k. Since

$$B^\lambda(z_{-k}, y_k) = \frac{\lambda}{1-\beta} + \sum_{\ell \neq k} F_\ell^\lambda(z_\ell) + F_k^\lambda(y_k),$$

cash additivity and (C.1) imply

$$
\begin{aligned}
&u(z_k) + \beta \rho_k^{z_k}\left(\left(B^\lambda(z_{-k}, y_k)\right)_{y_k \in S_k}\right) \\
&\quad = u(z_k) + \beta \rho_k^{z_k}\left(\left(F_k^\lambda(y_k)\right)_{y_k \in S_k}\right) + \beta\left[\frac{\lambda}{1-\beta} + \sum_{\ell \neq k} F_\ell^\lambda(z_\ell)\right] \\
&\quad \leq F_k^\lambda(z_k) + \lambda + \beta\left[\frac{\lambda}{1-\beta} + \sum_{\ell \neq k} F_\ell^\lambda(z_\ell)\right] \\
&\quad = \frac{\lambda}{1-\beta} + F_k^\lambda(z_k) + \beta \sum_{\ell \neq k} F_\ell^\lambda(z_\ell) \\
&\quad \leq \frac{\lambda}{1-\beta} + \sum_{\ell=1}^{N} F_\ell^\lambda(z_\ell) = B^\lambda(z).
\end{aligned}
$$

Thus B^λ is a Bellman supersolution for the original product problem, and therefore upper-bounds the value of every original feasible policy.

Now fix a product state x. Suppose first that the largest current index is unique. Let j be the unique schedule with

$$\Lambda_j(x_j) > \max_{i \neq j} \Lambda_i(x_i),$$

and choose λ strictly between the largest and second-largest current indices. Then

$$F_i^\lambda(x_i) = 0 \quad \text{for every } i \neq j, \qquad F_j^\lambda(x_j) > 0.$$

For schedule j, the active branch binds in (C.1):

$$F_j^\lambda(x_j) = u(x_j) - \lambda + \beta \rho_j^{x_j}\left(\left(F_j^\lambda(y_j)\right)_{y_j \in S_j}\right).$$

Using cash additivity and the fact that all non-j excess values at the current state are zero,

$$
\begin{aligned}
&u(x_j) + \beta \rho_j^{x_j}\left(\left(B^\lambda(x_{-j}, y_j)\right)_{y_j \in S_j}\right) \\
&\quad = u(x_j) + \beta \rho_j^{x_j}\left(\left(F_j^\lambda(y_j)\right)_{y_j \in S_j} + \frac{\lambda}{1-\beta}\mathbf{1}\right) \\
&\quad = u(x_j) + \beta \rho_j^{x_j}\left(\left(F_j^\lambda(y_j)\right)_{y_j \in S_j}\right) + \frac{\beta\lambda}{1-\beta} \\
&\quad = F_j^\lambda(x_j) + \lambda + \frac{\beta\lambda}{1-\beta} = F_j^\lambda(x_j) + \frac{\lambda}{1-\beta} = B^\lambda(x).
\end{aligned}
$$

Thus advancing the unique largest-index schedule attains the charged upper bound in the current Bellman comparison.

For any $k \neq j$, the on-hold branch binds weakly in (C.1):

$$0 \geq u(x_k) - \lambda + \beta \rho_k^{x_k}\left(\left(F_k^\lambda(y_k)\right)_{y_k \in S_k}\right).$$

Again using cash additivity,

$$\begin{aligned}
&u(x_k) + \beta\rho_k^{x_k}\left(\left(B^\lambda(x_{-k}, y_k)\right)_{y_k \in S_k}\right) \\
&\quad = u(x_k) + \beta\rho_k^{x_k}\left(\left(F_k^\lambda(y_k)\right)_{y_k \in S_k} + \left[\frac{\lambda}{1-\beta} + F_j^\lambda(x_j)\right]\mathbf{1}\right) \\
&\quad \le \lambda + \beta\left[\frac{\lambda}{1-\beta} + F_j^\lambda(x_j)\right] \\
&\quad = \frac{\lambda}{1-\beta} + \beta F_j^\lambda(x_j) \le \frac{\lambda}{1-\beta} + F_j^\lambda(x_j) = B^\lambda(x).
\end{aligned}$$

Thus no lower-index schedule can exceed the branch value obtained by advancing j.

If the largest current index is tied, let

$$\bar{\lambda}(x) := \max_{1 \le i \le N} \Lambda_i(x_i)$$

and choose any $j \in \arg\max_i \Lambda_i(x_i)$. At $\lambda = \bar{\lambda}(x)$, every maximal schedule is at its boundary, so the active and on-hold branches are both optimal for those schedules; every nonmaximal schedule has the on-hold branch weakly optimal. The same calculation therefore gives equality for the selected maximal schedule j and weak inequality for all other schedules. Hence any fixed tie-breaking rule among maximal-index schedules is valid.

This is the prevailing-charge form of the Weber–Whittle argument. The family of one-schedule subsidy problems is solved for all λ. At a product state, the relevant charge is chosen from the current vector of critical charges: if the maximal index is unique, choose any charge between the largest and second-largest indices; if it is tied, use the common boundary charge and the fixed tie-breaking rule. After the selected schedule moves and the product state changes, the prevailing charge is recomputed from the new current indices.

The selector constructed in this way advances exactly one schedule at every product state and leaves all other rested schedules fixed. It is therefore feasible for the original elapsed-calendar constrained problem. The separated charged problem is a relaxation, hence an upper bound, of the original problem; the prevailing-charge calculation above selects a relaxed optimizer that satisfies the original elapsed-calendar constraint. The relaxation is therefore tight, and the resulting original-feasible policy is optimal. This policy is precisely the max-index rule in (5.5). ■

D Proofs for Extensions and Special Cases

Throughout this appendix, D+R gives the recursive stopping representation from Theorem 1. Thus continuation families can be identified, on the rich continuation domain, with their continuation-value vectors

$$v = (U_y(a^y))_{y \in S} \in \mathbb{R}^S.$$

The deterministic normalization fixes constants:

$$\rho_x(t\mathbf{1}) = t. \tag{D.1}$$

Whenever R7 is imposed, Proposition 2 gives cash additivity:

$$\rho_x(v + c\mathbf{1}) = \rho_x(v) + c. \tag{D.2}$$

Consequently, if a one-time retirement value is $W/(1-\beta)$, the charged one-arm stopping problem can be written in excess-value form as

$$F^W(x) = \max\left\{0,\, u(x) - W + \beta\rho_x(F^W)\right\}. \tag{D.3}$$

The associated index is

$$I(x) = \inf\{W : F^W(x) = 0\}.$$

Proof of Proposition 3. Fix x. Under D+R, the induced continuation preference $\succeq_x^C$ is represented by ρ_x on continuation-value vectors. EU is the finite-state Anscombe–Aumann independence axiom applied to these continuation acts. Together with weak order, continuity, and monotonicity from the recursive one-step axioms, it implies that $\succeq_x^C$ has an affine expected-utility representation:

$$v \mapsto a_x \sum_{y\in S} \pi_x(y)v_y + b_x,$$

where $\pi_x(y) \geq 0$ and the weights are not all zero. Normalizing the weights gives a probability $p_x \in \Delta(S)$.

Since ρ_x represents the same continuation preference, there is a strictly increasing transform φ_x such that

$$\rho_x(v) = \varphi_x\left(\sum_{y\in S} p_x(y)v_y\right).$$

The constant normalization (D.1) gives

$$t = \rho_x(t\mathbf{1}) = \varphi_x(t),$$

so φ_x is the identity on the relevant range. Hence

$$\rho_x(v) = \sum_{y\in S} p_x(y)v_y.$$

This proves the expected-utility representation. Conversely, a linear expectation over a Markov kernel satisfies continuation independence, and it is cash additive. Thus the expected-utility case is the sharp iff specialization of the recursive representation.

Substituting the linear aggregator into (D.3) gives (6.7). For the ratio formula, fix $\tau \in \Upsilon_x^+$.

Continuing until τ and then retiring gives excess value

$$\mathbf{E}_x^p\left[\sum_{t=0}^{\tau-1}\beta^t(u(X_t)-W)\right]=\mathbf{E}_x^p\left[\sum_{t=0}^{\tau-1}\beta^t u(X_t)\right]-W\mathbf{E}_x^p\left[\sum_{t=0}^{\tau-1}\beta^t\right].$$

This continuation plan is weakly worthwhile exactly when W is no larger than the corresponding discounted reward-per-time ratio. Taking the supremum over positive finite stopping times gives (6.8). ■

Proof of Lemma 1. By D, R, and EU, the continuation aggregator at each active integer state $n = 1,\dots,T$ is linear on the finite continuation domain. Hence there is a probability vector

$$q_n \in \Delta(S_T^G)$$

such that, for every continuation-value vector $v \in \mathbb{R}^{S_T^G}$,

$$\rho_n(v)=\sum_{y\in S_T^G} q_n(y)v(y).$$

For $A \subseteq S_T^G$, write

$$q_n(A) := \sum_{y\in A} q_n(y).$$

First, the state-ordering part of GNL identifies the unresolved states as strictly between the two benchmark states. Indeed, for any active integer state n, the constant continuation family $[z]_n$ delivers the same consequence z at every next-state coordinate. Constant preservation of the continuation aggregator therefore gives the value $u(z)$. Thus

$$[\rho]_n \succ_n^C [j]_n \succ_n^C [\gamma]_n$$

implies

$$u(\rho) > u(j) > u(\gamma) \qquad \text{for every } j = 0,\dots,T.$$

In particular $u(\rho) > u(\gamma)$, so the normalized values

$$\mu_j := \frac{u(j)-u(\gamma)}{u(\rho)-u(\gamma)}$$

are well defined and satisfy

$$\mu_j \in (0,1), \qquad j = 0,\dots,T.$$

Next, consider the exhaustiveness condition. The continuation family

$$[\rho \text{ on } \{\rho, n-1\}; \gamma \text{ off } \{\rho, n-1\}]_n$$

has represented continuation value

$$q_n(\{\rho, n-1\})u(\rho) + \Big(1 - q_n(\{\rho, n-1\})\Big)u(\gamma).$$

The constant continuation family $[\rho]_n$ has value $u(\rho)$. The indifference in (6.9) therefore gives

$$q_n(\{\rho, n-1\})u(\rho) + \Big(1 - q_n(\{\rho, n-1\})\Big)u(\gamma) = u(\rho).$$

Since $u(\rho) > u(\gamma)$, this implies

$$q_n(\{\rho, n-1\}) = 1.$$

Hence, under the represented one-step law at n, all weight is placed on the two coordinates ρ and $n-1$.

Now consider the stationary relative-price condition. The good-news branch bet

$$[\rho \text{ on } \{\rho\}; \gamma \text{ off } \{\rho\}]_n$$

has represented value

$$q_n(\rho)u(\rho) + \Big(1 - q_n(\rho)\Big)u(\gamma).$$

The mixed constant continuation family

$$\alpha[n]_n + (1-\alpha)[\gamma]_n$$

has represented value

$$\alpha u(n) + (1-\alpha)u(\gamma).$$

Therefore

$$[\rho \text{ on } \{\rho\}; \gamma \text{ off } \{\rho\}]_n \succeq_n^C \alpha[n]_n + (1-\alpha)[\gamma]_n$$

is equivalent to

$$q_n(\rho)\Big(u(\rho) - u(\gamma)\Big) \geq \alpha\Big(u(n) - u(\gamma)\Big),$$

or, equivalently,

$$q_n(\rho) \geq \alpha\mu_n.$$

The strict comparison

$$[n]_n \succ_n^C [\rho \text{ on } \{\rho\}; \gamma \text{ off } \{\rho\}]_n \succ_n^C [\gamma]_n$$

implies

$$0 < q_n(\rho) < \mu_n.$$

Hence $q_n(\rho)/\mu_n \in (0,1)$.

The equivalence in (6.10), applied for all $\alpha \in (0,1)$, implies that the threshold

$$\frac{q_n(\rho)}{\mu_n}$$

is the same at every active integer state. Otherwise, if $q_n(\rho)/\mu_n \neq q_m(\rho)/\mu_m$, one could choose $\alpha \in (0,1)$ strictly between the two ratios and violate the stated equivalence. Let the common value be

$$\theta \in (0,1).$$

Then, for every $n = 1, \ldots, T$,

$$q_n(\rho) = \theta \mu_n.$$

Using $q_n(\{\rho, n-1\}) = 1$, we also obtain

$$q_n(n-1) = 1 - \theta \mu_n.$$

It remains to derive the no-news update. By successor-consequence consistency,

$$[n]_n \sim_n^C [\rho \text{ on } \{\rho\};\ n-1 \text{ on } \{n-1\};\ \gamma \text{ off } \{\rho, n-1\}]_n.$$

The left-hand side has value $u(n)$. Since the represented law places all weight on $\{\rho, n-1\}$, the right-hand side has value

$$q_n(\rho)u(\rho) + q_n(n-1)u(n-1).$$

Thus

$$u(n) = q_n(\rho)u(\rho) + q_n(n-1)u(n-1).$$

Subtract $u(\gamma)$ from both sides and divide by $u(\rho) - u(\gamma)$. Using the definitions of μ_n and μ_{n-1}, this gives

$$\mu_n = q_n(\rho) + (1 - q_n(\rho))\mu_{n-1}.$$

Substituting $q_n(\rho) = \theta \mu_n$, we get

$$\mu_n = \theta \mu_n + (1 - \theta \mu_n)\mu_{n-1}.$$

Solving for μ_{n-1} yields

$$\mu_{n-1} = \frac{(1-\theta)\mu_n}{1 - \theta \mu_n},$$

which is (6.12).

Equivalently,

$$\frac{\mu_{n-1}}{1-\mu_{n-1}} = (1-\theta)\frac{\mu_n}{1-\mu_n}.$$

Taking logarithms gives

$$\log\frac{\mu_{n-1}}{1-\mu_{n-1}} = \log\frac{\mu_n}{1-\mu_n} + \log(1-\theta),$$

which is (6.13). Hence, if

$$r_n := \log\frac{\mu_n}{1-\mu_n},$$

then

$$r_{n-1} = r_n + \log(1-\theta).$$

Writing

$$\delta := -\log(1-\theta) > 0,$$

we have

$$r_n = a + n\delta$$

for the constant $a = r_0$. Thus the primitive integer labels are represented as an affine log-likelihood scale, and the represented one-step law has the stationary good-news form

$$q_n(\rho) = \theta\mu_n, \qquad q_n(n-1) = 1-\theta\mu_n.$$

This is exactly the finite good-news process in which good news arrives with probability θ conditional on the good type and never arrives conditional on the bad type. ■

Proof of Proposition 4. Fix x. By R7 and Proposition 2, ρ_x is cash additive. VA says that if two continuation-value vectors are indifferent, then their objective mixture is weakly preferred. Thus ρ_x is quasiconcave. Cash additivity turns quasiconcavity into concavity. Let

$$a = \rho_x(v), \qquad b = \rho_x(w).$$

Then

$$\rho_x(v - a\mathbf{1}) = 0 = \rho_x(w - b\mathbf{1}).$$

By VA,

$$\rho_x\Big(\alpha(v-a\mathbf{1}) + (1-\alpha)(w-b\mathbf{1})\Big) \geq 0.$$

Using cash additivity,

$$\rho_x(\alpha v + (1-\alpha)w) \geq \alpha\rho_x(v) + (1-\alpha)\rho_x(w).$$

Thus ρ_x is concave, monotone, continuous, cash additive, and normalized.

Let $\phi_x(v) := -\rho_x(v)$. Then ϕ_x is convex, continuous, decreasing, and satisfies

$$\phi_x(v + c\mathbf{1}) = \phi_x(v) - c.$$

By finite-dimensional convex duality,

$$\phi_x(v) = \sup_{q \in \mathbb{R}^S} \{q \cdot v - \phi_x^*(q)\}.$$

If $\phi_x^*(q) < \infty$, cash additivity forces $q \cdot \mathbf{1} = -1$, while monotonicity forces $q \leq 0$. Hence $q = -p$ for some $p \in \Delta(S)$. Define

$$\alpha_x(p) := \phi_x^*(-p).$$

Then

$$\rho_x(v) = \min_{p \in \Delta(S)} \{p \cdot v + \alpha_x(p)\}.$$

The penalty is convex and lower semicontinuous by construction, and $\rho_x(0) = 0$ implies $\min_p \alpha_x(p) = 0$.

Conversely, any aggregator of the displayed variational form is monotone, cash additive, and concave:

$$\min_p \{p \cdot (v + c\mathbf{1}) + \alpha_x(p)\} = c + \min_p \{p \cdot v + \alpha_x(p)\}.$$

Cash additivity gives R7 through Proposition 2, and concavity gives VA. Thus the variational representation is the iff form of D+R+R7+VA on the rich continuation domain.

Substituting the variational aggregator into (D.3) gives (6.16). For the ratio expression, fix $\tau \in \mathcal{T}_x^+$. Under the rectangular interpretation of one-step penalties, continuing until τ gives excess value

$$\inf_{\mathbb{P}} \mathbf{E}_{\mathbb{P}} \left[\sum_{t=0}^{\tau-1} \beta^t (u(X_t) - W) + \sum_{t=0}^{\tau-1} \beta^{t+1} \alpha_{X_t}(p_t) \right].$$

The W-terms separate as

$$-W \mathbf{E}_{\mathbb{P}} \left[\sum_{t=0}^{\tau-1} \beta^t \right].$$

Thus the plan is weakly worthwhile exactly when W is no larger than the penalized ratio in (6.17). Taking the supremum over stopping times gives the variational index. ■

Proof of Proposition 5. Fix x. Under D+R, continuation schedules are represented on continuation-value vectors by ρ_x. CEU imposes independence only on comonotonic continuation families. By the finite-state Schmeidler representation theorem, continuity, monotonicity, and comonotonic independence imply that $\succeq_x^C$ is represented by a Choquet integral with respect to a normalized

capacity ν_x:

$$\rho_x(v) = \int v \, d\nu_x.$$

The deterministic normalization pins down constants, so the integral is on the same utility scale fixed by D+R.

Conversely, a finite-state Choquet integral with respect to a normalized capacity is continuous, monotone, and satisfies comonotonic independence. Thus the Choquet representation is the iff form of the CEU restriction on the continuation domain.

A normalized Choquet integral is cash additive:

$$\int (v + c\mathbf{1}) \, d\nu_x = \int v \, d\nu_x + c.$$

Substituting the Choquet aggregator into (D.3) gives (6.25), and the index is

$$I^{CH}(x) = \inf\{W : F^W(x) = 0\}.$$

If ν_x is additive, then

$$\int v \, d\nu_x = \sum_{y \in S} \nu_x(\{y\}) v_y,$$

so the model reduces to the expected-utility case. ■

Proof of Proposition 6. Fix an unopened box i. If the outside continuation value is b, opening the box gives

$$H_i(b) := -c_i + \beta \rho_i\Big((b \vee g_i(z))_{z \in Z_i}\Big).$$

A reservation value is a fixed point $b = H_i(b)$. Since ρ_i is monotone and cash additive, it is nonexpansive in the sup norm: if $\|v - w\|_\infty \le a$, then

$$v \le w + a\mathbf{1} \quad \text{and} \quad w \le v + a\mathbf{1},$$

so monotonicity and cash additivity imply

$$|\rho_i(v) - \rho_i(w)| \le a.$$

The map $b \mapsto (b \vee g_i(z))_{z \in Z_i}$ is one-Lipschitz, so H_i is β-Lipschitz. Since $\beta < 1$, H_i has a unique fixed point π_i, which satisfies (6.29).

Cash additivity gives the equivalent excess form. Since

$$\pi_i \vee g_i(z) = \pi_i + (g_i(z) - \pi_i)^+,$$

we have

$$\rho_i\Big((\pi_i \vee g_i(z))_{z \in Z_i}\Big) = \pi_i + \rho_i\Big(((g_i(z) - \pi_i)^+)_{z \in Z_i}\Big).$$

Substituting this into (6.29) gives

$$\pi_i = -c_i + \beta\pi_i + \beta\rho_i\Big(((g_i(z) - \pi_i)^+)_{z\in Z_i}\Big),$$

which is equivalent to (6.30).

For a revealed state z, the local state is absorbing, so using the revealed box forever gives value

$$g_i(z) = \frac{u(z)}{1-\beta}.$$

The corresponding per-period index is $u(z) = (1-\beta)g_i(z)$. At the unopened state, the one-time reservation value is π_i, so the per-period index is $(1-\beta)\pi_i$. This proves (6.31).

It remains to verify the Pandora rule. Let G be the best currently revealed absorbing value. Opening box i is worthwhile against outside value G exactly when

$$-c_i + \beta\rho_i\Big((G \vee g_i(z))_{z\in Z_i}\Big) \geq G.$$

The left-hand side minus G is weakly decreasing in G and is equal to zero at $G = \pi_i$. Hence box i is worth opening exactly when $G \leq \pi_i$, with strict preference when $G < \pi_i$. Therefore, if the best revealed value is below the largest reservation value, the optimal action is to open a box with maximal π_i. If the best revealed value is at least as large as every reservation value, the optimal action is to commit to the best revealed absorbing state. Backward induction on the number of unopened boxes gives the stated rule.

Under subjective expected utility,

$$\rho_i(v) = \sum_{z\in Z_i} p_i(z)v(z),$$

so (6.30) becomes (6.32). With $g_i(z) = z$ and the formal boundary case $\beta = 1$, this reduces to Weitzman's reservation equation. ■

References

Sydney N. Afriat. The construction of utility functions from expenditure data. *International Economic Review*, 8(1):67–77, 1967. doi: 10.2307/2525382.

Christopher M. Anderson. Ambiguity aversion in multi-armed bandit problems. *Theory and Decision*, 72(1):15–33, 2012. doi: 10.1007/s11238-011-9259-2.

Francis J. Anscombe and Robert J. Aumann. A definition of subjective probability. *The Annals of Mathematical Statistics*, 34(1):199–205, 1963. doi: 10.1214/aoms/1177704255.

Sarah Auster and Yeon-Koo Che. Pandora's box reopened: Robust search and choice overload. Discussion Paper 20984, CEPR, 2025. Also available as arXiv:2512.21192.

Dirk Bergemann and Juuso Välimäki. Bandit problems. In Steven N. Durlauf and Lawrence E.

Blume, editors, *The New Palgrave Dictionary of Economics*. Palgrave Macmillan, London, 2008. doi: 10.1057/978-1-349-95121-5_2075-1.

Donald A. Berry and Bert Fristedt. *Bandit Problems: Sequential Allocation of Experiments*. Monographs on Statistics and Applied Probability. Chapman and Hall, London, 1985.

Patrick Bolton and Christopher Harris. Strategic experimentation. *Econometrica*, 67(2):349–374, 1999. doi: 10.1111/1468-0262.00022.

Felipe Caro and Aparupa Das Gupta. Robust control of the multi-armed bandit problem. *Annals of Operations Research*, 317(2):461–480, 2022. doi: 10.1007/s10479-015-1965-7.

Madhav Chandrasekher, Mira Frick, Ryota Iijima, and Yves Le Yaouanq. Dual-self representations of ambiguity preferences. *Econometrica*, 90(3):1029–1061, 2022. doi: 10.3982/ECTA17502.

Xiaoyu Cheng. Relative maximum likelihood updating of ambiguous beliefs. *Journal of Mathematical Economics*, 99:102587, 2022. doi: 10.1016/j.jmateco.2021.102587.

Soo Hong Chew and Larry G. Epstein. Recursive utility under uncertainty. In M. Ali Khan and Nicholas C. Yannelis, editors, *Equilibrium Theory in Infinite Dimensional Spaces*, pages 352–369. Springer, Berlin, 1991.

Yuan Shih Chow, Herbert Robbins, and David Siegmund. *Great Expectations: The Theory of Optimal Stopping*. Houghton Mifflin, Boston, 1971.

Samuel N. Cohen and Tanut Treetanthiploet. Gittins' theorem under uncertainty. *Electronic Journal of Probability*, 27:1–48, 2022. doi: 10.1214/22-EJP742.

Eric V. Denardo, Haechurl Park, and Uriel G. Rothblum. Risk-sensitive and risk-neutral multiarmed bandits. *Mathematics of Operations Research*, 32(2):374–394, 2007. doi: 10.1287/moor.1060.0240.

Tommaso Denti. Posterior separable cost of information. *American Economic Review*, 112(10):3215–3259, 2022. doi: 10.1257/aer.20211252.

Tommaso Denti and Luciano Pomatto. Model and predictive uncertainty: A foundation for smooth ambiguity preferences. *Econometrica*, 90(2):551–584, 2022. doi: 10.3982/ECTA18009.

David Dillenberger, Juan Sebastián Lleras, Philipp Sadowski, and Norio Takeoka. A theory of subjective learning. *Journal of Economic Theory*, 153:287–312, 2014. doi: 10.1016/j.jet.2014.07.003.

David Dillenberger, R. Vijay Krishna, and Philipp Sadowski. Subjective information choice processes. *Theoretical Economics*, 18(2):529–559, 2023. doi: 10.3982/TE4531.

Avinash K. Dixit and Robert S. Pindyck. *Investment under Uncertainty*. Princeton University Press, 1994.

Laura Doval. Whether or not to open pandora's box. *Journal of Economic Theory*, 175:127–158, 2018. doi: 10.1016/j.jet.2018.01.005.

Larry G. Epstein and Martin Schneider. Recursive multiple-priors. *Journal of Economic Theory*, 113(1):1–31, 2003. doi: 10.1016/S0022-0531(03)00097-8.

Larry G. Epstein and Stanley E. Zin. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica*, 57(4):937–969, 1989.

doi: 10.2307/1913778.

Mira Frick, Ryota Iijima, and Yves Le Yaouanq. Objective rationality foundations for (dynamic) α-MEU. *Journal of Economic Theory*, 200:105394, 2022. doi: 10.1016/j.jet.2021.105394.

Itzhak Gilboa and David Schmeidler. Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18(2):141–153, 1989. doi: 10.1016/0304-4068(89)90018-9.

John C. Gittins. Bandit processes and dynamic allocation indices. *Journal of the Royal Statistical Society. Series B (Methodological)*, 41(2):148–177, 1979. doi: 10.1111/j.2517-6161.1979.tb01068.x.

John C. Gittins and David M. Jones. A dynamic allocation index for the sequential design of experiments. In Joseph Gani, editor, *Progress in Statistics*, pages 241–266. North-Holland, Amsterdam, 1974.

John C. Gittins, Kevin D. Glazebrook, and Richard R. Weber. *Multi-Armed Bandit Allocation Indices*. John Wiley & Sons, Chichester, 2 edition, 2011. ISBN 9780470670026.

Lars Peter Hansen and Thomas J. Sargent. Robust control and model uncertainty. *American Economic Review*, 91(2):60–66, 2001. doi: 10.1257/aer.91.2.60.

Lars Peter Hansen and Thomas J. Sargent. *Robustness*. Princeton University Press, Princeton, 2008.

Michael N. Katehakis and Arthur F. Jr. Veinott. The multi-armed bandit problem: Decomposition and computation. *Mathematics of Operations Research*, 12(2):262–268, 1987. doi: 10.1287/moor.12.2.262.

Godfrey Keller, Sven Rady, and Martin Cripps. Strategic experimentation with exponential bandits. *Econometrica*, 73(1):39–68, 2005. doi: 10.1111/j.1468-0262.2005.00564.x.

Michael Jong Kim and Andrew E. B. Lim. Robust multiarmed bandit problems. *Management Science*, 62(1):264–285, 2016. doi: 10.1287/mnsc.2015.2153.

Peter Klibanoff, Massimo Marinacci, and Sujoy Mukerji. A smooth model of decision making under ambiguity. *Econometrica*, 73(6):1849–1892, 2005. doi: 10.1111/j.1468-0262.2005.00640.x.

Tjalling C. Koopmans. Stationary ordinal utility and impatience. *Econometrica*, 28(2):287–309, 1960. doi: 10.2307/1907722.

Tjalling C. Koopmans, Peter A. Diamond, and Richard E. Williamson. Stationary utility and time perspective. *Econometrica*, 32(1/2):82–100, 1964. doi: 10.2307/1913736.

David M. Kreps and Evan L. Porteus. Temporal resolution of uncertainty and dynamic choice theory. *Econometrica*, 46(1):185–200, 1978. doi: 10.2307/1913656.

Jian Li. The K-armed bandit problem with multiple priors. *Journal of Mathematical Economics*, 80:22–38, 2019. doi: 10.1016/j.jmateco.2018.10.002.

Fabio Maccheroni, Massimo Marinacci, and Aldo Rustichini. Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica*, 74(6):1447–1498, 2006a. doi: 10.1111/j.1468-0262.2006.00716.x.

Fabio Maccheroni, Massimo Marinacci, and Aldo Rustichini. Dynamic variational preferences.

Journal of Economic Theory, 128(1):4–44, 2006b. doi: 10.1016/j.jet.2005.12.011.

Milad Malekipirbazari and Ozlem Cavus. Index policy for multiarmed bandit problem with dynamic risk measures. *European Journal of Operational Research*, 312(2):627–640, 2024. doi: 10.1016/j.ejor.2023.08.004.

Wojciech Olszewski and Richard Weber. A more general pandora rule? *Journal of Economic Theory*, 160:429–437, 2015. doi: 10.1016/j.jet.2015.10.009.

Goran Peskir and Albert N. Shiryaev. *Optimal Stopping and Free-Boundary Problems*. Lectures in Mathematics. ETH Z"urich. Birkh"auser, Basel, 2006. doi: 10.1007/978-3-7643-7390-0.

Michael Rothschild. A two-armed bandit theory of market pricing. *Journal of Economic Theory*, 9 (2):185–202, 1974. doi: 10.1016/0022-0531(74)90066-0.

Philipp Sadowski and Todd Sarver. Adaptive preferences: An evolutionary model of non-expected utility and ambiguity aversion. *Journal of Economic Theory*, 218:105840, 2024. doi: 10.1016/j.jet.2024.105840.

Paul A. Samuelson. A note on the pure theory of consumer's behaviour. *Economica*, 5(17):61–71, 1938. doi: 10.2307/2548836.

Todd Sarver. Dynamic mixture-averse preferences. *Econometrica*, 86(4):1347–1382, 2018. doi: 10.3982/ECTA12687.

David Schmeidler. Subjective probability and expected utility without additivity. *Econometrica*, 57(3):571–587, 1989. doi: 10.2307/1911053.

Albert N. Shiryaev. *Optimal Stopping Rules*, volume 8 of *Applications of Mathematics*. Springer-Verlag, Berlin, 1978.

Hal R. Varian. The nonparametric approach to demand analysis. *Econometrica*, 50(4):945–973, 1982. doi: 10.2307/1912771.

Richard R. Weber. On the Gittins index for multiarmed bandits. *The Annals of Applied Probability*, 2(4):1024–1033, 1992. doi: 10.1214/aoap/1177005588.

Martin L. Weitzman. Optimal search for the best alternative. *Econometrica*, 47(3):641–654, 1979. doi: 10.2307/1910412.

Peter Whittle. Multi-armed bandits and the Gittins index. *Journal of the Royal Statistical Society. Series B (Methodological)*, 42(2):143–149, 1980. doi: 10.1111/j.2517-6161.1980.tb01111.x.