

Bayesian Sequential Search: From Censored Observations to Full Information Revelation*

Ehud Lehrer[†] Daniel Z. Li[‡]

Monday 27th April, 2026

Abstract

This paper investigates the role of information coarsening in restoring the tractability of Bayesian sequential search. While learning in uncertain environments typically induces non-monotonic continuation search values, precluding simple reservation-rule optimality, we demonstrate that one-sided censorship serves as a stabilizing force. By dampening the informational impact of search realizations, censorship ensures that the negative fallback-value effect dominates the potentially destabilizing learning effect. We provide necessary and sufficient conditions for the search problem to remain monotone under lower-censorship and derive a state-free sufficient condition based solely on model primitives. Our results show that coarse information can sustain simple stopping rules even when full revelation fails to do so. The analysis extends to multi-distribution and infinite-horizon settings, characterizing optimal strategies via the minimal fixed point of the Bellman operator. Our results can be applied to a wide range of economic settings, such as job search, consumer search, and venture capital investment.

Keywords: sequential search; Bayesian learning; censored observation; full revelation.

JEL classification: C61; D83.

*We thank Eren Arbatli, Remzi Kaygusuz, and Leslie Reinhorn for helpful comments and discussions.

[†]Durham University Business School. Email: ehud.m.lehrer@durham.ac.uk

[‡]Durham University Business School. Tel: 01913346335. Email: daniel.li@durham.ac.uk.

1 Introduction

Search problems with learning arise in many economic environments in which a decision maker (DM) sequentially explores uncertain payoffs while updating beliefs during search. Classical search theory establishes that when the payoff distribution is known, the optimal stopping rule is myopic and follows simple reservation cutoffs (McCall, 1970; Weitzman, 1979). However, in environments with model uncertainty, Bayesian updating introduces a fundamental tension: new information can improve beliefs sufficiently to justify continued search even as the fallback option improves. As a result, the continuation value of search needs not to be monotone, and reservation rules may no longer characterize the optimal stopping behavior.

This paper studies Bayesian search in environments where learning is based on censored observations. In many economic settings, information is selectively revealed: agents observe precise outcomes only when they exceed a threshold, while unfavorable realizations generate coarse signals. For example, job seekers typically learn little from rejections beyond the fact that they fall below an acceptance threshold; consumers will further investigate product characteristics only after preliminary price screening; and investors will develop accurate estimates only for projects that pass current best prototype. Motivated by these environments, we consider a search problem with an unknown payoff distribution, and learning is subject to one-sided censorship.

The key finding is that censored learning can restore monotonicity in Bayesian search problems under moderate conditions. To develop the intuition, we decompose the inter-temporal change in marginal search value into a *fallback value effect* and a *learning effect*. The fallback value effect is always negative as an increase in the fallback option reduces the marginal gain of search, which drives monotonicity in classic search models. In contrast, the *learning effect* can be positive or negative, depending on whether the new information shifts the belief toward a more or less favorable distribution. In general, the learning effect can overturn monotonicity. Censored observations weaken this force: by pooling the informational content of unfavorable outcomes, they dampen positive belief and reduce the marginal search value. As a result, censored learning can restore the dominance of the fallback-value effect and re-establish monotonicity.

Specifically, we consider a model in which the payoff distribution is uncertain and belongs to a finite set of candidate distributions. A DM sequentially explores the uncertain payoffs with free

recall, and updates her beliefs inter-temporally using the Bayesian rule. The learning is subject to lower-censorship, that is, search outcomes lower than the current fallback payoff are censored, while those break-through outcomes are fully revealed.

We first provide a necessary and sufficient condition for monotonicity in Bayesian search problems with censored observations (Theorem 1). The condition is state-dependent, and it specifies when the negative fallback value effect always dominates the possibly positive learning effect under censored observations. We further derive a state-free sufficient condition that guarantees monotonicity. In particular, the sufficient condition depends only on model primitives, i.e., the set of candidate distributions, and is satisfied in common situations. Importantly, the condition is easy to verify, which makes the results applicable in a variety of economic settings.

While the two-distribution model provides a tractable benchmark for characterizing the stabilizing force of censorship, our analysis establishes that these properties are not artifacts of a restricted state space. We provide a generalized framework for Bayesian search involving a finite set of candidate distributions and an infinite search horizon. This extension addresses a fundamental challenge in undiscounted, infinite-horizon Markov Decision Processes (MDPs): the possibility that the Bellman equation admits multiple solutions, some of which, such as a constant function assigning the maximum possible payoff to all states, induce clearly suboptimal stopping behavior.

We resolve this technical ambiguity by proving that the minimal fixed point of the associated Bellman operator uniquely determines the optimal stopping rule. By constructing this solution as the pointwise limit of a sequence of value functions from finite-horizon problems, we establish (Theorem 2) the existence and optimality of a stopping strategy even in the absence of standard contraction-mapping conditions.

Furthermore, this generalized treatment allows us to incorporate an outside option, thereby expanding the model's applicability to settings in which a seeker may permanently forgo searching for a fixed payoff rather than accept the current fallback. By characterizing the optimal stopping rule as the smallest solution to this augmented Bellman equation, we provide the necessary theoretical closure for a wide range of economic applications. This formal foundation demonstrates that informational coarsening serves as a universal mechanism for restoring tractability, regardless of the number of underlying distributions or the presence of external alternatives.

Building on these foundations, we establish a comparison between censored learning and full information revelation. We show that monotonicity under full revelation implies monotonicity under censorship because a censored signal is a garbling of full revelation signals. We also provide a succinct necessary, and sufficient condition for monotonicity under full revelation, which relies on a single-crossing property and an endpoint condition on the location of the crossing point. The comparison formalizes the idea that censored observations can dampen the destabilizing effect of Bayesian learning and restore a tractable solution to the optimal search problems.

The theoretical results can be applied to a wide range of economic settings in which information is selectively revealed. In labor markets, a job seeker often learns little about a rejection beyond the fact that she falls below some acceptance threshold. A streak of rejections will shift her beliefs towards the unfavorable distributions continuously, and eventually induce a stop decision. In a consumer search, a consumer may simply ignore a product with a higher price without further investigating its other characteristics. For this dual form of upper-censorship, we also show that similar monotonicity properties can be maintained under moderate conditions.

The paper contributes to several strands of the literature. First, it relates to work on Bayesian search and learning, where optimal policies are typically non-monotone and history-dependent. We identify a new class of environments where informational coarsening restores monotonicity, making optimal policies more predictable and behaviorally grounded.

Second, we provide a significant technical contribution to the theory of undiscounted dynamic programming with large state spaces and continuation costs. In the context of infinite-state MDPs without discounting, where the Bellman equation may admit multiple, non-optimal solutions, we establish conditions that ensure the appropriate closure, thereby supporting a broad class of complex sequential decision problems.

Finally, the paper connects to the literature on sequential experimentation and selective disclosure. While imperfect information is typically viewed as a source of inefficiency, our analysis highlights a novel constructive role for informational frictions. By ensuring that the negative fallback-value effect consistently dominates the learning effect, censorship acts as a structural stabilizer that makes dynamic search problems more tractable and predictable.

The remainder of the paper is organized as follows. Section 2 reviews the related literature, and Section 3 details the model setup. Section 4 defines the one-step marginal search value

and the conditions for a monotone search problem. Section 5 derives the monotone conditions under lower-censorship and characterizes the resulting optimal stopping rule. Section 6 discusses practical applications, such as job search and consumer search. Section 7 establishes the formal dynamic programming foundation for the Bayesian search framework, providing the minimal fixed point theorem for environments with multiple candidate distributions and infinite horizons. It also provides a comparison to the full-revelation model and characterizes the specific restrictive conditions required for monotonicity when information is not coarsened.

2 Related Literature

In classical sequential search problems, a decision maker (DM) searches for a prize sequentially from a known distribution, each sampling at a cost. The DM's payoff is the expected value of the highest prize discovered net of gross search costs. This search problem is monotone, and the optimal search rule is myopic, i.e., the DM behaves as if she has just one more period to search, and takes the simple form of reservation cutoffs (McCall, 1970; Weitzman, 1979). However, when the prize distribution is unknown, a DM will update her belief of the underlying distribution during search, the above two properties may no longer hold, and solving the search problem can be challenging (Baucells and Zorc, 2025).

Bikhchandani and Sharma (1996) study a search problem with an unknown distribution. They propose two assumptions about the posterior distributions and show that, under Assumption 1, the search problem is monotone and hence the optimal search rule is myopic; furthermore, under Assumption 2, the reservation cutoffs for optimal stopping is history-independent. Assumption 1 states that if the n -th observed prize is also smaller than y , then the probability that the next observed prize is greater than y will decrease. This assumption formalizes the notion that the posterior probability of observing a higher prize, given that all the previous observations are lower, decreases with the number of observations. Assumption 2 states that the posterior probability of observing a prize greater than y , given that all previous observations are smaller than y , depends only on the number of previous observations. Assumption 2 implies that the reservation cutoff just depends on the number of observations and can be pre-determined. In other words, the reservation prize does not depend on the realization of a search history.

They focus mainly on *ad hoc* learning, rather than Bayesian learning models. They show that the two assumptions are satisfied in an *ad hoc* learning model, where the posterior distribution is a convex combination of the prior distribution and the empirical distribution, with the empirical weight a_n being increasing and $\lim_{n \rightarrow \infty} a_n = 1$. They also consider a special Bayesian learning model – the Dirichlet-multinomial distribution, where the posterior distribution is still Dirichlet-multinomial and can be represented as a convex combination of the prior and the empirical distribution, as in the *ad hoc* learning model. [Talmain \(1992\)](#) provides a closed-form solution to the optimal search rule in this case.¹

[Adam \(2001\)](#) extends the above problem to multiple unknown distributions. Specifically, a DM searches for a prize across a set of distinguishable and independent channels. For each channel, there is an unknown distribution, and she can update her belief based on previous search outcomes of this channel. He imposes a strong assumption (A1) that the reservation prize of each channel is decreasing over time, i.e., with a new observation of a channel i , either the new reservation cutoff is decreasing or the realized outcome is higher than the new reservation cutoff.² The author also suggests that the *ad hoc* learning model and Bayesian learning models in the case of the Dirichlet distribution and exponential distribution satisfy Assumption (A1). Under the assumption, he shows that a variant of Pandora’s rule *à la* [Weitzman \(1979\)](#) is optimal, and the reservation cutoffs are history-dependent.

[DeGroot \(1968\)](#) is a seminal paper that establishes a unified approach to Bayesian sequential search problems. He studies a search problem with intertemporal Bayesian updating, where the prize follows a normal distribution with an unknown mean and unit variance. He shows that it is optimal to stop if and only if the discovered prize is either sufficiently high or sufficiently low. [Baucells and Zorc \(2025\)](#) re-investigate the problem with both an unknown mean and variance of a normal distribution. The DM will update her beliefs about the mean and variance intertemporally. They show that neither the myopic property nor the reservation-prize property holds in this case. They develop a concise optimal stopping rule: rather than simply comparing

¹Due to the tractability of the Dirichlet multinomial model, it has been extensively applied in search models with Bayesian learning, such as in empirical studies of search behaviors ([Koulayev, 2013](#); [Santos et al., 2017](#)) and in technology adoption models ([Ulu and Smith, 2009](#); [Smith and Ulu, 2017](#)).

²In search problems with free recall and unknown distributions, there are a fallback value effect and a learning effect. The fallback value effect is always negative. Assumption (A1) rules out the possibility that a strong positive learning effect can dominate the negative fallback value effect, which makes the problem monotone.

the current highest prize to a threshold, the DM needs to compare the standardized value of the current highest prize to a threshold that depends on the standardized search cost.

The common insight is that, for Bayesian sequential search problems, except for the case of Dirichlet multinomial distribution, the problem is in general non-monotone. Hence, the optimal search rule does not have either the myopic or the reservation cutoff property. In this paper, we consider a Bayesian sequential search problem where information is censored in the case of negative outcomes. We show that, under moderate distributional assumptions, this Bayesian search problem under censorship is monotone, which allows a simple solution in the form of reservation cutoffs. Moreover, there are many real-world sittings that fit this setting.

There has been scarce literature studying Bayesian sequential search problems under censorship. [Ferguson and Phadia \(1979\)](#) is a seminal paper on Bayesian learning with censored data. They show that, under certain prior assumptions, the posterior distribution under data censorship preserves conjugacy for certain classes of priors, such as Dirichlet process priors and neutral-to-the-right priors. [He \(2022\)](#) shows that when DMs suffer from the Gambler’s Fallacy (i.e., the erroneous belief that a streak of bad outcomes signals a forthcoming “good streak”) and also learn from predecessors whose data are endogenously censored by stopping, they develop systematically pessimistic beliefs and tend to stop too early. Because predecessors stop once they hit a “good-enough” draw, their observed histories under-represent long good streaks; this asymmetry, combined with the fallacy bias, leads to sub-optimally low stopping thresholds.

Why Bayesian matters for an expected-utility maximizer. For a DM who maximizes expected utility under model uncertainty, Bayesian learning is the unique coherent rule that translates observed data, including censored events, into posterior beliefs consistent with the data-generating process. In sequential search, period- t continuation values hinge on the entire distribution of future draws, which the DM must assess through her posterior. When feedback is censored, naive (non-Bayesian) procedures that ignore the censoring mechanism mis-specify the likelihood and hence distort the posterior: they underweight the information content of “censored events” and overreact to observed breakthroughs. Bayesian updating corrects this by incorporating censored events through $F_i(y)$ or $\bar{F}_i(y)$, yielding posterior odds and Bayes factors that properly account for both revealed values and censored observations.

This coherence is not merely philosophical; it has direct policy implications. First, it ensures dynamic consistency: decisions that are optimal today remain optimal ex post when the same calculations are repeated with the updated posterior. Second, it enables tractable comparative statics: under primitive dominance relations (e.g., single-crossing order $L_1 \prec_{sc} L_2$), one can sign belief shifts and derive conditions under which the learning effect cannot overturn the (negative) fallback effect, thereby restoring monotonicity and validating myopic reservation rules. Third, it is essential for welfare analysis: the search length, the distribution of stopping states, and expected payoffs depend sensitively on how beliefs respond to censored information.

3 The Model

3.1 Model Setup

A Bayesian Decision Maker (DM) faces a number of covered and *ex ante* homogeneous boxes. Each box contains a random value X , which is an i.i.d. draw from an *unknown* distribution L defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. To inspect a box, the DM needs to pay a search cost $c > 0$. The DM inspects the boxes one-by-one sequentially, and updates her belief on L after each observation. The search is with free recall: she can reclaim any previous search outcome at no extra cost. The DM's objective is to maximize her expected search profit, which is equal to the maximum observed value net of gross search costs. For example, given a sequence of random values $\mathbf{X}_t = (X_1, \dots, X_t)$, the search profit

$$\Pi(\mathbf{X}_t) = \max_{\tau \in \{1, \dots, t\}} \{X_\tau\} - tc = Y_t - tc,$$

where $Y_t := \max_{\tau \leq t} \{X_\tau\}$ is defined as the *fallback value*, i.e., the highest discovered value till the end of period t , with $Y_0 \equiv 0$. Denote \mathcal{F}_t the σ -field generated by $\mathbf{X}_t = (X_1, X_2, \dots, X_t)$, which represents the information available till the end of period t , and $\mathcal{F}_t \subseteq \mathcal{F}_{t+1}$.

We now describe a Markov decision problem (MDP) associated with the search problem. At the end of period t , denote the posterior belief by $\alpha_t(\cdot) := \mathbb{P}(L \in \cdot | \mathcal{F}_t)$, and Y_t is the fallback value. The DM can either stop and claim the fallback value Y_t , or continue sampling another box at a cost c in which she obtains a new sample value X and updates her belief. Let $J(Y_t, \alpha_t)$

denote the value function of the MDP, and the Bellman equation can be written as

$$J(Y_t, \alpha_t) = \max \{Y_t, \mathbb{E}[J(\max\{Y_t, X\}, \alpha_{t+1})|\mathcal{F}_t] - c\}, \quad (1)$$

where $\alpha_{t+1} = \mathbb{P}(L \in \cdot | \mathcal{F}_{t+1})$ is the updated posterior belief after a new observation X in period $t + 1$. The function J appears on both sides of this equation, so an argument establishing the existence of such a function is required. We return to this issue later in Section 7.

3.2 A Simple Learning Model

Let X take k possible distinct values, i.e., $X \in K := \{1, 2, \dots, k\}$, and the distribution generating the data can be one of two distributions, i.e., $L \in \{L_1, L_2\}$. Let $L_1 = (p_1, p_2, \dots, p_k)$ and $L_2 = (q_1, q_2, \dots, q_k)$, with $\sum_j p_j = 1$, $\sum_j q_j = 1$. To avoid unnecessary complications, we assume $p_j, q_j > 0$ and $p_k < q_k$.³ For any $x \in K$, denote $F_1(x) := \sum_{j=1}^x p_j$ and $F_2(x) := \sum_{j=1}^x q_j$ as the cumulative mass functions. Given a posterior belief

$$\alpha_t := \mathbb{P}(L = L_1 | \mathcal{F}_t),$$

the posterior probability and cumulative mass functions are: for all $x \in K$,

$$\rho_x(\alpha_t) := \alpha_t p_x + (1 - \alpha_t) q_x, \quad \text{and} \quad F(x, \alpha_t) := \alpha_t F_1(x) + (1 - \alpha_t) F_2(x).$$

Upon a new observation $X = x$, by Bayes' rule, the updated posterior belief is

$$\alpha_{t+1} = \frac{\alpha_t p_x}{\rho_x(\alpha_t)} \quad \text{for all } x \in K. \quad (2)$$

Our main interest lies in Bayesian update under lower censorship, i.e., any sampled value $X = x$ smaller than the fallback value y is censored. In this case, the updated posterior belief is

$$\alpha_{t+1} = \begin{cases} \frac{\alpha_t F_1(y)}{F(y, \alpha_t)}, & \text{if } x \leq y, \\ \frac{\alpha_t p_x}{\rho_x(\alpha_t)}, & \text{if } x > y. \end{cases} \quad (3)$$

³The assumption $p_k < q_k$ simplifies the discussion in the proof of Theorem 4, yet does not change the results.

We introduce some stochastic orders between L_1 and L_2 , which suggest L_2 is 'better'.

Definition 1. (i) L_1 is smaller than L_2 in likelihood ratio order, denoted by $L_1 \prec_{lr} L_2$, if p_x/q_x is decreasing in x . (ii) L_1 is smaller than L_2 in single-crossing order, denoted by $L_1 \prec_{sc} L_2$, if the sequence $(p_x - q_x)_{x=1}^k$ changes its sign from positive to negative only once with increasing x . (iii) L_1 is smaller than L_2 in stochastic order, denoted by $L_1 \prec_{st} L_2$, if $F_1(x) \geq F_2(x)$ for all $x \in K$. (iv) L_1 is smaller than L_2 in increasing convex order, denoted by $L_1 \prec_{icx} L_2$, if $\Delta(x) := \sum_{z=x}^k (q_z - p_z)(z - x) \geq 0$ for all $x \in K$.

Remark 1. Regarding the relation of the four stochastic orders, we have $lr \Rightarrow sc \Rightarrow st \Rightarrow icx$.

Given a posterior belief α_t , denote the posterior distribution by $L(\alpha_t) := \alpha_t L_1 + (1 - \alpha_t) L_2$. The following result shows that, if L_1 and L_2 are ordered according to icx , st , sc , or lr order, then $L(\alpha_t)$ and $L(\alpha_{t+1})$ can also be ordered according to the same order. The result is useful for the comparative static analysis of optimal stopping rules.

Lemma 1. *If $L_1 \prec_o L_2$, $o \in \{icx, st, sc, lr\}$, then $L(\alpha_t) \prec_o L(\alpha_{t+1})$ if and only if $\alpha_t \geq \alpha_{t+1}$.*

Remark 2. The posterior belief α_t is the probability of L being the inferior distribution L_1 . Under the belief update rule (3), the belief shift $\alpha_{t+1} - \alpha_t$ is given by (10). Suppose $L_1 \prec_{sc} L_2$. If a search outcome is censored, then $\alpha_{t+1} - \alpha_t > 0$ and the DM gets more pessimistic as $L(\alpha_{t+1}) \prec_{sc} L(\alpha_t)$. If a search outcome is fully revealed, then the sign of $\alpha_{t+1} - \alpha_t$ is determined by $p_x - q_x$, which can be either positive or negative.

4 A Monotone Search Problem

In general, the search problem (1) does not have a simple closed-form solution. However, if the problem is *monotone*, the optimal stopping rule is myopic and is in the simple form of threshold cutoffs. Specifically, define the one-step marginal search value at the end of period t by

$$U_t := \mathbb{E}[Y_{t+1} | \mathcal{F}_t] - Y_t = \mathbb{E}[(X - Y_t)_+ | \mathcal{F}_t],$$

which measures the expected increment in the fallback value by sampling one more box in period $t + 1$ and is \mathcal{F}_t -measurable. The search problem is *monotone* if

$$U_t \geq U_{t+1} \quad \text{a.s.} \quad (4)$$

Under the monotone condition (4), if a myopic (i.e., one-step ahead) rule calls for stopping at the end of period t , i.e., $U_t \leq c$, it will also call for stopping at the end of period $t + 1$, no matter what the search outcome X in period $t + 1$ is. For monotone search problems, the optimal stopping rule is simple – it is optimal to stop if and only if $U_t \leq c$.

Remark 3. In classic search problems where the distribution L is known, the search problem (1) is obviously monotone. To see this, note that $U_t = \mathbb{E}[(X - Y_t)_+]$ is decreasing in the fallback value Y_t , which increases over time as the search is with free recall. Therefore, $U_t \geq U_{t+1}$ always. However, when L becomes unknown, the DM updates her belief inter-temporally, and the updated belief can either increase or decrease the one-step marginal search value. In this case, the search problem (1) is not monotone in general.

Specifically, in our learning model, the one-step marginal search value is

$$U(y, \alpha) := \mathbb{E}[(X - y)_+ | \alpha] = \sum_{z=y}^k \rho_z(\alpha) (z - y) = \alpha U_1(y) + (1 - \alpha) U_2(y),$$

where $U_i(y) := \mathbb{E}_{L_i}[(X - y)_+]$, $i = 1, 2$, is the one-step marginal search value under distribution L_i . We further define the gap between $U_2(y)$ and $U_1(y)$ by

$$\Delta(y) := U_2(y) - U_1(y) = \sum_{z=y}^k (q_z - p_z) (z - y). \quad (5)$$

The following lemmas provide some useful properties of $U(y, \alpha)$ and $\Delta(y)$.

Assumption 1. $L_1 \prec_{sc} L_2$.

Lemma 2. Under Assumption 1, *i)* $F_1(y) - F_2(y) \geq 0$; *ii)* $\Delta(y) \geq 0$ with equality only at $y = k$, and it is decreasing in y .

Lemma 3. *i) $U(y, \alpha)$ is strictly decreasing and strictly convex in y . ii) Under Assumption 1, $U(y, \alpha)$ is decreasing in α and supermodular in (y, α) .*

Given any state (y, α) and a search outcome $X = x$, the new fallback value $y' = \max\{x, y\}$ and the updated belief α' is derived from (2) if true outcomes are observed, or from (3) if the observation is subject to censorship. The search problem is monotone if

$$U(y, \alpha) \geq U(y', \alpha'), \quad \text{for any } (y, \alpha) \text{ and any } X = x. \quad (6)$$

We can decompose the difference into a fallback value effect and a learning effect, as follows

$$U(y', \alpha') - U(y, \alpha) = \underbrace{U(y', \alpha') - U(y, \alpha')}_{\text{i) fallback value effect}} + \underbrace{U(y, \alpha') - U(y, \alpha)}_{\text{ii) learning effect}}. \quad (7)$$

As shown in Lemma 3, the fallback value effect i) is always negative. Specifically,

$$i) = U(y', \alpha') - U(y, \alpha') = - \sum_{z=y}^{y'-1} \bar{F}(z, \alpha) < 0. \quad (8)$$

But the learning effect ii) can be either negative or positive, observing that

$$ii) = U(y, \alpha') - U(y, \alpha) = (\alpha - \alpha') \Delta(y). \quad (9)$$

Under Assumption 1, $\Delta(y) \geq 0$ and the sign of $\alpha - \alpha'$ then determines whether the learning effect ii) is positive or not. If the joint effect of i) and ii) is negative for any (y, α) and any $X = x$, then the search problem is monotone.

5 Bayesian Search under Lower-censorship

This Section investigates the search problem (1) under lower-censorship. A learning rule is subject to *lower-censorship*, if any search outcome lower than the current fallback value is censored. To be specific, given a fallback value y , if a new search outcome $X \leq y$, then its realization is

censored and the DM only learns the censored event that $X \leq y$.⁴ This learning model captures the key idea that information is usually censored or coarse in the case of negative outcomes, e.g., no news is bad news. There are many real-world situations that fit this learning model, and below we suggest some motivational examples.

Job search with screening. A job seeker searches for external jobs sequentially, and each application costs c . She is currently employed in a job with value y , and she can always keep and stay on a job offer received. The worker’s match value X_i for an firm i is an i.i.d draw from an unknown distribution L , and she updates her belief on L inter-temporally based on her application outcomes. The application is screened by a recruiter or a platform. If she receives a rejection, she will know the exact value of X_i while only learns that it is below a bar. If she is short-listed for an interview and even receives the offer, she will develop a more accurate estimate of X_i . This situation fits our setup of sequential search with lower-censorship.

Venture capital investment. A hi-tech startup searches for venture capital investment. A venture capitalist (VC) screens and selects projects with uncertain potential payoffs. Projects below the current benchmark (e.g., current best prototype) are discarded without full evaluation. Upon a rejection, the start-up only learns that the project is inferior to the current benchmark. However, when the project exceeds the current best prototype, the VC will fully investigate the project and produce a measurable evaluation of the project, and the start-up learns the true value of the project. This is also an example of learning with lower-censorship.

Product development. The same setup also applies for product development. A firm runs experiments on successive product versions. If the new version performs worse than the current benchmark, the detailed performance metrics are often not recorded or analyzed in detail (lower-censorship). As a consequence, the firm’s posterior belief about what “works” is more based on observed improvements, yet not the full distribution of trials.

Consumer price search. The setup can be extended to situations of *upper-censorship*, where a

⁴Bikhchandani and Sharma (1996) also considered search with right-censored data, yet their setup is quite different from ours. Specifically, they considered a sequence of observations $(X_\tau, \sigma_\tau)_{\tau=1}^t$ where $\sigma_\tau = 1$ for a complete observation (i.e., $X_\tau = x_\tau$) and $\sigma_\tau = 0$ for a censored observation (i.e., $X_\tau < x_\tau$). Our setup is different from theirs in at least two perspectives: i) we study Bayesian learning, yet they consider *ad hoc* learning rule; ii) in our setup, only the observations lower than current fallback y are censored, as a result, their sequence of observations may not be feasible in our model.

search outcome higher than a benchmark is censored. Relevant examples may include a consumer searching for a lower price, or a producer searching for a smaller production cost. Consider a consumer who is searching for a lower price. She is not certain of the overall market situation, and will update her belief inter-temporally based on the search outcomes. The search is with free recall, and the current best price is y . The consumer pays little attention to or simply ignores a price higher than y , e.g., without further investigating other characteristics of the commodity. If the sampled prices keep on being higher than y , she may get more and more pessimistic. Eventually, she will stop searching and accept the current price y . This is an example of sequential search under upper-censorship.

5.1 Monotonicity under lower-censorship

Under lower-censorship, the updated posterior belief α' is given by (3), and the belief shift

$$\alpha' - \alpha = \begin{cases} \frac{\alpha(1-\alpha)}{F(y;\alpha)} [F_1(y) - F_2(y)], & \text{if } x \leq y, \\ \frac{\alpha(1-\alpha)}{\rho_x(\alpha)} (p_x - q_x), & \text{if } x > y. \end{cases} \quad (10)$$

Under Assumption 1, let $\hat{x} := \max\{x \in K : p_x - q_x \geq 0\}$. Then under lower censorship,

$$\begin{cases} \alpha' > \alpha, & \text{if } x \leq \max\{y, \hat{x}\}, \\ \alpha' < \alpha & \text{if } x > \max\{y, \hat{x}\}. \end{cases} \quad (11)$$

Note that a censored outcome or an observed outcome smaller than \hat{x} will drive the belief towards the inferior distribution L_1 ; otherwise, the belief shifts to L_2 .⁵ The theorem below provides a necessary and sufficient condition for the search problem to be monotone. It also provides a state-free sufficient condition, which depends only on the distributions of L_1 and L_2 .

Theorem 1. *Under Assumption 1, the search problem under lower-censorship is monotone iff*

$$\sum_{z=y}^{x-1} \bar{F}(z, \alpha) \geq (\alpha - \alpha') \Delta(x), \quad \text{for any } (y, \alpha) \text{ and any } x > \max\{y, \hat{x}\}. \quad (12)$$

⁵Under lower censorship, $\alpha' = \alpha$ can happen when the outcome $x \in (y, \hat{x})$ and $p_x = q_x$.

Moreover, a state-free sufficient condition for (12) to hold is that

$$1 - F_1(x - 1) \geq \frac{\sqrt{q_x} - \sqrt{p_x}}{\sqrt{q_x} + \sqrt{p_x}} \Delta(x), \quad \text{for any } x > \hat{x}. \quad (13)$$

The fallback value effect of $U(y', \alpha') - U(y, \alpha')$ is always negative. Under Assumption 1, the learning effect of $U(y, \alpha') - U(y, \alpha)$ is positive for any $x > \max\{y, \hat{x}\}$, and is negative otherwise. Therefore, the joint effect can be positive only in the region of $x > \max\{y, \hat{x}\}$. Condition (12) ensures that the positive learning effect is dominated by the negative fallback value effect in this region. Condition (12) is both sufficient and necessary, yet it is difficult to verify in practice. The state-free sufficient condition (13) depends only on the model primitives, e.g., the distributions of L_1 and L_2 , and is easy to check.

Remark 4. The single-crossing order turns out to be useful in the study of search problems. First, it implies stochastic order, i.e., $F_1(y) \geq F_2(y)$, and therefore, a censored outcome drives the belief toward the inferior distribution L_1 as shown in (10). Second, it implies increasing convex order, that is, $\Delta(x) = U_2(x) - U_1(x) \geq 0$. The single-crossing order helps to characterize some comparative static results of the optimal stopping rule.

Remark 5. In the state-free sufficient condition (13), the LHS term of $1 - F_1(x - 1)$ is a lower bound of the negative fallback value effect, while the RHS is an upper bound of the positive learning effect. Note that, under Assumption 1 and that $x > \hat{x}$,

$$\frac{\sqrt{q_x} - \sqrt{p_x}}{\sqrt{q_x} + \sqrt{p_x}} = \frac{q_x - p_x}{(\sqrt{q_x} + \sqrt{p_x})^2} < \frac{q_x - p_x}{(2\sqrt{p_x})^2} = \frac{q_x - p_x}{4p_x}.$$

A stronger sufficient condition that implies (13) is as follows: under Assumption 1,

$$1 - F_1(x - 1) \geq \frac{q_x - p_x}{4p_x} \Delta(x), \quad \text{for any } x > \hat{x}. \quad (14)$$

The state-free sufficient condition (13), as well as (14), is not as restrictive as it first looks, and is satisfied in many common situations. We next provide an example.

Example 1 (Monotone under lower-censorship). Consider the following example:

X	1	2	3	4
L_1	$p_1 = 0.25$	$p_2 = 0.25$	$p_3 = 0.25$	$p_4 = 0.25$
L_2	$q_1 = 0.2$	$q_2 = 0.1$	$q_3 = 0.3$	$q_4 = 0.4$

Observe that $L_1 \prec_{sc} L_2$ and $\hat{x} = 2$. When $x = 4$, $\Delta(x) = 0$ and the sufficient condition (13) holds obviously. When $x = 3$, $1 - F_1(x - 1) = 0.5$ and

$$\frac{\sqrt{q_x} - \sqrt{p_x}}{\sqrt{q_x} + \sqrt{p_x}} \Delta(x) = \frac{\sqrt{0.3} - \sqrt{0.25}}{\sqrt{0.3} + \sqrt{0.25}} (0.4 - 0.25) (4 - 3) \simeq 0.00068.$$

Condition (13) also holds. Therefore, the search problem under lower-censorship is monotone.

Remark 6. [Bikhchandani and Sharma \(1996\)](#) propose two assumptions on the learning rule, which are sufficient for a sequential search problem to be monotone. In particular, they show that the two assumptions hold for the *ad hoc* learning rule, where the posterior distribution is a weighted average of the sample distribution and the prior distribution, with the sample weight monotonically converging to 1 when the sample size converges to infinity.

Consider a search history $\mathbf{x}_t = (x_1, x_2, \dots, x_{t-1}, x_t)$ with its maximum value $y < k$. By removing period t observation x_t , we have the period $t - 1$ history $\mathbf{x}_{t-1} = (x_1, x_2, \dots, x_{t-1})$. Assumption 1 of BS states that, for any outcome value $y' > y$, the posterior probability

$$\Pr(X = y' | \mathbf{x}_t) \leq \Pr(X = y' | \mathbf{x}_{t-1}).$$

It formalizes the notion that if a DM observes more outcomes that are lower than y , then the posterior probability of observing a value higher than y decreases. Their Assumption 1 suggests that the learning effect (as defined in our model) is always negative.

Furthermore, consider two search histories of the same length, e.g., $\mathbf{x}_t = (x_1, x_2, \dots, x_t)$ and $\mathbf{z}_t = (z_1, z_2, \dots, z_t)$, and let $y < k$ be the maximum observation of the two search histories. Assumption 2 of BS states that, for any $y' > y$,

$$\Pr(X = y' | \mathbf{x}_t) = \Pr(X = y' | \mathbf{z}_t).$$

That is, the posterior probability of observing a value higher than y , given that all the previous observations are lower than y , depends only on the number of previous observations. This assumption is pivotal to their result of a history-independent cutoff for optimal stopping.

Neither Assumption 1 nor Assumption 2 of BS holds in our model. We consider a Bayesian learning model under lower-censorship, which is distinct from the *ad hoc* learning rule. Moreover, given the censorship rule in our model, the sequence of observations matters in determining the posterior beliefs. We next show this using the previous Example 1.

Let α_t denote the posterior belief after observing the search history \mathbf{x}_t , and the prior belief $\alpha = 0.5$. Consider two search histories: $\mathbf{x}_2 = (x_1, x_2) = (1, 3)$ and $\tilde{\mathbf{x}}_2 = (\tilde{x}_1, \tilde{x}_2) = (2, 1)$. Along the search history $\mathbf{x}_2 = (1, 3)$, from (3), the posterior beliefs are

$$\alpha_1 = \frac{\alpha p_1}{\alpha p_1 + (1 - \alpha)q_1} = \frac{5}{9}, \quad \alpha_2 = \frac{\alpha_1 p_3}{\alpha_1 p_3 + (1 - \alpha_1)q_3} = \frac{25}{49}.$$

Assumption 1 of BS does not hold as $\alpha_2 < \alpha_1$, i.e., it is more likely to observe a higher value 4 with more observations of values strictly less than 4.

Along the search history $\tilde{\mathbf{x}}_2 = (2, 1)$, from (3), the posterior beliefs are

$$\tilde{\alpha}_1 = \frac{\alpha p_2}{\alpha p_2 + (1 - \alpha)q_2} = \frac{5}{7}, \quad \tilde{\alpha}_2 = \frac{\tilde{\alpha}_1 F_1(2)}{\tilde{\alpha}_1 F_1(2) + (1 - \tilde{\alpha}_1)F_2(2)} = \frac{25}{31}.$$

Assumption 2 of BS does not hold either as $\alpha_2 \neq \tilde{\alpha}_2$. In short, the sufficient condition of (13) in our model accommodates the case where the learning effect can be positive, which is not possible under Assumption 1 of BS, and that the cutoff values for optimal stopping is history-dependent, which is different from BS where the cutoffs are history-independent under their Assumption 2.

5.2 Optimal Stopping Rule: Characterizations

For a monotone search problem, the optimal stopping rule is myopic and takes the form of threshold cutoffs. We define a unique cutoff value $y^*(\alpha) \in \bar{K} := K \cup \{0\}$ by

$$y^*(\alpha) = \min \{z \in \bar{K} : U(z, \alpha) \leq c\}. \quad (15)$$

The uniqueness of $y^*(\alpha)$ is ensured as $z \mapsto U(z, \alpha)$ is strictly decreasing and $U(k, \alpha) = 0$. Similarly, we can define a cutoff value for each distribution L_i by $y_i^* = \min \{z \in \bar{K} : U_i(z) \leq c\}$. If the search problem (1) is monotone, it is optimal to stop if and only if $y \geq y^*(\alpha)$.

Lemma 4. *Under Assumption 1 and condition (13), the cutoff value $y^*(\alpha)$ for optimal stopping*

$$y_1^* \leq y^*(\alpha) \leq y_2^*, \quad (16)$$

and is decreasing in both α and c .

Here are some remarks on the cutoff value $y^*(\alpha)$ under lower-censorship.

- Given a fallback value y at the end of period t , consider two possible search outcomes in period $t + 1$, $x, x' \leq y$. Both outcomes are censored, and from (3), the posterior belief remains the same across them. Therefore, for any two search outcomes $x, x' \leq y$, the cutoff values for optimal stopping remain the same in period $t + 1$.
- Given a fallback value y at the end of period t and a search outcome $X_{t+1} = x$. The posterior belief α' is derived from (3), and from (10), we have

$$\begin{cases} y^*(\alpha') \leq y^*(\alpha) & \text{if } x \leq \max\{y, \hat{x}\}, \\ y^*(\alpha') > y^*(\alpha) & \text{if } x > \max\{y, \hat{x}\}. \end{cases}$$

Moreover, $y^*(\alpha') = y^*(\alpha)$ can happen only when $y < x \leq \hat{x}$ and $p_x = q_x$.

We can also represent the stopping cutoff in terms of belief α . It is clear from (16) that, regardless of the belief α , the DM will continue searching for any $y < y_1^*$, which is the lowest possible stopping cutoff, and will stop for any $y \geq y_2^*$. Define $\mathcal{Y} := \{y \in K : y \geq y_1^*\}$ as the set of fallback values with which the DM may stop searching.

Lemma 5 (Cutoff belief for optimal stopping). *For any $y \in \mathcal{Y}$, define a cutoff belief*

$$\bar{\alpha}(y) := \inf\{\alpha \in [0, 1] : U(y, \alpha) \leq c\}. \quad (17)$$

Under Assumption 1 and condition (13), it is optimal to stop if and only if $\alpha \geq \bar{\alpha}(y)$.

To study the transitions of posterior beliefs, it is more convenient to use odds ratios. Define the odds ratio at the end of period t by $O_t := O(\alpha_t) = \frac{\alpha_t}{1-\alpha_t}$, and under lower-censorship,

$$\frac{O_{t+1}}{O_t} = \begin{cases} \frac{F_1(y)}{F_2(y)}, & \text{if } x \leq y, \\ \frac{p_x}{q_x}, & \text{if } x > y, \end{cases} \quad (18)$$

Given the belief cutoff $\bar{\alpha}(y)$, we can define the cutoff odds ratio by $\bar{O}(y) := \frac{\bar{\alpha}(y)}{1-\bar{\alpha}(y)} \in (0, \infty)$. Then it is optimal to stop if and only if $O(\alpha) := \frac{\alpha}{1-\alpha} \geq \bar{O}(y)$.

Note that, for any $y < k$, $R(y) := \frac{F_1(y)}{F_2(y)} > 1$ by Assumption 1 and $q_k > p_k$. If there are m consecutive censored search outcomes, the updated odds ratio is $O(\alpha)R(y)^m$. For any $y \in \mathcal{Y}$ and $\alpha \in (0, 1)$, define the stopping bound for ‘consecutive-censored’ periods by

$$m_y(\alpha) := \min \left\{ m \in \mathbb{N}_0 : O(\alpha) R(y)^m \geq \bar{O}(y) \right\}. \quad (19)$$

That is, given a state (y, α) , $m_y(\alpha)$ is the upper-bound of search periods if the sample values keep on being censored. For example, for any $y \geq y_2^*$, $\bar{\alpha}(y) = 0$ and hence, $\bar{O}(y) = 0$ and $m_y(\alpha) = 0$. For $y = y_1^*$, if $\bar{\alpha}(y_1^*) < 1$, then $\bar{O}(y_1^*)$ is finite, and the DM must stop within $m_\alpha(y_1^*)$ periods if the sample values keep on being censored. In the ‘knife-edge’ condition of $\bar{\alpha}(y_1^*) = 1$, $\bar{O}(y_1^*) = \infty$, and we apply the convention that $m_{y_1^*}(\alpha) = \infty$. We assume away the knife-edge condition of $\bar{\alpha}(y_1^*) = 1$ by introducing the following assumption.

Assumption 2. $U_1(y_1^*) < c$, or equivalently, $\bar{\alpha}(y_1^*) < 1$.

Under Assumption 2, for any fallback value $y \in \mathcal{Y}$, if the sample values keep on being censored, then there exists a finite bound for the maximum number of search periods. If the sample values are not censored, the fallback value y will increase. As \mathcal{Y} is finite, the DM will also stop for sure within a finite number of periods. We formalize the result as follows.

Proposition 1. *Suppose Assumptions 1, 2 and condition (13) hold. Starting from any state (y, α) with $y \in \mathcal{Y}$ and $\alpha \in (0, 1)$, the optimal policy must stop within a finite number of $\bar{n}(y, \alpha)$ periods on all the possible realizations of search histories.*

Moreover, $\bar{n}(y, \alpha) := 0$ for all $y \geq y_2^*$ and can be recursively determined as follows:

$$\bar{n}(y_2^* - 1, \alpha) := m_{y_2^* - 1}(\alpha) + 1, \quad (20)$$

$$\bar{n}(y, \alpha) := \max \left\{ m_y(\alpha), \max_{x \in \{y+1, \dots, k\}} \bar{n}(x, \alpha'_x) + 1 \right\}, \quad \text{for } y \leq y_2^* - 2, \quad (21)$$

where $\alpha'_x := \frac{\alpha p_x}{\alpha p_x + (1-\alpha)q_x}$ from (3).

Remark 7. We assume $\alpha \in (0, 1)$ as there is no intertemporal learning if $\alpha = 0$ or 1. Assumption 2 assumes away the knife-edge condition of $\bar{\alpha}(y_1^*) = 1$, in which case the upper bound of “consecutive-censored” periods can be infinite, i.e., $m_{y_1^*}(\alpha) = \infty$. With free recall, the fallback value y_t increases over time, and censored events become more likely when the search continues. Under Assumption 1, censoring is more indicative of L_1 as odds multiplier $R(y) > 1$. As a result, sequences with many censored observations tend to push beliefs toward L_1 , whereas occasional high uncensored breakthroughs, e.g., $x > \max\{y, \hat{x}\}$, tend to push beliefs toward L_2 while simultaneously raising y_t and bringing stopping closer.

6 Applications

6.1 Job Search

The theoretical framework of sequential search under lower-censorship can be readily applied to job search models, where coarse information is provided in the case of negative outcomes. Consider a job seeker who is currently employed in a job with a value y . She believes she deserves a better job, and searches for external job opportunities sequentially. For each job application, she needs to pay a cost c such as time, effort, or opportunity cost.

There is a continuum of *ex-ante* homogeneous firms in a competitive market. Each firm lists a job with a base value y . The match value between the job seeker and the firm i , denoted by $X_i \in K$, is an i.i.d. draw from an unknown distribution $L \in \{L_1, L_2\}$. The firm i will make an offer if $X_i > y$, and the package value of the job contract is calibrated to the job seeker’s qualification, i.e., the realized match value $X_i = x$.

The recruitment process consists of two steps: i) initial HR screening, and ii) on-site interview.

By screening, we assume a firm i learns the following binary signals for free, i.e., whether the match value $X_i > y$ or not. Onsite interview generates a precise estimate of the match value, i.e., $X_i = x$, yet there is a cost $\epsilon > 0$ for the firm to arrange an interview, such as travel, accommodation, and other logistics costs. Obviously, the firm will simply reject an unqualified application, e.g., $X_i \leq y$, in which case the job seeker learns nothing beyond the fact that her qualification is evaluated below the benchmark y .

Within our framework, we may assume $y > y_1^*$ where y_1^* can be interpreted as an official minimum wage under Assumption 1. The job seeker's initial belief α is such that $y^*(\alpha) > y$, and therefore, she will continue to search as her current job value y is smaller than the stopping cutoff $y^*(\alpha)$. If she keeps on receiving rejections, her posterior belief will keep on increasing as shown in (18). When the updated posterior belief $\alpha' > \bar{\alpha}(y)$, the job seeker will eventually stop searching and stay at her current job (Lemma 5). Below is one example.

Example 2 (Job Search). Let $c = 1$, $y = 2$ and $\alpha = 0.05$. Consider the following example of

$$L_1 = (0.4, 0.4, 0.1, 0.1), \text{ and } L_2 = (0.1, 0.1, 0.4, 0.4).$$

One observe that $\hat{x} = 2$ and the sufficient condition (13) is satisfied. Therefore, it is a monotone search problem under lower-censorship.

Given the belief $\alpha = 0.05$, the stopping cutoff $y^*(\alpha) = 3$ from (15), which is great than the current job value $y = 2$. Therefore, the job seeker continues to look for outside jobs. If she keeps on receiving rejections, the updated posterior belief will keep on increasing. The belief cutoff for optimal stopping is $\bar{\alpha}(y) = \frac{2}{9}$ by (17), and the corresponding cutoff odd ratio is $\bar{O}(y) = \frac{2}{7}$. From (19), we have $R(y) = \frac{F_1(y)}{F_2(y)} = 4$, $O(\alpha) = \frac{1}{19}$ and $m_y(\alpha) = 2$. That is, if the job seeker receives two consecutive rejections, she will stop searching and stay at her current job.

6.2 Consumer Price Search

A consumer searches for a lower price among many *ex ante* homogeneous retailers, and the search is subject to *upper-censorship*, i.e., any sampled price that is higher than the current lowest price is censored. The setting remains the same as in Section 3, and we are still interested in the conditions for the dual problem to be monotone.

After a sample of random prices $\mathbf{X}_t = (X_1, \dots, X_t)$, the search payoff is $\Pi(\mathbf{X}_t) = v - (Y_t + tc)$, where $Y_t := \min_{\tau \in \{1, \dots, t\}} \{X_\tau\}$ is defined as the *fallback price* with $Y_0 = \infty$, and v is a given product value.⁶ As before, let $\alpha_t := \mathbb{P}(L = L_1 | \mathcal{F}_t)$ denote the posterior belief, and $J^u(Y_t, \alpha_t)$ denote the value function under 'upper censorship'. The Bellman equation can be written as

$$J(Y_t, \alpha_t) = \min\{Y_t, \mathbb{E}[J(\min\{Y_t, X\}, \alpha_{t+1}) | \mathcal{F}_t] + c\}. \quad (22)$$

With an outcome $X = x$ in period $t + 1$, the updated posterior belief under *upper-censorship* is

$$\alpha_{t+1} = \begin{cases} \frac{\alpha_t p_x}{\alpha_t p_x + (1 - \alpha_t) q_x}, & \text{if } x \leq y - 1, \\ \frac{\alpha_t \bar{F}_1(y-1)}{\alpha_t \bar{F}_1(y-1) + (1 - \alpha_t) \bar{F}_2(y-1)}, & \text{if } x > y - 1, \end{cases} \quad (23)$$

where $\bar{F}_i(y) := 1 - F_i(y)$, $i = 1, 2$, are the survival functions. In the case of consumer price search, the one-step marginal search value is defined by

$$U(y, \alpha) := \mathbb{E}[(X - y)_- | \alpha] = \alpha U_1(y) + (1 - \alpha) U_2(y)$$

where $U_i(y) := \mathbb{E}_{L_i}[(X - y)_-]$, $i = 1, 2$. As before, define $\Delta(y) = U_2(y) - U_1(y)$. Lemma 6 below summarizes some properties of $\Delta(y)$ and $U(y, \alpha)$.

Lemma 6. (1) Under Assumption 1, $\bar{F}_1(y) \leq \bar{F}_2(y)$, and $\Delta(y) \leq 0$ and is decreasing in y . (2) $U(y, \alpha)$ is strictly increasing and strictly convex in y ; moreover, under Assumption 1, $U(y, \alpha)$ is increasing in α , and supermodular in (y, α) .

Proposition 2 provides conditions under which the search problem (22) is monotone.

Proposition 2. Under Assumption 1, the consumer price search problem under upper censorship is monotone if and only if

$$\sum_{z=x}^{y-1} F(z, \alpha) \geq (\alpha - \alpha') \Delta(x), \quad \text{for any } (y, \alpha) \text{ and any } x < \min\{y, \hat{x}\}, \quad (24)$$

⁶In the hope of not causing confusion, we keep the same notation as in Section 5

where α' is derived from (23). Moreover, a state-free sufficient condition for (24) to hold is

$$F_2(x) \geq \frac{\sqrt{q_x} - \sqrt{p_x}}{\sqrt{q_x} + \sqrt{p_x}} \Delta(x), \quad \text{for all } x < \hat{x}. \quad (25)$$

Example 3 (Monotone under upper-censorship). Consider the previous example of

$$L_1 = (0.25, 0.25, 0.25, 0.25) \quad \text{and} \quad L_2 = (0.2, 0.1, 0.3, 0.4).$$

Observe that $L_1 \prec_{sc} L_2$ and $\hat{x} = 3$. For $x = 1$, $\Delta(1) = 0$, and condition (25) holds obviously. For $x = 2$, $F_2(2) = 0.3$ and $\Delta(2) = \sum_{z=1}^2 (q_z - p_z)(2 - z) = (q_1 - p_1)(2 - 1) = -0.05$. It is obvious that $F_2(2)$ is greater than

$$\frac{\sqrt{q_2} - \sqrt{p_2}}{\sqrt{q_2} + \sqrt{p_2}} \Delta(2) = -\frac{\sqrt{0.1} - \sqrt{0.25}}{\sqrt{0.1} + \sqrt{0.25}} \times 0.05 \simeq 0.0113,$$

and (25) also holds. Therefore, the search problem under upper-censorship is monotone.

Remark 8. Given a state (y, α) , define a unique cutoff value $y^*(\alpha) \in \bar{K}$ that solves

$$y^*(\alpha) = \max \{z \in \bar{K} : U(z, \alpha) \leq c\}.$$

For a monotone consumer search problem (22), it is optimal to stop if and only if $y \leq y^*(\alpha)$.

7 Bayesian Search with Full Revelation

7.1 A Bayesian Search with Multiple Distributions

We extend the model to accommodate a broader class of distributions. The goal of this extension is to clarify the relationship between finite-horizon search problems and their infinite-horizon counterpart, including the associated Markov decision process (MDP), its Bellman equation, and the existence of optimal stopping strategies.

As before, the random variable $X \in K := \{1, 2, \dots, k\}$, yet the true data-generating model is $L \in L_1, \dots, L_\ell$. The DM holds an initial belief $\vec{\alpha}_0$ over the set of possible distributions. The DM chooses how long to search in order to maximize expected payoff. At any point, her

decision depends on two state variables: the current fallback value y and the posterior belief $\vec{\alpha}$. Accordingly, the value function J satisfies the Bellman equation

$$J(y, \vec{\alpha}) = \max \{y, \mathbb{E}[J(\max\{y, X_{\vec{\alpha}}\}, \text{post}(\vec{\alpha} | X_{\vec{\alpha}}))] - c\}, \quad (26)$$

where $X_{\vec{\alpha}}$ is distributed according to the mixture $\sum_{i=1}^{\ell} \vec{\alpha}(i)L_i$, and $\text{post}(\vec{\alpha} | X_{\vec{\alpha}})$ denotes the posterior belief induced by prior $\vec{\alpha}$ and realization $X_{\vec{\alpha}}$ (defined analogously to (1)).

This formulation gives rise to several nontrivial technical challenges. First, the induced MDP typically has a countably infinite state space. Starting from $y = 0$ and prior $\vec{\alpha}_0$, the search process may generate countably many distinct state-belief pairs $(y, \vec{\alpha})$. MDPs with countable, non-compact state spaces fall outside the scope of many standard dynamic programming techniques.

Second, unlike more familiar MDP environments, the problem lacks an immediate fixed-point argument ensuring the existence of an optimal strategy or stopping rule. Classical existence results based on compactness in finite-state models (e.g., Brouwer’s fixed-point theorem), as well as contraction-mapping arguments commonly used in discounted problems, do not apply directly in the present setting.

A further difficulty is that equation (26) may admit multiple solutions. While any solution J induces a stopping rule, namely, stop if $J(y, \vec{\alpha}) = y$ and continue otherwise, not every such rule is optimal. To illustrate, consider the function $J(y, \vec{\alpha}) = k$ for all $(y, \vec{\alpha})$. This function trivially satisfies (26), since it assigns the maximal possible payoff in every state. The induced strategy stops only when $y = k$ and otherwise continues searching, independently of the posterior belief or the likelihood of improvement. This stopping rule is clearly suboptimal. The following theorem characterizes the solution to (26) that uniquely determines the optimal strategy.

Theorem 2. *The optimal stopping rule of the Bayesian search problem is determined by the smallest solution of (26).*

The proof of Theorem 2 is deferred to the Appendix. Before presenting it, we briefly outline the central difficulty and the key structural features of the model that allow us to overcome it.

As in the discussion above, the state space is the product $K \times \Delta(\{1, \dots, \ell\})$, where states consist of a fallback value and a posterior belief over the set of possible outcome distributions.

For any continuous function $f : K \times \Delta(\{1, \dots, \ell\}) \rightarrow \mathbb{R}$, define the Bellman operator

$$Tf(y, \vec{\alpha}) = \max \{y, \mathbb{E}[f(\max\{y, X_{\vec{\alpha}}\}, \text{post}(\vec{\alpha} \mid X_{\vec{\alpha}}))] - c\}. \quad (27)$$

The operator T maps functions into functions on the same space. Accordingly, any solution J to the Bellman equation (26) is a fixed point of T , i.e., $J = TJ$.

In standard Markov decision problems, such as discounted models with compact state spaces, the Bellman operator admits a unique fixed point, which coincides with the optimal continuation value function and fully characterizes the optimal strategy (see, e.g., [Lucas and Stokey \(1989\)](#)). In the present model, by contrast, fixed points need not be unique. Existence itself is not problematic: the constant function equal to k discussed above provides a trivial example. The challenge is instead to identify the fixed point that corresponds to optimal behavior.

Our approach is to construct a particular fixed point and to show that it is minimal, in the sense that it is pointwise smaller than any other fixed point of T . We then prove that the stopping rule induced by this minimal fixed point is optimal.

The key property enabling this construction is the monotonicity of the fallback value: along any search path, the fallback value can only increase. This monotonicity allows us to construct an increasing sequence of value functions, each corresponding to a finite-horizon Bayesian search problem, and to show that each function in the sequence induces an optimal stopping time for its respective horizon (Lemma 12). We then prove that the pointwise limit of this sequence is a fixed point of the Bellman operator (Lemma 10) and that it is minimal among all fixed points (Lemma 11). Finally, and most importantly, we show that the stopping rule induced by this limit function is optimal for the unrestricted Bayesian search problem without an exogenously imposed horizon (Lemma 15).

The main drawback of this approach is its technical complexity. This limitation motivates the broader objective of the paper

7.2 A Bayesian Search Problem with an Outside Option

As before, the DM chooses how long to continue searching so as to maximize her expected payoff. In this section, the DM is endowed with an outside option. In addition to the standard choices

of stopping and accepting the current fallback value or continuing the search, she may quit permanently at any time and obtain a fixed payoff O .

At each point in time, her decision is summarized by two state variables: the maximum between the current fallback value y and O and the posterior belief $\vec{\alpha}$. Accordingly, the value function J_O satisfies the following Bellman equation:

$$J_O(\max\{y, O\}, \vec{\alpha}) = \max\{y, O, \mathbb{E}[J_O(\max\{y, O, X_{\vec{\alpha}}\}, \text{post}(\vec{\alpha} | X_{\vec{\alpha}}))] - c\}. \quad (28)$$

The first two terms inside the maximization correspond to stopping immediately and obtaining the larger of the current fallback value y and the outside option O . The third term represents the continuation value: the DM pays the search cost c , observes a realization $X_{\vec{\alpha}}$ drawn according to belief $\vec{\alpha}$, updates her belief to $\text{post}(\vec{\alpha} | X_{\vec{\alpha}})$, and updates the fallback value to the best payoff available so far, namely $\max\{y, O, X_{\vec{\alpha}}\}$.

Finally, note that when the outside option satisfies $O = 0$, the Bellman equation (28) reduces to (26), so the two formulations coincide. The technique we develop for the proof of Theorem 2 enables us to demonstrate the following result.

Theorem 3. *The optimal stopping rule of the Bayesian search problem with outside option O is determined by the smallest solution of (28).*

7.3 Full Revelation Model with Two Distributions

Reconsider the model in Section 3, yet now assume all search outcomes are fully revealed. That is, the DM learns the true values of all sampled boxes. Under full revelation, the updated belief is given by (2). We are still interested in monotone search problems, that is, when condition (6) holds. Intuitively, the monotone conditions under full revelation should be more restrictive than those under censorship. The formal result is given as follows.

Proposition 3. *If the search problem is monotone under full revelation, then it is monotone under lower-censorship.*

Remark 9. In fact, the result of Proposition 3 holds for *any* form of censorship. The key insight is that the posterior of a censored outcome can be represented as a conditional mean of the

posteriors of the relevant revealed outcomes, and that $U(y, \alpha)$ is linear in α . Therefore, if the monotone condition (6) holds for all search outcomes under full revelation, it should also hold under any form of censorship. Note that Proposition 3 does not rely on any assumption on the stochastic orders between L_1 and L_2 .

This insight also implies that the maximum expected search profit under full revelation is always (weakly) greater than that under censorship. Because a censored signal is a garbling of full-revelation signals, a DM with full revelation can always replicate the strategies and outcomes of a DM under censorship by intentionally pooling those censored outcomes, e.g., $X \leq y$ in the case of lower-censorship. This result does not depend on search costs or risk-aversion.

Remark 10. The welfare impact of censorship can be categorized into two primary effects:

- **Informational Loss:** For censored events, learning occurs from truncated likelihood terms, e.g., $F(y, \alpha)$, rather than full realizations. This loss of granularity prevents full learning of search outcomes, and hence the optimal stopping rule under full revelation. Moreover, in a monotone search problem, the existence of an upper bound $\bar{n}(y, \alpha)$ also suggests that the optimal stopping rule prevents full learning of the underlying distribution.
- **Policy Distortions:** Naive procedures that ignore the censoring mechanism would mis-specify the likelihood, underweighting the information content of "no message" and over-reacting to breakthroughs. Our Bayesian framework corrects this, yielding posterior odds that properly account for censored events.

Theorem 4. *The search problem with full revelation is monotone if and only if*

$$L_1 \prec_{sc} L_2 \text{ and } \hat{x} \geq k - 1. \tag{29}$$

Remark 11. Condition (29) for monotonicity is quite restrictive. It requires that a single crossing occurs at the very top, i.e., $\hat{x} \geq k - 1$. In our previous example of

$$L_1 = (0.25, 0.25, 0.25, 0.25) \quad \text{and} \quad L_2 = (0.2, 0.1, 0.3, 0.4).$$

As $\hat{x} = 2 < k - 1$, condition (29) is not satisfied. Therefore, the search problem is not monotone under full revelation, yet it is monotone under either lower- or upper-censorship.

- When $k = 2$, condition (29) holds trivially. Therefore, when $k = 2$, a search problem under full revelation is always monotone.
- Observe that condition (29) implies $p_j \geq q_j$ for all $j \leq k - 1$, and therefore, α_t increases over time. We have, if the search problem with full revelation is monotone, then the stopping cutoff $y^*(\alpha_t)$ is decreasing over time.

Under lower-censorship, we establish a finite stopping bound, $\bar{n}(y, \alpha)$, which ensures that for any $y \geq y_1^*$ and $\alpha \in (0, 1)$, the DM will stop within $\bar{n}(y, \alpha)$ periods on every realization of search outcomes. We can establish a similar stopping bound, $\bar{n}^f(y, \alpha)$, in the case of full revelation. Moreover, it is greater than the stopping bound under lower censorship.

Lemma 7. *Suppose Assumption 2 and (29) hold, and $p_j \neq q_j$ for all $j \in K$. Starting from any state (y, α) with $y \in \mathcal{Y}$ and $\alpha \in (0, 1)$, the optimal policy under full revelation stops within a finite number of $\bar{n}^f(y, \alpha)$ periods on all possible realizations of search outcomes. Moreover,*

$$\bar{n}^f(y, \alpha) \geq \bar{n}(y, \alpha).$$

8 Summary

This paper investigates the role of information coarsening in maintaining the tractability of Bayesian sequential search problems. While traditional Bayesian updating in uncertain environments often leads to non-monotone continuation values, thereby precluding simple reservation-price rules, the paper demonstrates that one-sided censorship serves as a stabilizing force. By limiting the informational impact of unfavorable outcomes, censorship ensures that the negative fallback-value effect consistently outweighs the potentially destabilizing learning effect.

In addition, the paper provides necessary and sufficient conditions (Theorem 1) for a search problem to remain monotone under lower-censorship, alongside a state-free sufficient condition based solely on model primitives. These results highlight a paradoxical benefit of "coarse" information: partial disclosure can sustain simple, optimal stopping rules even when full revelation would cause them to fail. This mechanism is shown to be robust across multi-distribution and

infinite-horizon settings, where optimal strategies are characterized by the minimal fixed point of the Bellman operator.

The analysis extends these theoretical insights to practical economic contexts, including job search with offer filtration, consumer price search, and venture capital evaluation. In these environments, informational frictions such as rejections or preliminary screenings provide only coarse signals, which effectively moderate the influence of learning on continuation values. Ultimately, the paper suggests that the inherent limitations of censored data can, in fact, preserve the monotone structure necessary for reservation-rule optimality in complex learning environments.

Section 7 establishes the formal dynamic programming foundation for the Bayesian search framework, extending the results from the simple two-distribution benchmark to a generalized environment with multiple candidate distributions. The primary contribution of this section is the resolution of several non-trivial technical challenges inherent in infinite-state Markov Decision Processes (MDPs) with no discount. In this setting, the Bellman equation may admit multiple solutions, some of which induce clearly sub-optimal stopping rules. Theorems 2 and 3 provide the necessary theoretical closure by proving that the optimal stopping rule is uniquely determined by the minimal fixed point of the Bellman operator. This is achieved constructively by demonstrating that the pointwise limit of a sequence of finite-horizon value functions converges to the smallest solution of the functional equation.

Furthermore, Section 7 provides the rigorous justification for the "stabilizing force" of censorship through a direct comparison with the full revelation model

References

- ADAM, K. (2001): "Learning while searching for the best alternative," *Journal of Economic Theory*, 101, 252–280.
- BAUCELLS, M. AND S. ZORC (2025): "Search in the dark: The case with recall and gaussian learning," *Operations Research*, 73, 2572–2590.
- BIKHCHANDANI, S. AND S. SHARMA (1996): "Optimal search with learning," *Journal of Economic Dynamics and Control*, 20, 333–359.

- DEGROOT, M. H. (1968): “Some problems of optimal stopping,” *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 30, 108–122.
- FERGUSON, T. S. AND E. G. PHADIA (1979): “Bayesian nonparametric estimation based on censored data,” *The annals of statistics*, 163–186.
- HE, K. (2022): “Mislearning from censored data: The gambler’s fallacy and other correlational mistakes in optimal-stopping problems,” *Theoretical Economics*, 17, 1269–1312.
- KOULAYEV, S. (2013): “Search with dirichlet priors: estimation and implications for consumer demand,” *Journal of Business & Economic Statistics*, 31, 226–239.
- LUCAS, R. E. AND N. L. STOKEY (1989): *Recursive methods in economic dynamics*, Harvard University Press.
- MCCALL, J. J. (1970): “Economics of information and job search,” *The Quarterly Journal of Economics*, 84, 113–126.
- SANTOS, B. D. L., A. HORTAÇSU, AND M. R. WILDENBEEST (2017): “Search with learning for differentiated products: Evidence from e-commerce,” *Journal of Business & Economic Statistics*, 35, 626–641.
- SMITH, J. E. AND C. ULU (2017): “Risk aversion, information acquisition, and technology adoption,” *Operations Research*, 65, 1011–1028.
- TALMAIN, G. (1992): “Search from an unknown distribution an explicit solution,” *Journal of Economic Theory*, 57, 141–157.
- ULU, C. AND J. E. SMITH (2009): “Uncertainty, information acquisition, and technology adoption,” *Operations Research*, 57, 740–752.
- WEITZMAN, M. L. (1979): “Optimal search for the best alternative,” *Econometrica: Journal of the Econometric Society*, 641–654.

Appendix: Proofs

Proof of Lemma 1. Consider the posterior distributions of $L(\alpha_t)$ and $L(\alpha_{t+1})$. We have

$$\rho_j(\alpha_{t+1}) - \rho_j(\alpha_t) = (\alpha_{t+1} - \alpha_t)(p_j - q_j).$$

The result is then obvious for *icx*, *st*, and *sc* from their definitions. For the *lr* order, let

$$l_j \equiv \frac{\rho_j(\alpha_{t+1})}{\rho_j(\alpha_t)} = \frac{\alpha_{t+1} \cdot (p_j - q_j) + q_j}{\alpha_t \cdot (p_j - q_j) + q_j}$$

and the result is implied by $l_{j+1} - l_j \propto (\alpha_{t+1} - \alpha_t) \left(\frac{p_{j+1}}{q_{j+1}} - \frac{p_j}{q_j} \right)$. \square

Proof of Lemma 2. Observe $L_1 \prec_{sc} L_2 \Rightarrow L_1 \prec_{st} L_2 \Rightarrow L_1 \prec_{icx} L_2$. Moreover,

$$\begin{aligned} \Delta(y+1) - \Delta(y) &= \sum_{z=y+1}^k (q_z - p_z)(z - y - 1) - \sum_{z=y}^k (q_z - p_z)(z - y) \\ &= \sum_{z=y+1}^k (p_z - q_z) = \bar{F}_1(y) - \bar{F}_2(y) \leq 0, \end{aligned}$$

where $\bar{F}_i(y) = 1 - F_i(y)$ is the survival function. Note that $\Delta(k) = 0$ and $\Delta(k-1) = q_k - p_k > 0$, and hence $\Delta(y) > 0$ for all $y < k$ by monotonicity. \square

Proof of Lemma 3. i) For any $\alpha \in [0, 1]$, we have

$$U(y+1, \alpha) - U(y, \alpha) = \sum_{z=y+1}^k \rho_z(\alpha)(z - y - 1) - \sum_{z=y}^k \rho_z(\alpha)(z - y) = - \sum_{z=y+1}^k \rho_z(\alpha) = -\bar{F}(y, \alpha)$$

where $F(y, \alpha) := \sum_{z=1}^y \rho_z(\alpha)$ and $\bar{F}(y, \alpha) = 1 - F(y, \alpha)$. Then for any $y' > y$,

$$U(y', \alpha) - U(y, \alpha) = \sum_{z=y}^{y'-1} [U(z+1, \alpha) - U(z, \alpha)] = - \sum_{z=y}^{y'-1} \bar{F}(z, \alpha) < 0.$$

The second order difference gives the convexity as follows

$$[U(y+2, \alpha) - U(y+1, \alpha)] - [U(y+1, \alpha) - U(y, \alpha)] = \bar{F}(y, \alpha) - \bar{F}(y+1, \alpha) = \rho_{y+1}(\alpha) > 0.$$

ii) By Lemma 2, $\frac{\partial U}{\partial \alpha} = U_1(y) - U_2(y) = -\Delta(y) \leq 0$, and for any $y' > y$,

$$\frac{\partial}{\partial \alpha} [U(y', \alpha) - U(y, \alpha)] = \Delta(y) - \Delta(y') \geq 0,$$

as $\Delta(y)$ is decreasing in y . □

Proof of Theorem 1. For any given state (y, α) and any search outcome $X = x$, the new fallback value $y' = \max\{x, y\}$ and the new belief α' is derived by (3). We are interested in the conditions for (6) to hold. There are two cases.

- If $x \leq y$, the outcome x is censored, and the new fallback value $y' = y$.

$$U(y', \alpha') - U(y, \alpha) = U(y, \alpha') - U(y, \alpha) = (\alpha - \alpha') \Delta(y) \leq 0,$$

because $\Delta(y) \geq 0$ from Lemma 2, and $\alpha' - \alpha \propto F_1(y) - F_2(y) > 0$ from (10).

- If $x > y$, the outcome x is revealed and $y' = x$. From the decomposition expression (7),

– Term i) measures the fallback value effect. As shown in Lemma 3,

$$\text{i) } = U(x, \alpha) - U(y, \alpha) = - \sum_{z=y}^{x-1} \bar{F}(z, \alpha) < 0. \quad (30)$$

– Term ii) measures the learning effect, and

$$\text{ii) } = U(x, \alpha') - U(x, \alpha) = (\alpha - \alpha') \Delta(x). \quad (31)$$

We know $\alpha' - \alpha \propto p_x - q_x$ and $\Delta(x) \geq 0$. There are further two subcases. (1) If $x \leq \hat{x}$, i.e., $p_x - q_x \geq 0$, then $\alpha' - \alpha \geq 0$. Term ii) ≤ 0 , and hence $U(y', \alpha') - U(y, \alpha) \leq 0$. (2) If $x > \hat{x}$, i.e., $p_x - q_x < 0$, then $\alpha' - \alpha < 0$. The learning effect ii) ≥ 0 . The sign of $U(y', \alpha') - U(y, \alpha)$ is determined by the relative magnitudes of the negative fallback value effect i) and the positive learning effect ii). From (30) and (31), $U(y', \alpha') - U(y, \alpha) \leq 0$ iff (12) holds.

We next show the sufficient condition (13) implies (12). Consider any $x > \max\{y, \hat{x}\}$. First,

$$\sum_{z=y}^{x-1} \bar{F}(z, \alpha) \geq \bar{F}(y, \alpha) = 1 - F(y, \alpha) \geq 1 - F_1(y) \geq 1 - F_1(x-1),$$

where $F(y, \alpha) = \alpha F_1(y) + (1 - \alpha) F_2(y) \leq F_1(y)$ because $F_1(y) \geq F_2(y)$ by Assumption 1. Second, observe that $r_x := \frac{p_x}{q_x} < 1$ for all $x > \max\{y, \hat{x}\}$, and we have

$$\alpha - \alpha' = \frac{\alpha(1 - \alpha)}{\alpha r_x + (1 - \alpha)} (1 - r_x)$$

achieves its maximum at $\alpha^* = (1 + \sqrt{r_x})^{-1}$. It then follows that

$$(\alpha - \alpha') \Delta(x) \leq \frac{\alpha^*(1 - \alpha^*)}{\alpha^* r_x + (1 - \alpha^*)} (1 - r_x) \Delta(x) = \frac{1 - \sqrt{r_x}}{1 + \sqrt{r_x}} \Delta(x)$$

We then have the state-free sufficient condition (13) for monotonicity. \square

Proof of Lemma 4. The results are directly implied by Lemma 2 and Lemma 3. \square

Proof of Lemma 5. $U(y, \alpha) = \alpha U_1(y) + (1 - \alpha) U_2(y)$ is decreasing in α by Assumption 1. Moreover, for any $y \in \mathcal{Y}$, $U(y, 1) = U_1(y) \leq c$, and the set $\{\alpha : U(y, \alpha) \leq c\}$ is nonempty. Therefore $\bar{\alpha}(y)$ is well-defined and satisfies: $\alpha \geq \bar{\alpha}(y) \Leftrightarrow U(y, \alpha) \leq c$. We then have, in a monotone search problem, it is optimal to stop iff $\alpha \geq \bar{\alpha}(y)$. \square

Proof of Proposition 1. For any $y \geq y_2^*$, stopping is immediate, and therefore, $\bar{n}(y, \alpha) = 0$ for all $y \geq y_2^*$. We next prove the result by backward induction.

We start from $y = y_2^* - 1$. If the next observation X is uncensored, then the renewed fallback value $X \geq y_2^*$, and stopping occurs after one period. If the next observation is censored, then the fallback value remains y and the odds ratio after one period is $O_1(y) = O(\alpha)R(y)$ from (18). After $m_y(\alpha)$ consecutive censored observations, the odds ratio exceed $\bar{O}(y)$ by (19); then the optimal policy stops after $m_y(\alpha)$ periods. Therefore, starting from state $(y_2^* - 1, \alpha)$, the optimal policy stops within $\bar{n}(y_2^* - 1, \alpha) := m_{y_2^* - 1}(\alpha) + 1$ periods.

As an induction hypothesis, suppose the claim holds for every fallback $j \in \{y + 1, \dots, y_2^* - 1\}$ and all belief $\alpha \in (0, 1)$, with corresponding finite bounds $\bar{n}(j, \alpha)$. Let us consider a state (y, α) .

Case 1: The search outcomes remain censored for $m_y(\alpha)$ consecutive periods. Then the fallback value stays at y and odds ratios evolve as $O_{t+1} = O_t R(y)$ from (18). By definition of (19), after $m_y(\alpha)$ consecutive censored outcomes we have $O \geq \bar{O}(y)$, hence the optimal policy stops in $m_y(\alpha)$ periods.

Case 2: An uncensored outcome occurs after $m < m_y(\alpha)$ consecutive censored periods. Let the first uncensored outcome be $X = x > y$, and at that moment the fallback jumps to $y' = x \geq y + 1$ and the updated posterior belief, denoted by $\alpha'_x \in (0, 1)$, is derived from (3). If $y' \geq y_2^*$, then stopping is immediate. Otherwise, by the induction hypothesis the policy stops within $\bar{n}(y', \alpha'_x)$ additional periods. Hence, the optimal policy stops within

$$1 + \max_{x \in \{y+1, \dots, k\}} \bar{n}(x, \alpha'_x)$$

periods after the uncensored event. Taking the maximum of the bounds in the two cases yields a finite uniform bound, which completes the proof. \square

Proof of Lemma 6. (1) $L_1 \prec_{sc} L_2$ implies $L_1 \prec_{st} L_2$ and $L_1 \prec_{icx} L_2$. Moreover, observe that

$$\Delta(y+1) - \Delta(y) = \sum_{z=1}^{y+1} (q_z - p_z)(y+1-z) - \sum_{z=1}^y (q_z - p_z)(y-z) = \sum_{z=1}^y (q_z - p_z) \leq 0.$$

(2) First, for monotonicity, we have

$$U(y+1, \alpha) - U(y, \alpha) = \sum_{z=1}^{y+1} \rho_z(\alpha)(y+1-z) - \sum_{z=1}^y \rho_z(\alpha)(y-z) = \sum_{z=1}^y \rho_z(\alpha) = F(y, \alpha) > 0.$$

Then for any $y' > y$,

$$U(y', \alpha) - U(y, \alpha) = \sum_{z=y}^{y'-1} [U(z+1, \alpha) - U(z, \alpha)] = \sum_{z=y}^{y'-1} F(z, \alpha) > 0.$$

The second order difference gives the convexity as follows.

$$[U(y+2, \alpha) - U(y+1, \alpha)] - [U(y+1, \alpha) - U(y, \alpha)] = F(y+1, \alpha) - F(y, \alpha) = \rho_{y+1}(\alpha) > 0.$$

Second, it is clear that $\frac{\partial U}{\partial \alpha} = U_1(y) - U_2(y) = -\Delta(y) \geq 0$, and for $y' > y$,

$$\frac{\partial}{\partial \alpha} [U(y', \alpha) - U(y, \alpha)] = \Delta(y) - \Delta(y') \geq 0,$$

as $\Delta(y)$ is decreasing in y . □

Proof of Proposition 2. For any given state (y, α) and a search outcome $X = x$. The new fallback value $y' = \min\{x, y\}$ and the new belief α' is derived from (23). We are interested in deriving the conditions for (6) to hold. There are two cases.

- If $x \geq y$, the search outcome x is censored. The new fallback value $y' = y$ and

$$U(y', \alpha') - U(y, \alpha) = U(y, \alpha') - U(y, \alpha) = -(\alpha' - \alpha) \Delta(y) \leq 0,$$

because $\Delta(y) \leq 0$ and $\alpha' - \alpha \propto \bar{F}_1(y) - \bar{F}_2(y) \leq 0$ by (23).

- If $x < y$, the outcome x is revealed and $y' = x$. From the decomposition expression (7),

– Term i) measures the fallback value effect. As shown in Lemma 6,

$$\text{i) } = U(x, \alpha) - U(y, \alpha) = - \sum_{z=x}^{y-1} F(z, \alpha) < 0. \quad (32)$$

– Term ii) measures the learning effect, and

$$\text{ii) } = U(x, \alpha') - U(x, \alpha) = (\alpha - \alpha') \Delta(x) \quad (33)$$

We know $\Delta(x) \leq 0$ and $\alpha' - \alpha \propto p_x - q_x$. There are two further subcases. (1) If $x \geq \hat{x}$, i.e., $p_x - q_x \leq 0$, then $\alpha' - \alpha \leq 0$. In this case, ii) ≤ 0 , and hence $U(y', \alpha') - U(y, \alpha) \leq 0$. (2) If $x < \hat{x}$, i.e., $p_x - q_x > 0$, then $\alpha' - \alpha > 0$, and the learning effect ii) ≥ 0 . The sign of $U(y', \alpha') - U(y, \alpha)$ is determined by the relative magnitudes of the negative fallback value effect i) and the positive learning effect ii). From (32) and (33), $U(y', \alpha') - U(y, \alpha) \leq 0$ if and only if (24) holds.

We next show (25) implies condition (24). For any $x < \min\{y, \hat{x}\}$, first observe that

$$\sum_{z=x}^{y-1} F(z, \alpha) \geq F(x, \alpha) \geq F_2(x),$$

where we apply $F(x, \alpha) = \alpha F_1(x) + (1 - \alpha) F_2(x) \geq F_2(y)$ by Assumption 1. Second, for $x < \min\{y, \hat{x}\}$, denote $r_x := \frac{p_x}{q_x} > 1$ and we have

$$\alpha - \alpha' = \frac{\alpha(1 - \alpha)}{\alpha r_x + (1 - \alpha)} (1 - r_x) < 0$$

which achieves its minimum at $\alpha^* = (1 + \sqrt{r_x})^{-1}$. As $\Delta(x) \leq 0$, it then follows

$$(\alpha - \alpha') \Delta(x) \leq \frac{\alpha^*(1 - \alpha^*)}{\alpha^* r_x + (1 - \alpha^*)} (1 - r_x) \Delta(x) = \frac{1 - \sqrt{r_x}}{1 + \sqrt{r_x}} \Delta(x)$$

We then have the sufficient condition (25) for monotonicity. □

Proof of Proposition 3. Suppose the monotone condition (6) holds under full revelation. Now consider lower-censorship. For any state (y, α) and any new outcome $X = x$:

- If $x > y$, then the value x is revealed and the posterior belief α' is the same as that under full revelation, and therefore, condition (6) is satisfied.
- If $x \leq y$, then the value is censored and the posterior belief, by the law of total expectation,

$$\alpha' = \mathbb{P}(L = L_1 | \alpha, X \leq y) = \sum_{x=1}^y \mathbb{P}(X = x | X \leq y) \mathbb{P}(L = L_1 | \alpha, X = x) = \mathbb{E}(\alpha'(X) | X \leq y).$$

Since $U(y, \alpha)$ is linear in α , we have

$$U(y, \alpha') = U(y, \mathbb{E}(\alpha'(X) | X \leq y)) = \mathbb{E}[U(y, \alpha'(X)) | X \leq y].$$

As condition (6) is satisfied under full revelation, then for all $x \leq y$, $U(y, \alpha) \geq U(y, \alpha(x))$.

Taking conditional expectation then gives $U(y, \alpha) \geq \mathbb{E}[U(y, \alpha'(X)) | X \leq y] = U(y, \alpha')$. □

Proof of Theorem 4. Sufficient part. Under condition (29), $q_x \leq p_x$ for all $x \leq k - 1$. Given any state (y, α) and any new search outcome $X = x$, there are two cases.

- If $x \leq y$, then the new fallback value $y' = y$ and

$$U(y', \alpha') - U(y, \alpha) = (\alpha' - \alpha)[U_1(y) - U_2(y)] \propto (q_x - p_x) \Delta(y),$$

because $\Delta(y) \geq 0$ for all $y \in K$ under Assumption 1, i.e., $L_1 \prec_{sc} L_2$, and $\alpha - \alpha' \propto q_x - p_x$ by (2). If $y = k$, then $\Delta(y) = 0$ and (6) is satisfied. If $y < k$, then $x \leq y \leq k - 1$, and hence $\alpha - \alpha' \leq 0$ under condition (29). The monotone condition (6) is also satisfied.

- If $x > y$, then $y' = x$ and, from the decomposition expression (7),
 - the fallback value effect i) $= U(x, \alpha) - U(y, \alpha) < 0$ (Lemma 3).
 - the learning effect ii) $= (\alpha' - \alpha)[U_1(x) - U_2(x)] \propto (q_x - p_x) \Delta(x)$. In this case, if $x \leq \hat{x}$, then ii) is negative as $q_x \leq p_x$ under condition (29). If $x > \hat{x}$, then $x = k$ also under condition (29), and $\Delta(x) = \Delta(k) = 0$.

Necessary part. Suppose the monotone condition (6) holds. Fix any $y \in K$ and any $x \leq y$. We have $y' = y$ and α' is derived from (2). It then follows from (6) that

$$U(y', \alpha') - U(y, \alpha) = (\alpha' - \alpha)[U_1(y) - U_2(y)] \propto (q_x - p_x) \Delta(y) \leq 0.$$

For $y = k$, $\Delta(k) = 0$ and the above inequality holds. For $y = k - 1$, $\Delta(k - 1) = q_k - p_k > 0$ and the above inequality implies, for all $x \leq k - 1$, $(q_x - p_x)(q_k - p_k) \leq 0$, and then $q_x - p_x \leq 0$. This gives condition (29). \square

Proof of Lemma 7. Define $r_j = \frac{p_j}{q_j}$. From condition (29) and $p_j \neq q_j$ for any $j \in K$, we have $r_j > 1$ for any $j < k$. For any given state (y, α) with $y \in \mathcal{Y} \setminus \{k\}$ and $\alpha \in (0, 1)$, define $\underline{r}(y) := \min_{j \leq y} \{r_j\} > 1$.

If there are m consecutive search outcomes that are smaller than y , the updated new odds ratio is weakly greater than $O(\alpha)\underline{r}(y)^m$. For any $y \in \mathcal{Y}$ and $\alpha \in (0, 1)$, define the stopping upper

bound for “consecutive-smaller” periods by

$$\bar{m}_y(\alpha) := \min \left\{ m \in \mathbb{N}_0 : O(\alpha) \cdot \underline{r}(y)^m \geq \bar{O}(y) \right\}. \quad (34)$$

That is, given a state (y, α) , under full revelation, $\bar{m}_y(\alpha)$ is the upper-bound of search periods if the sample values keep on being smaller than y . For example, for any $y \geq y_2^*$, $\bar{\alpha}(y) = 0$ and hence, $\bar{O}(y) = 0$ and $\bar{m}_y(\alpha) = 0$. For $y = y_1^*$, $\bar{\alpha}(y_1^*) < 1$ by Assumption 2, then $\bar{O}(y_1^*)$ is finite, and the DM must stop within $\bar{m}_\alpha(y_1^*)$ periods if the sample values keep on being smaller.

The remaining proof for a finite bound $n^f(y, \alpha)$ remains the same as that of Proposition 1 and hence is omitted here. To show $n^f(y, \alpha) \geq n(y, \alpha)$, it suffices to show $\bar{m}_\alpha(y) \geq m_\alpha(y)$, or equivalently, $\underline{r}(y) \leq R(y) = \frac{F_1(y)}{F_2(y)}$. By definition, for any $j = 1, \dots, y$, $\underline{r}(y) \leq r_j$ and hence $q_j \underline{r}(y) \leq p_j$. Summing up the inequalities gives

$$\underline{r}(y) \sum_{j=1}^y q_j \leq \sum_{j=1}^y p_j,$$

and therefore, $\underline{r}(y) \leq \frac{F_1(y)}{F_2(y)} = R(y)$. □

Proof of Theorem 2

Recall the space $K \times \Delta(\{1, \dots, \ell\})$ and the operator T defined on it. For two real functions f, g defined on this space, we say that $f \leq g$, if it occurs pointwise: for every $(y, \vec{\alpha})$, $f(y, \vec{\alpha}) \leq g(y, \vec{\alpha})$. In the sequel, a function is referred to as a real function defined on $K \times \Delta(\{1, \dots, \ell\})$.

Lemma 8. *Suppose that f, g are two functions such that $f \leq g$. Then, $Tf \leq Tg$.*

Proof. Recall the definition of T is (27). Tf and Tg evaluated at the same point $(y, \vec{\alpha})$. Share the same fallback value and the same posteriors $\text{post}(\vec{\alpha} | X_{\vec{\alpha}})$. But the due to the assumption,

$$f(\max\{y, X_{\vec{\alpha}}\}, \text{post}(\vec{\alpha} | X_{\vec{\alpha}})) \leq g(\max\{y, X_{\vec{\alpha}}\}, \text{post}(\vec{\alpha} | X_{\vec{\alpha}}))$$

with probability 1, and therefore the expectation over these two random variables preserves the same inequality. Thus, $Tf \leq Tg$. □

We now define a sequence of functions inductively on which the rest of the proof resides. Let $f_0(y, \vec{\alpha}) = y$ for every $(y, \vec{\alpha})$. Then for every integer n , set $f_{n+1} = Tf_n$.

Lemma 9. *The sequence $(f_n)_n$ is increasing (e.g., for every $\vec{\alpha}$ and n , $f_n \leq f_{n+1}$).*

Proof. We proceed by induction on n . It is immediate that $f_0 \leq f_1$, since f_1 is defined as the maximum of f_0 and something else.

Assume now that $f_{n-1} \leq f_n$. The previous lemma implies that $f_n = Tf_{n-1} \leq Tf_n = f_{n+1}$, as desired. \square

Since $(f_n)_n$ is bounded and monotonically increasing, the limit exists. Denote it by f_∞ .

Lemma 10. *f_∞ is a fixed point of T .*

Proof. Since $f_n \leq f_\infty$ for every n and T is monotonic, we have $Tf_n \leq Tf_\infty$ for every n . Taking limits yields $f_\infty \leq Tf_\infty$.

Suppose, toward a contradiction, that there exist $(y, \vec{\alpha})$ and $\varepsilon > 0$ such that

$$f_\infty(y, \vec{\alpha}) + \varepsilon < Tf_\infty(y, \vec{\alpha}).$$

Since $f_\infty = \lim_n f_n$, we have

$$Tf_\infty(y, \vec{\alpha}) = \lim_{n \rightarrow \infty} \max\{y, \mathbb{E}[f_n(\max\{y, X_{\vec{\alpha}}\}, \text{post}(\vec{\alpha} | X_{\vec{\alpha}}))] - c\} > f_\infty(y, \vec{\alpha}) + \varepsilon.$$

Hence, for n sufficiently large,

$$f_{n+1}(y, \vec{\alpha}) = \max\{y, \mathbb{E}[f_n(\max\{y, X_{\vec{\alpha}}\}, \text{post}(\vec{\alpha} | X_{\vec{\alpha}}))] - c\} > f_\infty(y, \vec{\alpha}) + \varepsilon/2.$$

This is impossible, since $f_n \leq f_\infty$ for all n . Therefore, $f_\infty = Tf_\infty$. \square

Lemma 11. *If J is a fixed point of T , then $f_\infty \leq J$.*

Proof. Clearly, $f_0 \leq J$. By Lemma 8 and the construction, $f_n = T^n f_0$, we have

$$f_n = T^n f_0 \leq T^n J = J \quad \text{for all } n.$$

Taking limits preserves the inequality, and hence $f_\infty \leq J$. □

We now turn to finite-horizon Bayesian search problems. Suppose that the search process proceeds as usual until time n . If the decision maker has not stopped before time n , the existing fallback payoff is the outcome of the dynamic programming. Our goal is to relate the optimal strategy of the n -period search problems to the sequence $(f_n)_n$.

Fix n and define the stopping time τ_n as follows. Consider any history of search outcomes of length t that has not yet led to stopping, and let $(y, \vec{\alpha})$ denote the induced fallback payoff and posterior belief. The stopping rule is:

$$\tau_n = t \quad \text{if and only if} \quad f_n(y, \vec{\alpha}) = y.$$

In other words, $f_n(y, \vec{\alpha}) = y$ implies stopping. If $f_n(y, \vec{\alpha}) > y$, the stopping time remains unspecified, which corresponds to continuing the search.

The same definition applies with f_∞ , and the corresponding stopping time is denoted by τ_∞ . Note that τ_n is formally defined on the space of outcome histories; however, the induced state $(y, \vec{\alpha})$ fully determines, via f_n , whether the rule prescribes stopping or continuation.

Remark 12. Because the sequence (f_n) is increasing, whenever τ_n dictates continuation, so does τ_{n+1} and therefore, τ_∞ . But it might happen that after a certain history, τ_n dictates stopping, while τ_{n+1} dictates continuation. It implies that among the histories of a fixed length, N , the set of histories that follow stopping is diminishing with n .

Lemma 12. *The stopping time τ_n is optimal for the n -period search problem.*

Proof. Consider stage $n - 1$ of the search process, where the state is $(y, \vec{\alpha})$. Only one search opportunity remains. The decision maker can either stop and receive y , or continue and obtain the expected payoff

$$\mathbb{E}[\max\{y, X_{\vec{\alpha}}\}] - c.$$

The optimal choice yields $f_1(y, \vec{\alpha})$, and τ_1 implements this choice.

Now consider stage $n - 2$, where two search opportunities remain, and the state is again

$(y, \vec{\alpha})$. The decision maker can either stop with a payoff y , or continue and obtain

$$\mathbb{E}[f_1(\max\{y, X_{\vec{\alpha}}\}, \text{post}(\vec{\alpha} \mid X_{\vec{\alpha}}))] - c.$$

The maximum of these two payoffs is $f_2(y, \vec{\alpha})$, which is implemented by τ_2 .

Proceeding by backward induction, we conclude that τ_n is optimal for the n -period search problem, and that the maximal expected net payoff starting from $(y, \vec{\alpha})$ is $f_n(y, \vec{\alpha})$ (with the usual initial condition $(0, \vec{\alpha}_0)$). \square

Lemma 13. *For any integer N there exists an $\varepsilon(N) > 0$ such that for every n ,*

$$\mathbb{P}(\tau_n > N) < \varepsilon(N),$$

and $\varepsilon(N) \rightarrow 0$ when $N \rightarrow \infty$.

Proof. If this were not the case, then for there is $\varepsilon > 0$ s.t. for infinitely many N there would exist n such that

$$\mathbb{P}(\tau_n > N) \geq \varepsilon.$$

In that event, with probability at least ε the search cost c is incurred at least N times, implying an expected payoff no greater than

$$k - \varepsilon N c.$$

For N sufficiently large, this quantity becomes negative, which contradicts Lemma 12. \square

Corollary 1. *For every integer N there is an integer S s.t. all τ_n , $n \geq S$ and τ_∞ agree on histories shorter of N .*

Proof. Remark 12 and Lemma 13 imply that the collection of histories at which stopping occurs is monotone nondecreasing in n . Since the set of histories of length less than N is finite, this monotone sequence must stabilize. Consequently, there exists S such that for all $n \geq S$, the stopping times coincide on every history of length less than N . \square

Lemma 14. (i) $\mathbb{E}[\tau_\infty] < \infty$.

(ii) $\mathbb{E}[\tau_\infty \mid \tau_\infty > N]$ tends to 0 as $N \rightarrow \infty$.

Proof. If $\mathbb{E}[\tau_\infty] = \infty$, there is an N s.t. $\mathbb{E}[\tau_\infty \mid \tau_\infty \leq N] > 2k/c$. By Corollary 1, there is an S s.t. all τ_n , $n \geq S$ and τ_∞ agree on histories shorter of N . It implies that the yield of such τ_n is not greater than

$$\begin{aligned} & \mathbb{E}[X_{\tau_n} - \tau_n c \mid \tau_n \leq N] \mathbb{P}(\tau \leq N) + \mathbb{P}(\tau > N)k \leq \\ & \mathbb{E}[X_{\tau_n}] \mathbb{P}(\tau \leq N) - (2k/c)c + \mathbb{P}(\tau > N)k \leq k - 2k < 0. \end{aligned}$$

This is impossible, and we conclude (i). (ii) is an immediate consequence of it. \square

The next lemma shows that the minimal fixed point f_∞ characterizes the optimal strategy of the infinite-horizon problem.

Lemma 15. *The stopping time τ_∞ is optimal for the unbounded search problem.*

Proof. Fix N . By Lemma 13 and Corollary 1 there is N and S s.t. the payoff τ_n yields is at most

$$\begin{aligned} & \mathbb{E}[X_{\tau_n} - \tau_n c \mid \tau_n \leq N] \mathbf{p}(\tau_n \leq N) + \mathbf{p}(\tau_n > N)k \leq \\ & \mathbb{E}[X_{\tau_n} - \tau_n c \mid \tau_n \leq N] + \varepsilon(N)k, \end{aligned} \tag{35}$$

for every $n \geq S$.

On the other hand, τ_∞ yields at least

$$\begin{aligned} & \mathbb{E}[X_{\tau_\infty} - \tau_\infty c \mid \tau_\infty \leq N] \mathbb{P}(\tau_\infty \leq N) + \mathbb{E}[X_{\tau_\infty} - \tau_\infty c \mid \tau_\infty > N] \mathbb{P}(\tau_\infty > N) \geq \\ & \mathbb{E}[X_{\tau_n} - \tau_\infty c \mid \tau_n \leq N] \mathbb{P}(\tau_n \leq N) - c \mathbb{E}[\tau_\infty \mid \tau_\infty > N]. \end{aligned} \tag{36}$$

The difference between (35) and (36) is

$$\varepsilon(N)k - c \mathbb{E}[\tau_\infty \mid \tau_\infty > N],$$

which converges to 0 as $N \rightarrow \infty$. Indeed, as $N \rightarrow \infty$: (a) $\varepsilon(N) \rightarrow 0$ by Lemma 13; and (b) $\mathbb{E}[\tau_\infty \mid \tau_\infty > N] \rightarrow 0$ by Lemma 14. We conclude that τ_∞ yields at least $f_n(0, \vec{\alpha}_0)$ for every n and therefore $f_n(0, \vec{\alpha}_0)$.

It remains to show that there is no strictly better stopping time for the unbounded search problem. If there would be such, there would then be a finite time, n at which the search process could have stopped and yield strictly more than $f_n(0, \vec{\alpha}_0)$. This would contradict Lemma 12. We conclude that τ_∞ is an optimal stopping rule for the unbounded Bayesian search problem, as desired \square

The previous lemma completes the proof of Theorem 2. First, the lowest fixed point of the operator T is f_∞ , which is obtained constructively as the limit of the sequence $(f_n)_n$. Second, the stopping rule τ_∞ induced by f_∞ is optimal for the unbounded search problem and constitutes the minimal solution to (26).

An additional outcome of the proof is a transparent connection between finite-horizon search problems and their optimal stopping rules, on the one hand, and the unbounded search problem and its optimal stopping rule, on the other.

Proof of Theorem 3

As in the proof of Theorem 2 we define a sequence of functions inductively. Let $f_0(\max\{y, O\}, \vec{\alpha}) = \max\{y, O\}$ for every $(y, O, \vec{\alpha})$. Then for every integer n , set $f_{n+1} = Tf_n$, where

$$Tg(\max\{y, O\}, \vec{\alpha}) = \max\{y, O, \mathbb{E}[g(\max(y, O, X_{\vec{\alpha}}), \text{post}(\vec{\alpha} | X_{\vec{\alpha}}))] - c\}. \quad (37)$$

The proof follows the footsteps of the previous one.