Rational Search with Learning[*]

T. Fung Ngai[†]

November 17, 2026

[Click here for most recent version]

Abstract

I explore Bayes rational search for prizes (such as wages or prices) with learning. A searcher Sam is impatient and pays a cost for every search. He draws prize offers from a distribution with an uncertain mean, where observed draws constitute informative signals of the true mean via a location family with the monotone likelihood ratio property. A higher draw is myopically more inviting, but also raises expectations of better future draws. I show that a logconcave prior on the mean prize guarantees that a standard reservation rule is optimal. By contrast, in a highly uncertain environment with a heavy-tailed prior distribution, a sufficiently patient Sam accepts some prizes below those he rejects. I show that if Sam searches long enough, a reservation policy is once again optimal.

With a prior over finitely many possible means, reservation policy failures arise with enough dispersion.

Keywords: Sequential search, optimal stopping, learning, reservation wage, logconcavity, heavy tails.
JEL codes: D81, D83.

[*]I am deeply grateful to my advisors Lones Smith, Benjamin Bernard, and Marek Weretka for their invaluable guidance. The author would also like to thank Matteo Camboni, Daniel Quint, Marzena Rostek, and Rodrigo Yañez as well as participants in UW–Madison Theory Lunch, HKEA Biennial Conference, and UW–Madison Theory Seminar for their helpful comments and feedback.

[†]Department of Economics, University of Wisconsin–Madison. E-mail: tngai@wisc.edu.

1 Introduction

The simplest dynamic search problem confronts a decision maker with a random i.i.d. sequence of offers (prices, wages, prizes) from a known distribution; a searcher must choose a single one, and until that moment, he pays either an explicit cost or suffers a delay time cost. In the job search problem, for instance, it is optimal to set a reservation wage, accepting all wages above and rejecting all wages below (McCall, 1970). Information acquisition is another forward-looking dynamic optimization. Its optimal solution likewise employs cutoffs — to stop when sufficiently confident of a state of the world (Moscarini and Smith, 2001). This paper develops a theory for the marriage of these paradigms, such as price or wage search with learning about the distribution. While one could actively learn by doing market research, it's simplest to assume that people just extract information from their search experience. This is especially relevant when searching for the first time, or after a time span when the world has evolved.

In this paper, we consider search without recall, such as looking for a job and selling a house. We specifically assume that the prize offers are drawn from a location family, and the searcher is uncertain of the location parameter, which we call the state. So, each wage or price offer is an informative signal about the state. For instance, Sam is a fresh graduate who is uncertain about the demand for his skills. His wage offer combines a persistent component—the general quality of his employment opportunities—with a component specific to the employer making the offer (i.e., the noise). DeGroot (1968) finds that if the noise distribution and the prior belief on state are both Gaussian, a standard cutoff rule is optimal. This paper provides sharp and robust sufficient conditions on when a reservation policy is optimal and when it is not, and subsumes the double Gaussian setting as a special case.

In models without learning, the search optionality is well-understood. The expected value of having an additional draw depends only on the known distribution. By contrast, when the distribution of offers is unknown, each observation serves a dual role: it provides an immediate payoff opportunity and simultaneously conveys information about the state of the world. While searching, the searcher learns about the prize distribution and updates his belief. We employ stochastic dynamic programming and associated recursion logic to characterize the Bellman value function (Proposition 1 and Lemma 2).

Learning fundamentally alters the nature of search. The searcher accepts an offer iff the acceptance payoff exceeds the expected rejection gain, which is the discounted value of continued search. While it is clear that the acceptance gain rises linearly in the offer, the rejection value increment depends on the rise in the posterior belief. A

higher offer may substantially shift beliefs up that the rejection gain in Bellman values swamps the linear gain in acceptance value. We capture that information effect by a payoff-score covariance, and show that a reservation policy is optimal if that covariance is less than one (Lemma 4). The payoff-score covariance can be decomposed into three terms: the expected marginal payoff, a boundary term, and a prior-shape term, where the expected marginal payoff is at most one and the boundary term nonpositive. We prove that when the prior belief admits a logconcave density, the prior-shape term must also be nonpositive, implying that the covariance is at most one. We thereby obtain the sharp result that a reservation policy is optimal with logconcave belief, such as Gaussian (Proposition 2). Since belief updating preserves logconcavity, we show that if the belief density is logconcave at some date, then reservation policies are optimal thenceforth after every finite history (Corollary 1).

Beliefs with logconcave density are thin-tailed. But many important distributions commonly used in economics are heavy-tailed, and they are often used to model uncertain environments in which extreme or rare events occur far more frequently than a Gaussian distribution would predict. We consider a broad class of heavy-tailed beliefs, namely, those with eventually decreasingly logconvex density. This is a large class, including many widely used distributions such as lognormal, Pareto, and Student's t-distributions. For these beliefs, we show that a sufficiently patient searcher accepts some medium offers but rejects higher ones over an interval of search costs (Proposition 3). Nonetheless, if the belief density has a bounded log-curvature (as the three heavy-tailed belief examples do), it will be logconcave after a finite horizon as belief updating fosters logconcavity, and reservation policies are optimal thenceforth (Proposition 4).

The tail of the prior belief density matters because it determines how much Sam attributes a good offer to the particular opportunity and to the environment in which he is searching. In a job search setting, an unusually favorable offer can have two explanations: (a) This particular employer is offering Sam an unusually good opportunity; and (b) Sam's employment prospects are generally better than he thought. With a logconcave prior, as the offer rises, Sam tends to attribute that to employer-specific luck. Heavy-tailed prior beliefs allow very favorable observations to be explained predominantly by a favorable persistent environment. That effect can be so strong that Sam continues searching.

Last but not least, with a discrete prior belief over finitely many possible states, we show that a reservation policy fails to be optimal if the states are dispersed enough. The main body presents that result for a simple two-state setting (Proposition 5), which can be easily extended to finitely many states as illustrated in the appendix.

Related Literature. Among the first to point out that a reservation policy may not be optimal in general, Rothschild (1974) explores a model with only a finite number of possible prizes, and finds that the reservation property holds if the searcher has a Dirichlet prior on the multinomial distributions. Convenient though it is, the updating rule implied by the Dirichlet prior is very special: the information gained is strictly local in the sense that observing an offer raises the likelihood of that particular offer, but reduces the likelihood of *all* other offers *by the same proportion*. Bikhchandani and Sharma (1996) generalize the idea of local information gain to non-Bayesian learning. In contrast, our model allows for a more natural possibility that observing a high offer may reduce the likelihood of a low offer more than that of a medium one.

Albright (1977) studies search with learning for some specific conjugate families. Rosenfield and Shapiro (1981) and Seierstad (1992) provide sufficient conditions for the optimality of a reservation policy on the predictive distribution of offer as a function of a vector of observed offers. Gershkov and Moldovanu (2012) propose a condition that implies those of Rosenfield and Shapiro and Seierstad. Employing dynamic programming and associated recursion logic, this paper provides simple conditions on the belief for when a reservation policy is optimal and when it is not.

Burdett and Vishwanath (1988) impose endogenous assumptions on the value function under which a reservation policy is optimal and the reservation wage declines if the search cost is large enough. Not only does this paper illustrate that those endogenous assumptions do not hold in general, but our results also provide primitive foundations for them when they do hold. Similarly, Adam (2001) considers search with recall and makes an assumption to rule out strong positive learning effects.

On the experimental side, Cox and Oaxaca (2000) test finite-horizon search from unknown wage offer distributions in the laboratory. In their design, offers are drawn from one of two discrete uniform distributions that differ by a location shift—a discrete-offer analogue of the two-state belief in Section 7. Subjects' search behavior is consistent with a risk neutral Bayesian in the large majority of trials, whether or not a reservation policy is optimal.

The remainder of our paper is organized as follows. In Section 2, we lay out our model of search with learning, and in Section 3 we characterize the optimal policy by the value function. Section 4 provides conditions on when a reservation policy is optimal and when it is not. In Section 5, we show that a reservation policy is optimal when the belief has a logconcave density while in Sections 6 and 7 we show its failure under a class of heavy-tailed beliefs and discrete beliefs, respectively. The Appendix includes all proofs not included in the text.

2 The Model

Timing and the Search Problem. Periods are $t = 0, 1, 2, \ldots$ and payoffs are discounted by a factor $\beta \in (0, 1)$. At the beginning of each period t, the searcher Sam chooses between *quitting*, which ends the search with payoff zero, and *searching*.[1] If he searches, he pays the cost $c \geq 0$, observes an offer x_t, and chooses between *accepting*, which ends the search with payoff x_t, and *rejecting*, after which period $t + 1$ begins.

States and Offers. The state space Θ is a nonempty Borel subset of $\mathbb{R}$, possibly unbounded. Conditional on a state $\theta \in \Theta$, the offer in each period $t \geq 0$ is a random variable $x_t = \theta + \varepsilon_t$ drawn from a location family with location parameter θ. The noise terms $(\varepsilon_t)_{t=0}^{\infty}$ are iid with density g that is twice continuously differentiable and positive on $\mathbb{R}$.[2] So, the offers $(x_t)_{t=0}^{\infty}$ are conditionally iid with density $f(x|\theta) = g(x-\theta)$. We assume the signals $f(x|\theta)$ have the monotone likelihood ratio property (MLRP), and note that this has the intuitive property that a higher offer is unambiguously encouraging and shifts up beliefs in the sense of First-Order Stochastic Dominance (Whitt, 1979; Milgrom, 1981). Equivalently, the density g is logconcave. And because of the positivity of g, no offer serves as a perfectly revealing signal.

Beliefs. A belief μ is a Borel probability measure on Θ, and each belief yields a predictive density $f_\mu(x) = \int_\Theta f(x|\theta)\mu(d\theta)$ of the next offer. In the initial period $t = 0$, Sam holds a prior belief μ_0 satisfying the first moment existence condition $\int_\Theta |\theta|\mu_0(d\theta) < \infty$. In each period $t \geq 0$ after the offer x_t has been observed, Sam's belief is updated from μ_t to μ_{t+1} via Bayes' rule:

$$\mu_{t+1}(d\theta) = \mu_t^{x_t}(d\theta) := \frac{f(x_t|\theta)\,\mu_t(d\theta)}{f_{\mu_t}(x_t)}. \tag{1}$$

This model allows general beliefs, including both continuous and discrete distributions as special cases. We now argue that the existence of the posterior moments follows.

Lemma 1 (Posterior moments)**.** *Let $\mathcal{P}_1(\Theta)$ be the set of Borel probability measures μ on Θ satisfying $\int_\Theta |\theta|\mu(d\theta) < \infty$. Then $\mu \in \mathcal{P}_1(\Theta)$ implies $\mu^x \in \mathcal{P}_1(\Theta)$ for every $x \in \mathbb{R}$. It follows from $\mu_0 \in \mathcal{P}_1(\Theta)$ that $\mu_t \in \mathcal{P}_1(\Theta)$ for all $t \geq 1$.*

Proof. Since g is a logconcave density, $\|g\|_\infty = \sup_{\varepsilon\in\mathbb{R}} g(\varepsilon) < \infty$. Fix $\mu \in \mathcal{P}_1(\Theta)$ and

[1] We name the searcher Sam for Karlin (1962), an early contributor to the search literature. Sam is also an acronym for "search and matching."

[2] It covers Gaussian and logistic noise, among others.

$x \in \mathbb{R}$. Then

$$\int_\Theta |\theta|\, \mu^x(d\theta) = \frac{\int_\Theta |\theta| g(x-\theta)\, \mu(d\theta)}{f_\mu(x)} \leq \frac{\|g\|_\infty}{f_\mu(x)} \int_\Theta |\theta|\, \mu(d\theta) < \infty.$$

Hence $\mu^x \in \mathcal{P}_1(\Theta)$. Starting from $\mu_0 \in \mathcal{P}_1(\Theta)$, induction using (1) yields $\mu_t \in \mathcal{P}_1(\Theta)$ for all $t \geq 1$. □

Optimization. Denote by $h_t = (x_0, \ldots, x_{t-1}) \in \mathcal{H}_t = \mathbb{R}^t$ the offer history observed as of date t. A *policy* $\pi \in \Pi$ consists of Borel functions $\pi_t^s \colon \mathcal{H}_t \to \{0,1\}$ and $\pi_t^a \colon \mathcal{H}_{t+1} \to \{0,1\}$. Reaching period t with history h_t, Sam searches if $\pi_t^s(h_t) = 1$ and quits otherwise; and if he searches, he accepts x_t if $\pi_t^a(h_t, x_t) = 1$ and rejects it otherwise. The state-contingent *policy payoff* of π is the present value of the negative costs and the offer,

$$J_\pi(\theta) = \mathbb{E}_\theta\Big[\sum_{t=0}^{\infty} \beta^t \Big(\prod_{j=0}^{t-1} \pi_j^s(h_j)[1 - \pi_j^a(h_j, x_j)]\Big) \pi_t^s(h_t)\big\{-c + \pi_t^a(h_t, x_t)\, x_t\big\}\Big]. \tag{2}$$

The *policy value* of π is its expected policy payoff,

$$V_\pi(\mu_0) = \int_\Theta J_\pi(\theta)\, \mu_0(d\theta). \tag{3}$$

Sam solves

$$\bar{v}(\mu_0) = \max_{\pi \in \Pi} V_\pi(\mu_0). \tag{4}$$

The optimal policy is the argmax of this, and is unique up to indifference at reached histories. We select the optimal policy $\pi^*(\mu_0)$ that accepts when indifferent and quits when indifferent. This policy is picked for convenience, and the choice does not affect the main results of this paper. We define the state-contingent *optimal payoff* $J_{\mu_0}(\theta) := J_{\pi^*(\mu_0)}(\theta)$ given the prior belief μ_0 to be the policy payoff of $\pi^*(\mu_0)$.

3 The Dynamic Programming Solution

First, as is standard, Sam has an *optimal value* $\bar{v}(\mu)$.[3] Moreover, it obeys the recursion

$$\bar{v}(\mu) = \max\Big\{0,\ -c + \int_{\mathbb{R}} \max\{x, \beta \bar{v}(\mu^x)\} f_\mu(x)\, dx\Big\}. \tag{5}$$

[3]Sam faces a controlled Markov process where the distribution of the next offer and subsequent belief updating depend on the past only through the current belief. By Bertsekas and Shreve (1978), the belief is a sufficient statistic, and it is sufficient to use a Markov policy.

The outer maximum is the decision to quit, taken before the cost is sunk; the inner maximum is the decision to accept, taken after the offer is seen.

Proposition 1 (Optimal value)**.** *For every $\beta \in (0,1)$ and $c \geq 0$, the optimal value $\bar{v}$ is continuous and convex on $\mathcal{P}_1(\Theta)$. Moreover:*

(a) (cost monotonicity) *$c \mapsto \bar{v}_c(\mu)$ is convex, nonincreasing, and strictly decreasing wherever it is positive;*

(b) (discount factor monotonicity) *$\beta \mapsto \bar{v}_\beta(\mu)$ is nondecreasing.*

The proof in Appendix A argues that $\bar{v}$ is the unique fixed point of a contraction $\mathcal{T}$ with modulus β defined by

$$(\mathcal{T}v)(\mu) := \max\left\{0,\ -c + \int_{\mathbb{R}} \max\{x, \beta v(\mu^x)\} f_\mu(x)\, dx\right\}. \tag{6}$$

As is standard in information economics, the convexity of the optimal value captures the idea that information is valuable: given a chance to learn about the state of the world, Jensen's inequality implies that learning weakly raises Sam's expected optimal payoff, as Sam can adapt his policy to what he learns while retaining the option to ignore the new information and follow his original policy.

Next, we discuss the monotonicity of the optimal value. Since $\bar{v}$ is a function of belief, which is not a scalar, we introduce the likelihood ratio ordering, in which the optimal value is monotone. Let μ and ν be mutually absolutely continuous.[4] We say ν is larger than μ in the likelihood ratio order ($\nu \succeq_{\text{lr}} \mu$) if the Radon–Nikodym derivative $d\nu/d\mu$ is nondecreasing.[5]

Lemma 2 (Better news, higher value)**.** *Let the beliefs $\mu, \nu \in \mathcal{P}_1(\Theta)$ be mutually absolutely continuous.*

(a) *$\mu^{x_2} \succeq_{\text{lr}} \mu^{x_1}$ whenever $x_2 > x_1$;*

(b) *$\bar{v}(\nu) \geq \bar{v}(\mu)$ whenever $\nu \succeq_{\text{lr}} \mu$.*

Part (a) follows from the logconcavity of the noise density g, which implies the monotone likelihood ratio property of $f(x|\theta) = g(x - \theta)$. A higher offer is better news about the state in the sense of the likelihood ratio order. For part (b), we employ

[4]Since no signal is perfectly revealing, beliefs at any two dates reached in the same search are mutually absolutely continuous.

[5]We can generalize the likelihood ratio ordering to general beliefs: Let μ, ν have densities ϕ_μ, ϕ_ν with respect to the measure $\lambda = \mu + \nu$. We write $\nu \succeq_{\text{lr}} \mu$ if $\phi_\nu(\theta_2)\phi_\mu(\theta_1) \geq \phi_\nu(\theta_1)\phi_\mu(\theta_2)$ for $\lambda \otimes \lambda$-almost every pair $\theta_2 \geq \theta_1$.

the recursion logic of dynamic programming to prove the monotonicity of $\bar{v}$ in the likelihood ratio ordering.

For this search with learning problem, there are two decisions Sam needs to make in each period: (a) to quit or to search, and (b) to accept or to reject. While the second decision is in general very complex, the first one is relatively easy to characterize. For the rest of this section, we characterize the optimal quit behavior.

Note that for a classical search problem without learning, the search value is constant over time as the offer distribution is known and there is no information gain by observing a new offer. So, the quit behavior is simple. When the constant search value is negative, the searcher quits at the very beginning. Otherwise, he keeps searching until he accepts an offer that is better than the reservation offer.

For search with learning, we denote its *search value* by

$$\bar{v}^s(\mu) := -c + \int_{\mathbb{R}} \max\{x, \beta \bar{v}(\mu^x)\} f_\mu(x)\, dx, \tag{7}$$

which depends on the belief and is time-variant in general. So, unlike the full-information benchmark, it is possible for Sam to quit in the middle of the search after seeing too many unsatisfactory offers. Having the option to quit or to search, the optimal value can thus be written as $\bar{v}(\mu) = \max\{0, \bar{v}^s(\mu)\}$.

Lemma 3 (Quitting and search value)**.** *Let $\mu \in \mathcal{P}_1(\Theta)$. Then:*

(a) Every negative offer is rejected, i.e., $A(\mu) \subseteq [0, \infty)$;

(b) The search value $\bar{v}^s$ solves the Bellman equation of a searcher without *a quitting option who observes the original offer x but receives $x^+ = \max\{0, x\}$ upon acceptance:*

$$\bar{v}^s(\mu) = -c + \int_{\mathbb{R}} \max\big\{x^+, \beta \bar{v}^s(\mu^x)\big\} f_\mu(x)\, dx; \tag{8}$$

(c) The search value $\bar{v}^s$ is nondecreasing in the likelihood-ratio order;

(d) The set of beliefs $\{\mu \in \mathcal{P}_1(\Theta) : \ \bar{v}_{\beta,c}(\mu) = 0\}$ at which quitting is optimal is a lower set in the likelihood ratio order. The set is empty if $c = 0$.

(e) Fix $\beta \in (0, 1)$. There exists a unique threshold $\hat{c} \in (0, \infty)$ such that Sam searches iff $c < \hat{c}$.

(f) Fix $c \geq 0$. There exists a unique threshold $\hat{\beta} \in [0, 1]$ such that Sam searches iff $\beta > \hat{\beta}$.

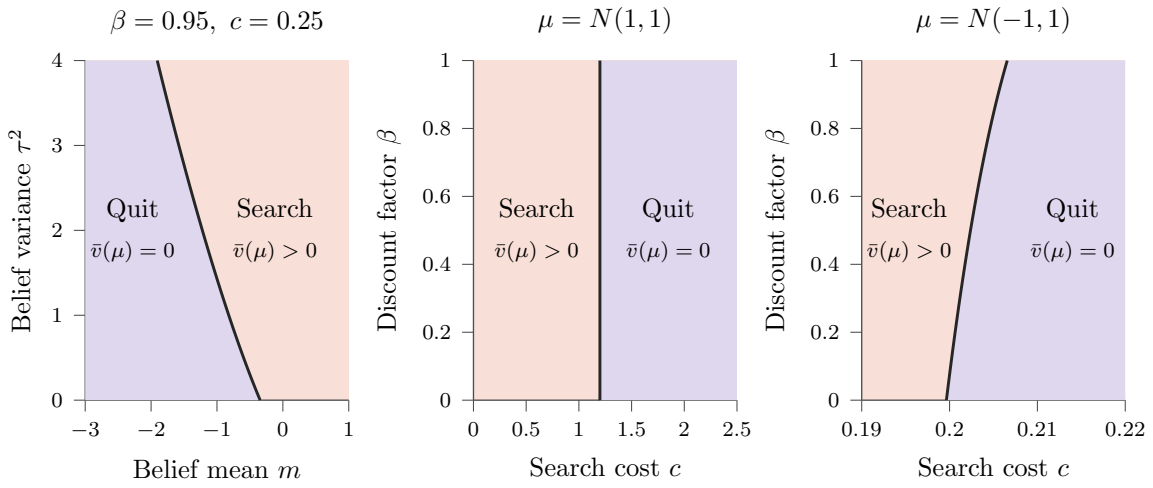


Figure 1: **Quitting Region with Gaussian Beliefs.** The noise is $\varepsilon_t \sim N(0,1)$. *Left panel:* the belief is $\mu = N(m, \tau^2)$, $\beta = 0.95$, and $c = 0.25$. Note that the Gaussian distribution is its own conjugate prior, so any resulting posterior distribution is also Gaussian. *Middle and right panels:* the beliefs are $\mu = N(1,1)$ and $\mu = N(-1,1)$, respectively. The lavender region is the set of beliefs at which quitting is optimal; search is strictly optimal in the coral region. The boundary is given by $\bar{v}^s(\mu) = 0$.

Proof. (a) Note that $\bar{v}(\mu) = \max\{0, \bar{v}^s(\mu)\} \geq 0$. If $x < 0$ then $x < 0 \leq \beta\bar{v}(\mu^x)$, so $x \notin A(\mu)$. (b) Since $\beta\bar{v}(\mu^x) \geq 0$ and $\bar{v}(\mu^x) = \max\{0, \bar{v}^s(\mu^x)\}$, we have

$$\max\{x, \beta\bar{v}(\mu^x)\} = \max\{x^+, \beta\bar{v}(\mu^x)\} = \max\{x^+, 0, \beta\bar{v}^s(\mu^x)\} = \max\{x^+, \beta\bar{v}^s(\mu^x)\}.$$

(c) Fix $\mu, \nu \in \mathcal{P}_1(\Theta)$ with $\nu \succeq_{\text{lr}} \mu$. By Lemma 2 and its proof, we know $\bar{v}(\nu^x) \geq \bar{v}(\mu^x)$, the predictive density f_ν first-order stochastically dominates f_μ, and the map $x \mapsto \max\{x, \beta\bar{v}(\mu^x)\}$ is nondecreasing. Thus, by (8),

$$\bar{v}^s(\nu) \geq -c + \int_{\mathbb{R}} \max\{x, \beta\bar{v}(\mu^x)\} f_\nu(x)\, dx \geq \bar{v}^s(\mu),$$

as desired. (d) By part (c), if $\bar{v}^s(\nu) \leq 0$ and $\nu \succeq_{\text{lr}} \mu$, then $\bar{v}^s(\mu) \leq \bar{v}^s(\nu) \leq 0$, which shows the claim on lower set. Finally, if $c = 0$, then (8) implies $\bar{v}^s(\mu) \geq \int_{\mathbb{R}} x^+ f_\mu(x)\, dx > 0$ since $f_\mu > 0$. (e) By part (a) of Proposition 1, the map $c \mapsto \bar{v}_{\beta,c}(\mu)$ is nonincreasing and convex, hence continuous on $(0, \infty)$. Let $m_+ := \int_{\mathbb{R}} x^+ f_\mu(x)\, dx \in (0, \infty)$. Since $\max\{0, m_+ - c\} \leq \bar{v}_{\beta,c}(\mu) \leq \max\{0, m_+/(1-\beta) - c\}$, monotonicity and continuity imply the existence of the unique threshold $\hat{c} \in [m_+, m_+/(1-\beta)]$. (f) By part (b) of Proposition 1, the map $\beta \mapsto \bar{v}_{\beta,c}(\mu)$ is nondecreasing, so its strictly positive set is an upper interval. □

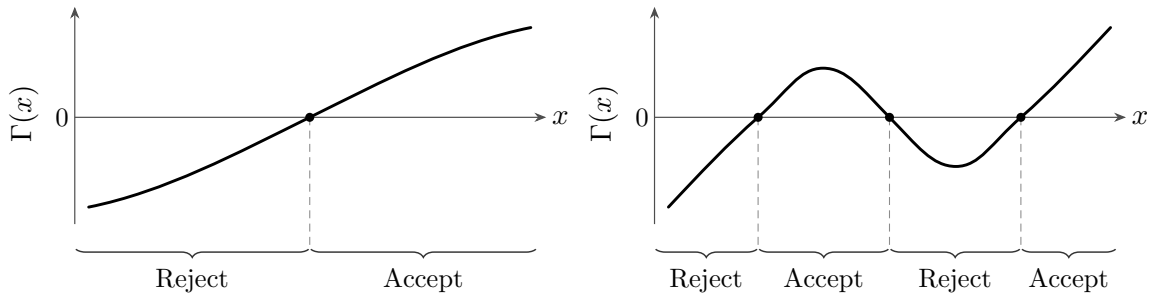


Figure 2: **Net Acceptance Gain and Reservation Policy.** The braces below each plot indicate the optimal action. In the left panel, an increasing net acceptance gain yields an upper interval of acceptable offers, thereby an optimal reservation policy. In the right panel, three zero crossings yield alternating rejection and acceptance regions.

In short, the optimal decision to quit employs a belief cutoff in the belief space in the sense of likelihood ratio order: to quit when sufficiently pessimistic about the state of the world. In the next section, we talk about the optimal decision to accept.

4 Analysis

Facing an offer x, Sam gains the linear payoff x itself if he accepts, and as he updates his belief from μ to μ^x, the expected payoff if he rejects is $\beta\bar{v}(\mu^x)$. We say a *reservation policy* is optimal at μ if the set of acceptable offers $A(\mu) := \{x \in \mathbb{R} : \ x \geq \beta\bar{v}(\mu^x)\}$ is an upper interval: whatever is accepted, anything better is accepted too. *Reservation failure* at μ means that some $x_1 < x_2$ have $x_1 \in A(\mu)$ and $x_2 \notin A(\mu)$—a better offer is rejected while a worse one is taken. This section builds the tool that decides between the two.

With prior belief μ, Sam's net gain of accepting x is $\Gamma(x) := x - \beta\bar{v}(\mu^x)$. So, Sam optimally accepts x iff $\Gamma(x) \geq 0$. If Γ is an increasing function, Sam should reject every small x and accept all the large ones, thereby forming a reservation policy (see the left panel of Figure 2). If it crosses zero three times, for example, as shown in the right panel of Figure 2, then it is optimal to reject a low prize, accept a medium prize, reject a higher one, and accept the highest. Below, we introduce a statistical concept and establish a criterion for the optimality of a reservation policy using the monotonicity of Γ.

The (Stein) score function is the logarithmic derivative of $f(x \mid \theta)$ with respect to

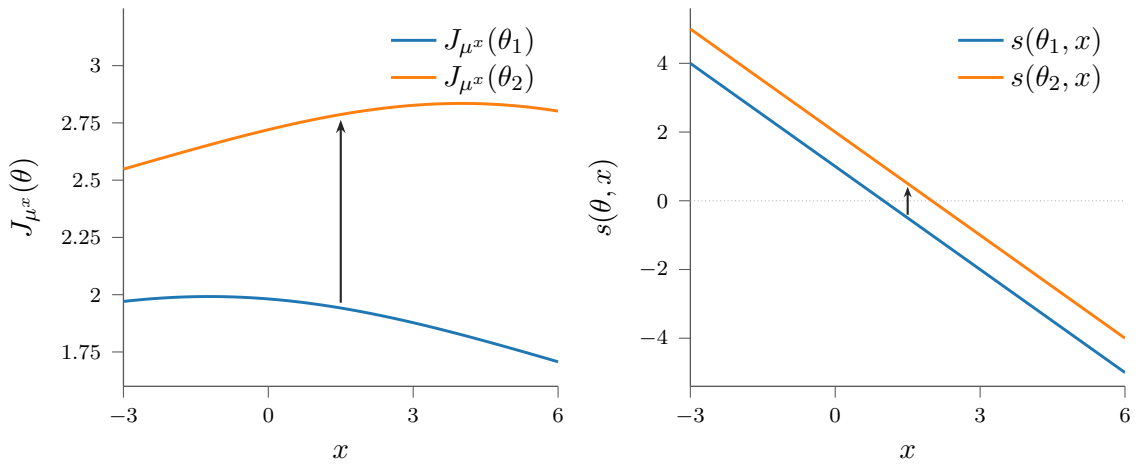


Figure 3: **Optimal Payoffs and the Score.** Fix $\varepsilon_t \sim N(0,1)$, $\mu = N(3/2, 1/4)$, $\theta_1 = 1$, $\theta_2 = 2$, $\beta = 0.95$, and $c = 0$. The left panel plots the numerically computed payoffs $J_{\mu^x}(\theta_1)$ and $J_{\mu^x}(\theta_2)$ under the policy $\pi^*(\mu^x)$. The right panel plots the exact scores $s(\theta_1, x) = 1 - x$ and $s(\theta_2, x) = 2 - x$. Each arrow indicates the change from θ_1 to θ_2.

x, namely,

$$s(\theta, x) := \frac{\partial}{\partial x} \log f(x \mid \theta) = (\log g)'(x - \theta) = \frac{g'(x-\theta)}{g(x-\theta)}. \tag{9}$$

It measures the proportional change in the likelihood per unit increase in the signal. By the log-supermodularity of $f(x \mid \theta)$ (or equivalently, by the logconcavity of g), the score $s(\theta, x)$ is nondecreasing in θ for every x.

Lemma 4 (The payoff-score covariance criterion)**.** *Fix $\mu \in \mathcal{P}_1(\Theta)$.*

(a) If $\beta \operatorname{Cov}_{\mu^x}(J_{\mu^x}(\theta), s(\theta, x)) \leq 1$ for almost every x, then the set of acceptable offers $A(\mu)$ is an upper interval: a reservation policy is optimal at μ;

(b) If x is an indifferent offer, i.e., $x = \beta\bar{v}(\mu^x)$, and $\beta \operatorname{Cov}_{\mu^x}(J_{\mu^x}(\theta), s(\theta, x)) > 1$, then $x - k$ is strictly accepted and $x + k$ is strictly rejected for all $k > 0$ sufficiently small: a reservation policy fails at μ.

In the appendix we show that, roughly speaking, the derivative of the net acceptance gain $\Gamma(x) = x - \beta\bar{v}(\mu^x)$ is $\Gamma'(x) = 1 - \beta \operatorname{Cov}_{\mu^x}(J_{\mu^x}(\theta), s(\theta, x))$. So, Γ is monotone if

$$\beta \operatorname{Cov}_{\mu^x}(J_{\mu^x}(\theta), s(\theta, x)) \leq 1.$$

We may think of the right-hand side as the payoff effect and the left-hand side as information effect. The failure of the reservation policy — accepting a medium offer but

rejecting a higher one — occurs only if the information effect exceeds the payoff effect. While it is clear that the acceptance gain rises linearly in the offer, the rejection value increment depends on the rise in posterior beliefs on the high state. A higher offer may shift beliefs upward so substantially that the rejection gain in Bellman values swamps the linear gain in acceptance value. The constant on the right-hand side represents the linear gain of payoff, and $\beta \operatorname{Cov}_{\mu^x}(J_{\mu^x}(\theta), s(\theta, x))$ captures the sensitivity of the change in the optimal value. The covariance is large when a higher offer shifts probability toward states in which continued search pays especially well. At an indifferent offer, an information effect greater than one makes nearby lower offers strictly acceptable and nearby higher offers strictly rejectable.

The nonnegativity of the information effect $\beta \operatorname{Cov}_{\mu^x}(J_{\mu^x}(\theta), s(\theta, x)) \geq 0$ follows from the lemma below.

Lemma 5 (Optimal payoffs covary positively with increasing statistics)**.** *Let $\mu \in \mathcal{P}_1(\Theta)$ and φ be a Borel function that is nondecreasing μ-almost everywhere.*[6] *Then*

$$\operatorname{Cov}_\mu\big(J_\mu(\theta), \varphi(\theta)\big) \geq 0 \tag{10}$$

whenever the covariance exists.

5 Optimality of a Reservation Policy

This section considers a prior belief μ that admits a density p supported on an interval with end points $-\infty \leq \underline{\theta} < \overline{\theta} \leq \infty$. For such a belief, define the *boundary penalty* by

$$B_\mu(J_\mu) := p(\underline{\theta})\big[J_\mu(\underline{\theta}) - \bar{v}(\mu)\big] + p(\overline{\theta})\big[\bar{v}(\mu) - J_\mu(\overline{\theta})\big], \tag{11}$$

with an endpoint term omitted when that endpoint is infinite. At a finite endpoint, the density value denotes its one-sided limit from inside the support.

[6] We say φ is nondecreasing μ-almost everywhere if $\varphi(\theta_1) \leq \varphi(\theta_2)$ for $\mu \otimes \mu$-almost every pair $\theta_1 \leq \theta_2$. If μ has a density, then it is enough that φ be nondecreasing Lebesgue almost everywhere.

Lemma 6 (The payoff-score covariance decomposition)**.** *Fix $x \in \mathbb{R}$, and let $\mu \in \mathcal{P}_1(\Theta)$ admit a density p supported on an interval with endpoints $-\infty \leq \underline{\theta} < \overline{\theta} \leq \infty$. Suppose p is strictly positive and locally absolutely continuous on $(\underline{\theta}, \overline{\theta})$ and has finite total variation.*[7] *Then*

$$\operatorname{Cov}_{\mu^x}\big(J_{\mu^x}(\theta), s(\theta, x)\big) = \mathbb{E}_{\mu^x}[J'_{\mu^x}(\theta)] + \operatorname{Cov}_{\mu^x}\big(J_{\mu^x}(\theta), (\log p)'(\theta)\big) + B_{\mu^x}(J_{\mu^x}), \quad (12)$$

with $0 \leq \mathbb{E}_{\mu^x}[J'_{\mu^x}(\theta)] \leq 1$, $B_{\mu^x}(J_{\mu^x}) \leq 0$, and $\operatorname{Cov}_{\mu^x}\big(J_{\mu^x}(\theta), s(\theta, x)\big) \geq 0$.

The payoff-score covariance $\operatorname{Cov}_{\mu^x}\big(J_{\mu^x}(\theta), s(\theta, x)\big)$ can be decomposed into three terms: the *expected marginal payoff* $\mathbb{E}_{\mu^x}[J'_{\mu^x}(\theta)]$ from the state, a *prior-shape* term $\operatorname{Cov}_{\mu^x}\big(J_{\mu^x}(\theta), (\log p)'(\theta)\big)$, and the *boundary penalty* $B_{\mu^x}(J_{\mu^x})$. The expected marginal payoff from the state is at most one: raising the state raises each offer by the same amount in expectation, but Sam collects at most one of them, with the accepted offer discounted if acceptance occurs later. And the boundary penalty is nonpositive. Thus only the prior-shape term can push the covariance above one. Under logconcavity, $(\log p)'$ is nonincreasing, so Lemma 5 makes this remaining term nonpositive as well.

Proposition 2 (Optimality of reservation policies for logconcave beliefs)**.** *Let μ be a prior belief on the state space Θ that admits a logconcave density. Then a reservation policy is optimal at μ.*

Proof. Let p be the density of μ. Then since p is a logconcave density, it is strictly positive and locally absolutely continuous on a (possibly unbounded) interval and has finite first absolute moment and total variation.

Since $(\log p)'$ is nonincreasing almost everywhere, by Lemma 5,

$$\operatorname{Cov}_{\mu^x}\big(J_{\mu^x}(\theta), (\log p)'(\theta)\big) = -\operatorname{Cov}_{\mu^x}\big(J_{\mu^x}(\theta), -(\log p)'(\theta)\big) \leq 0$$

for all $x \in \mathbb{R}$. It follows from Lemma 6 that $\operatorname{Cov}_{\mu^x}\big(J_{\mu^x}(\theta), s(\theta, x)\big) \leq 1$. So, by Lemma 4, a reservation policy is optimal at μ. □

Corollary 1 (Reservation policies after every history)**.** *Belief updating preserves logconcavity. As a result, if the belief density is logconcave at some date, then reservation policies are optimal thenceforth after every finite history.*

Proof. Let p be a logconcave density and fix any $x \in \mathbb{R}$. By Bayes' rule (1),

$$\log p^x(\theta) = \log p(\theta) + \log g(x - \theta) - \log f_\mu(x).$$

[7] The density p has finite total variation if $\int_{\underline{\theta}}^{\overline{\theta}} |p'(\theta)|\, d\theta < \infty$.

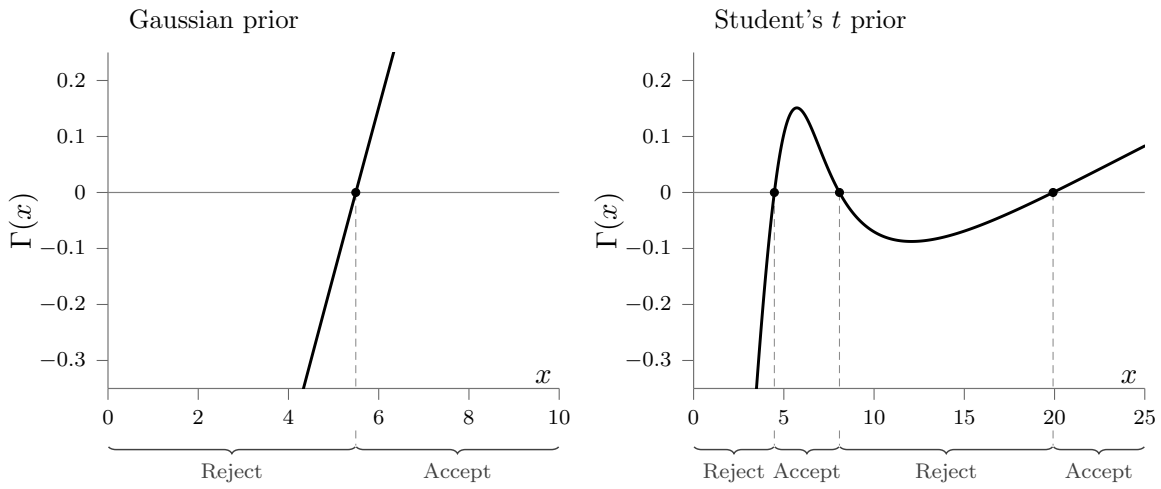


Figure 4: **Numerical Net Acceptance Gain and Optimal Policy.** Both panels set $\beta = 0.99$, $c = 0.025$, and $\varepsilon_t \sim N(0,1)$. The Gaussian prior is $N(2,3)$, and the Student's t prior $2 + t_3$ also has mean two and variance three.

The first term is concave by assumption, and the second is concave in θ because $\log g$ is concave. The last term is independent of θ, so $\log p^x$ is concave.[8] Starting at any date with a logconcave belief density, induction therefore gives a logconcave belief after every finite subsequent history. Applying Proposition 2 yields the desired result. □

The set of logconcave beliefs is absorbing under Bayesian updating: once the belief becomes logconcave, it remains so thereafter. As a result, having a logconcave belief means more than a one-shot optimality of a reservation policy only for the upcoming period, but the perpetual optimality of the standard cutoff rule for all the periods to come.

6 Failure of a Reservation Policy

Throughout this section, let the state space be $\Theta = (\underline{\theta}, \infty)$ with $\underline{\theta} \geq -\infty$, and assume the belief $\mu \in \mathcal{P}_1(\Theta)$ has a density p that is strictly positive and C^3 and has finite total variation. These conditions are widely satisfied, and the flexible state space setting is compatible with such distributions with different supports as the lognormal, Pareto, and Student's t. And we assume the noise density g has a bounded log-curvature,

[8]The logconcavity of the posterior density also follows from the proof of Theorem 7 in Prékopa (1973).

meaning that $|(\log g)''| \leq \Lambda$ for some $\Lambda > 0$. Every Gaussian density meets this requirement with $\Lambda = 1/\sigma^2$.

Now, we introduce the key conditions of the prior density which lead to reservation policy failure. We say p is decreasingly logconvex at a point θ if

$$(\log p)'''(\theta) \leq 0 \leq (\log p)''(\theta) \tag{13}$$

and p is eventually decreasingly logconvex if (13) holds for all θ sufficiently large.[9] Next, a belief μ is said to have a heavy tail if $\mathbb{E}_\mu[e^{t\theta}] = \infty$ for every $t > 0$.[10] These two conditions move away from logconcavity in two aspects: curvature and tail heaviness (as a logconcave function has thin tails). A wide range of distributions are covered, including lognormal, Pareto, and Student's t.

The importance of the tail of a distribution is not to be underestimated. For instance, Weitzman (2007) illustrates that the equity premium goes to infinity when agents price assets using fat-tailed distribution, and Morris and Yildiz (2019) show that a large shock under fat tails causes beliefs to shift dramatically, prompting sudden behavioral tipping points.

Proposition 3 (Reservation policy failure)**.** *Suppose p is eventually decreasingly logconvex and has a heavy tail. There exist a cost interval $[\underline{c}, \overline{c}]$ (with $0 < \underline{c} < \overline{c} < \infty$) and a discount factor threshold $\bar{\beta} \in (0,1)$ such that for every $c \in [\underline{c}, \overline{c}]$ and every $\beta \in (\bar{\beta}, 1)$, a reservation policy fails to be optimal.*

Proof sketch. Let $\beta \uparrow 1$. We show that there exists an offer x large enough such that $\beta \operatorname{Cov}_{\mu^x}\big(J_{\mu^x}(\theta), s(\theta, x)\big) > 1$. We then show that there exists an interval of search costs with which the offer x is an indifferent offer, i.e., $x = \beta \bar{v}(\mu^x)$.

Since p is eventually decreasingly logconvex and has a heavy tail, when $x \to \infty$, the belief has the "asymptotically fiducial" behavior studied by Dawid (1973): The prior becomes locally flat and the posterior distribution of $x - \theta$ approaches the noise distribution, as if the prior is uninformative.

Recall from Lemma 6 that

$$\operatorname{Cov}_{\mu^x}\big(J_{\mu^x}(\theta), s(\theta, x)\big) = \mathbb{E}_{\mu^x}[J'_{\mu^x}(\theta)] + \operatorname{Cov}_{\mu^x}\big(J_{\mu^x}(\theta), (\log p)'(\theta)\big) + B_{\mu^x}(J_{\mu^x}).$$

First, the expected marginal payoff $\mathbb{E}_{\mu^x}[J'_{\mu^x}(\theta)]$ from the state is close to one. Because offers satisfy $x_t = \theta + \varepsilon_t$, raising θ by a small amount raises every potential offer

[9] Note that a probability density that is positive on $\mathbb{R}$ cannot be globally logconvex, as a globally logconvex positive function cannot decay to zero in both tails.

[10] The set of heavy-tailed probability measures with eventually decreasingly logconvex densities is a dense set in $\mathcal{P}_1(\Theta)$.

by that same amount. As x grows large, the posterior belief μ^x becomes optimistic, so quitting becomes unlikely. As Sam will eventually accept an offer, the expected marginal payoff becomes the extra gain from the expected discounted offer that will be accepted when every potential offer raises by one. That goes to one as Sam is sufficiently patient.

Next, the boundary penalty $B_{\mu^x}(J_{\mu^x})$ is close to zero. Since the upper end point is $\overline{\theta} = \infty$, here we have only one-sided boundary $B_{\mu^x}(J_{\mu^x}) = p^x(\underline{\theta})\big[J_{\mu^x}(\underline{\theta}) - \bar{v}(\mu^x)\big]$. As the offer x becomes large, the posterior density $p^x(\underline{\theta}) = p(\underline{\theta})g(x-\underline{\theta})/f_\mu(x)$ at the boundary $\underline{\theta}$ becomes negligible, so the entire boundary term vanishes.

Last, the prior-shape term $\mathrm{Cov}_{\mu^x}\big(J_{\mu^x}(\theta), (\log p)'(\theta)\big)$ is positive. Logconvexity means that $(\log p)'$ is nondecreasing. If p is globally logconvex and has a heavy tail, then the prior-shape term is positive, and basically we are done. But p is only eventually logconvex. Nonetheless, when x is large enough, the posterior μ^x places almost all its weight in this logconvex tail, contributing positively, while the remaining states may contribute negatively. Decreasing log-curvature lets the proof control an interval of plausible states using the curvature at its upper end, making the positive effect persist across a useful range of states. Combined with heavy-tailedness, we show this to be strong enough to dominate the negative contribution. □

The result in Proposition 3 concerns the optimal acceptance rule conditional on searching and observing an offer. One simple sufficient condition under which it is optimal to search for every $c \in [\underline{c}, \overline{c}]$ and $\beta \in (\bar{\beta}, 1)$ is that $\mathbb{E}_\mu\theta + \mathbb{E}\varepsilon_0 \geq 0$. Thus, with mean-zero noise, a nonnegative prior mean is sufficient.

To provide a sense of how large the search cost needs to be for a reservation policy to fail, let ε_0 and ε_1 be iid with density g, and define the Gini mean difference $D_g := \mathbb{E}|\varepsilon_1 - \varepsilon_0| > 0$, $\tilde{c}_g := \frac{D_g}{12}$, and $\tilde{\tilde{c}}_g := \mathbb{E}\big[(\varepsilon_0 - \mathbb{E}\varepsilon_0 + D_g)^+\big]$. In the proof, we show that $\underline{c} \geq \tilde{c}_g$ and $\overline{c} \leq \tilde{\tilde{c}}_g$. For example, when we have Gaussian noise, the constants are $D_g = 2\sigma/\sqrt{\pi}$, $\tilde{c}_g = \sigma/(6\sqrt{\pi})$, and $\tilde{\tilde{c}}_g = \sigma\left[\phi(2/\sqrt{\pi}) + (2/\sqrt{\pi})\,\Phi(2/\sqrt{\pi})\right]$. For standard normal noise ($\sigma = 1$), we have $D_g \approx 1.128$, $\tilde{c}_g \approx 0.094$, and $\tilde{\tilde{c}}_g \approx 1.193$.

As we have seen, a reservation policy is not in general optimal for heavy-tailed priors. Nevertheless, the failure need not persist as the belief evolves. We will show that under two weak assumptions, if Sam searches long enough, a standard reservation policy is again optimal.

The first is a condition on the noise density. We say g is strongly logconcave if there exists $\eta > 0$ such that $\eta \leq -(\log g)''$. Note that when η is zero, that is simply the definition for logconcavity. Every Gaussian density is strongly logconcave, with $\eta = 1/\sigma^2$, so it is compatible with our canonical example of the noise.

The second assumption requires that the prior density p have a log-curvature that

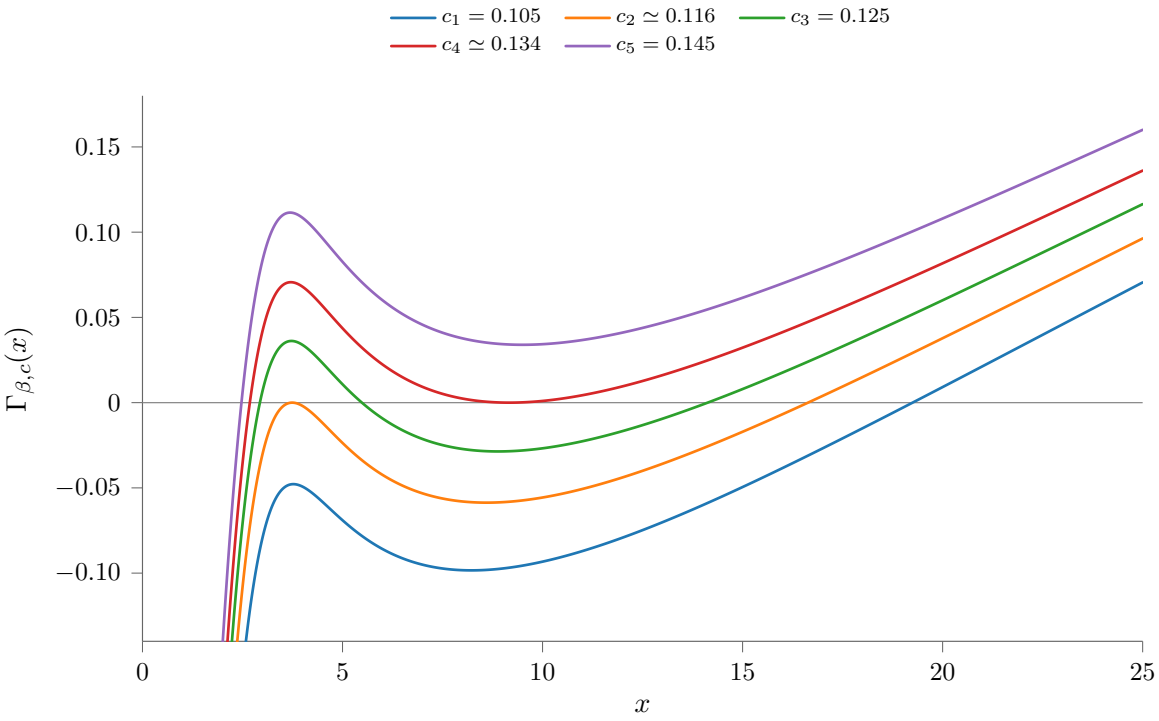


Figure 5: **Search Costs and Reservation Policy Failure.** Fix $\varepsilon_t \sim N(0,1)$, $\log \theta \sim N(0.5,1)$, and $\beta = 0.993$. This figure displays the net gain of each offer x. The five curves are ordered from bottom to top as $c_1 < c_2 < c_3 < c_4 < c_5$. The c_2 curve touches zero at a local maximum, and the c_4 curve touches zero at a local minimum. The middle curve has three zero crossings, indicating the violation of the reservation property.

is bounded above, i.e., $(\log p)'' \leq K$ for some $K > 0$.[11] This condition is met by many in the class of heavy-tailed distributions we consider in Proposition 3, such as the lognormal, Pareto, and Student's t distributions.

Proposition 4 (Reservation policies after repeated learning)**.** *Let g be strongly logconcave. Then belief updating fosters logconcavity. As a result, if the belief density p has a bounded above log-curvature at some date, it will be logconcave after a finite horizon. Reservation policies are optimal thenceforth.*

Proof. By assumption, $(\log p)'' \leq K$ and $-(\log g)'' \geq \eta > 0$. Relabel the date in question as date 0, and let p be the density of μ_0. Applying Bayes' rule (1) and taking derivatives, we obtain the second derivative of the posterior log-density after observing

[11] Such densities are also called semi-logconcave in some statistics and computer science literature.

Period-0 prior		Gaussian noise		
Family	Parameters	$\sigma = 1$	$\sigma = 2$	$\sigma = 4$
Lognormal	$m = 0.5,\ s = 1$	1	1	2
Lognormal	$m = 0.5,\ s = 2$	7	28	110
Lognormal	$m = 0,\ s = 2$	19	75	297
Pareto	$\alpha = 2,\ a = 2$	1	3	12
Pareto	$\alpha = 2,\ a = 0.5$	12	48	192
Pareto	$\alpha = 2,\ a = 0.25$	48	192	768
Student's t	$\nu = 3,\ s = 0.5$	1	3	11
Student's t	$\nu = 3,\ s = 0.1$	17	67	267
Student's t	$\nu = 3,\ s = 0.05$	67	267	1067

Table 1: **Periods until posterior logconcavity.** Noise is Gaussian, $\varepsilon_t \sim N(0, \sigma^2)$.

any offer history $h_t = (x_0, \ldots, x_{t-1})$:

$$(\log p)''(\theta) + \sum_{j=0}^{t-1} (\log g)''(x_j - \theta) \leq K - t\eta,$$

which is nonpositive for every $t \geq \lceil K/\eta \rceil$, implying that the posterior density is logconcave. Therefore Proposition 2 implies that a reservation policy is optimal at each of these logconcave posterior beliefs. □

Here we provide some intuition on why belief updating fosters logconcavity. Suppose Sam has seen several offers. An arbitrarily high θ would require explaining all those offers as unusually bad luck. With Gaussian noise—or, more generally, strongly logconcave noise—such explanations become increasingly costly in likelihood terms: Explaining the offer with a progressively higher state requires progressively worse luck, with an increasing marginal likelihood penalty. Thus, enough observations make the belief to have diminishing proportional change in the state, which means the belief is logconcave.

7 Discrete Beliefs

This section discusses the failure of a reservation policy for discrete beliefs, which are special cases of this paper's model. In particular, consider $\Theta = \{\theta_L, \theta_H\}$ where $0 \leq \theta_L < \theta_H$. The results in this section can be extended easily to a finite number of

states, as illustrated in the appendix.

Let δ_θ be the Dirac measure on state θ and

$$\mu_q := (1-q)\delta_{\theta_L} + q\delta_{\theta_H} \tag{14}$$

be the two-state belief that puts probability $q \in (0,1)$ on the high state. This belief has a finite first absolute moment, so relevant results in Sections 3 and 4 apply.

For any prior probability q, write

$$q^x := \mu_q^x(\{\theta_H\}) = \frac{qf(x|\theta_H)}{qf(x|\theta_H) + (1-q)f(x|\theta_L)} \tag{15}$$

for the posterior probability of the high state, so that $\mu_q^x = \mu_{q^x}$.

Proposition 5 (Reservation Failure on a Half-Plane)**.** *Fix $\beta \in (0,1)$ and $c \geq 0$, and let g be strongly logconcave. There exists $\Delta_0 > 0$ such that whenever the dispersion $\theta_H - \theta_L \geq \Delta_0$, a reservation policy fails to be optimal under some prior belief $q \in (0,1)$ on the high state.*

Proof sketch. In this two-state setting, we can obtain a simple form for the payoff-score covariance by expanding it under the two-point belief μ_q^x, namely, for any prior probability $q \in (0,1)$ and offer $x \in \mathbb{R}$,

$$\operatorname{Cov}_{\mu_q^x}(J_{\mu_q^x}(\theta), s(\theta,x)) = q^x(1-q^x)\big[J_{\mu_q^x}(\theta_H) - J_{\mu_q^x}(\theta_L)\big]\big[s(\theta_H,x) - s(\theta_L,x)\big]. \tag{16}$$

Since g is strongly logconcave, there exists $\eta > 0$ such that $\eta \leq -(\log g)''$. So, for every $x \in \mathbb{R}$,

$$s(\theta_H,x) - s(\theta_L,x) = \int_{x-\theta_H}^{x-\theta_L} -(\log g)''(\varepsilon)\, d\varepsilon \geq (\theta_H - \theta_L)\eta.$$

In the appendix, using (16), we show that under some prior belief, $\beta \operatorname{Cov}_{\mu_q^x}(J_{\mu_q^x}(\theta), s(\theta,x))$ can be arbitrarily large for some indifferent offer x as the dispersion $\theta_H - \theta_L$ grows. The desired result then follows from Lemma 4(b). □

8 Conclusion

This paper revisits the classic search problem by incorporating learning about the offer-generating process and characterizing optimal behavior in this environment. When the distribution of offers is known, optimal stopping takes a simple form: accept if and only if the offer exceeds a reservation level. Once the distribution itself is uncertain,

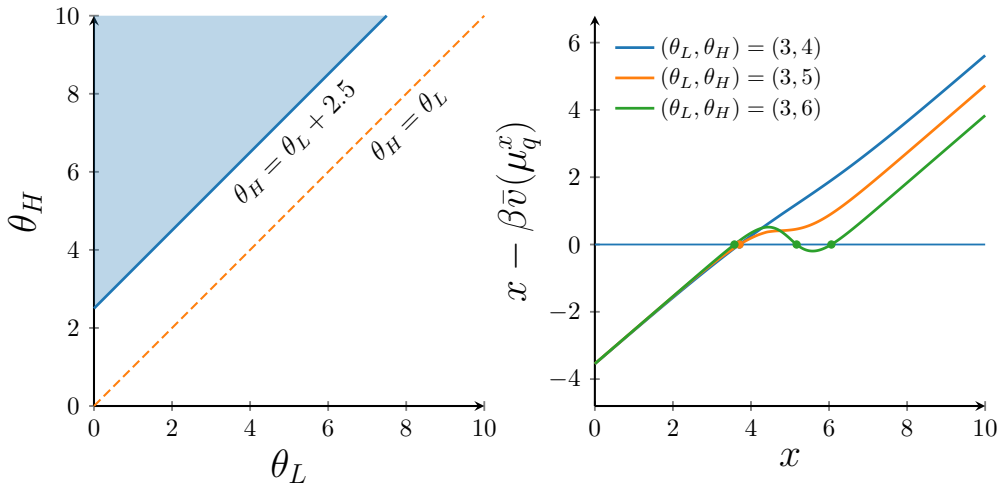


Figure 6: **Reservation Failure on a Half-Plane with Gaussian Noise.** Fix $\varepsilon_t \sim N(0,1)$, $\beta = 0.95$, and $c = 0$, as in Figure 7. *Left panel:* the shaded region is $\{(\theta_L, \theta_H) \in \mathbb{R}^2_+ : \theta_H - \theta_L \geq 2.5\}$. At each state pair in this region, a reservation policy fails for some prior $q \in (0,1)$. The solid boundary is included, and the dashed diagonal marks $\theta_H = \theta_L$. *Right panel:* the curves plot the numerically computed net acceptance gain $x - \beta\bar{v}(\mu_q^x)$ at prior $q = 0.2$ for $(\theta_L, \theta_H) = (3,4)$, $(3,5)$, and $(3,6)$. The horizontal line marks zero, and the dots mark zero crossings.

however, this logic breaks down. Each observed offer conveys information about the underlying environment, and the decision to accept or reject must trade off immediate payoff against the value of improved beliefs.

Logconcave belief density guarantees an optimal reservation policy for every search cost and discount factor, and Bayesian updating preserves this conclusion after every history. By contrast, for the class of heavy-tailed beliefs studied here, sufficiently patient searchers accept some medium offers but reject higher ones over an interval of search costs. Failure also arises with discrete beliefs. Repeated learning can restore the reservation policy: sufficiently many observations make the posterior logconcave, regardless of the realized history. A cutoff rule is then optimal at every subsequent history.

Our findings have implications for a wide range of economic settings in which agents face unfamiliar or evolving environments. Workers entering a new occupation and consumers purchasing an unfamiliar good provide natural settings in which learning can generate qualitatively new patterns of policy that are absent from standard search models.

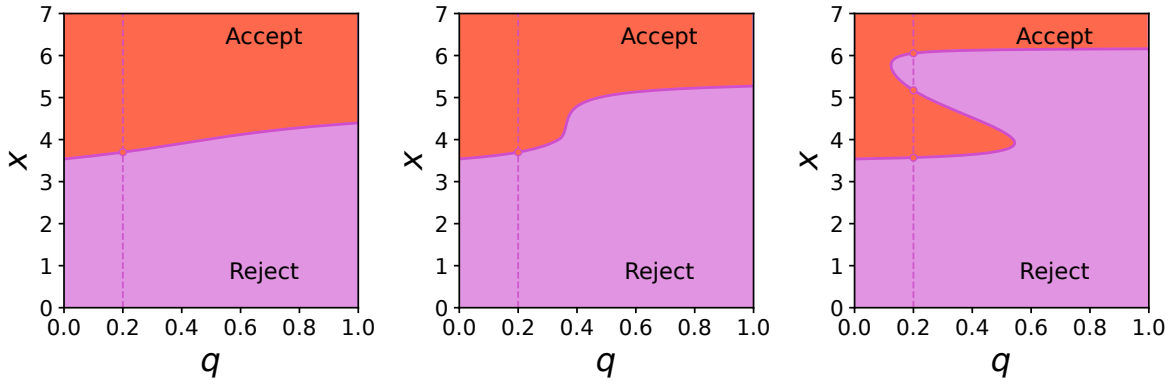


Figure 7: **Set of Acceptable Offers with Gaussian Noise:** $\theta_L = 3$. Fix $\varepsilon_t \sim N(0,1)$, $\beta = 0.95$, and $c = 0$. From left to right, $\theta_H = 4, 5, 6$. The coral region denotes acceptance, $x \geq \beta \bar{v}(\mu_{q^x})$, and the lavender region denotes rejection. The horizontal axis is the prior probability q of the high state, and the vertical axis is the offer x. The dashed line marks $q = 0.2$, and the dots mark indifferent offers. The right panel exhibits failure of the reservation policy.

A Bellman Contraction: Proofs for Section 3

This appendix specifies the topology and function spaces used in Proposition 1.

We equip $\mathcal{P}_1(\Theta)$, the set of Borel probability measures on Θ with finite first absolute moment, with the Wasserstein-1 distance

$$W_1(\mu, \nu) := \inf_{\xi \in \Xi(\mu,\nu)} \int_{\Theta^2} |\theta - \vartheta| \, \xi(d\theta, d\vartheta),$$

where $\Xi(\mu, \nu)$ is the set of joint probability measures with marginals μ and ν. Lemma 8.3(a) of Bickel and Freedman (1981) gives that convergence in W_1 is equivalent to weak convergence together with convergence of the first absolute moments: for $\mu_n, \mu \in \mathcal{P}_1(\Theta)$,

$$W_1(\mu_n, \mu) \to 0 \quad \Longleftrightarrow \quad \mu_n \Rightarrow \mu \quad \text{and} \quad \int_\Theta |\theta| \, \mu_n(d\theta) \to \int_\Theta |\theta| \, \mu(d\theta). \tag{17}$$

Define the weight

$$\omega(\mu) := \int_\Theta (1 + |\theta|) \, \mu(d\theta) = 1 + \mathbb{E}_\mu |\theta|, \tag{18}$$

the weighted supremum norm

$$\|v\|_\omega := \sup_{\mu \in \mathcal{P}_1(\Theta)} \frac{|v(\mu)|}{\omega(\mu)}, \tag{19}$$

and let

$$C_\omega(\Theta) := \{v\colon \mathcal{P}_1(\Theta) \to \mathbb{R} :\ v \text{ is } W_1\text{-continuous},\ \|v\|_\omega < \infty\},$$

which is a Banach space equipped with the norm $\|\cdot\|_\omega$ (Boyd, 1990, Section 3). The weight allows value functions to be unbounded across beliefs while controlling their growth by the first absolute moment.

We first verify that the operator $\mathcal{T}$ defined in (6) maps the Banach space $C_\omega(\Theta)$ into itself.

Lemma A.1. *(a) For every $\mu \in \mathcal{P}_1(\Theta)$,*

$$\int_{\mathbb{R}} \omega(\mu^x) f_\mu(x)\, dx = \omega(\mu), \qquad \int_{\mathbb{R}} |x| f_\mu(x)\, dx \le \mathbb{E}_\mu|\theta| + \mathbb{E}|\varepsilon_0| < \infty.$$

(b) The maps $(\mu, x) \mapsto f_\mu(x)$ and $(\mu, x) \mapsto \mu^x$ are jointly continuous on $\mathcal{P}_1(\Theta) \times \mathbb{R}$.

(c) For every $v \in C_\omega(\Theta)$, we have $\mathcal{T}v \in C_\omega(\Theta)$ and

$$\|\mathcal{T}v\|_\omega \le c + 1 + \mathbb{E}|\varepsilon_0| + \beta\|v\|_\omega.$$

(d) Every fixed point $v \in C_\omega(\Theta)$ of $\mathcal{T}$ equals the value $\bar{v}$ in (4), which can be attained by a policy at every initial belief.

Proof. (a) By Lemma 1, all posteriors lie in $\mathcal{P}_1(\Theta)$. Logconcavity of g gives $\mathbb{E}|\varepsilon_0| < \infty$. Bayes' rule and Tonelli's theorem yield

$$\int_{\mathbb{R}} \omega(\mu^x) f_\mu(x)\, dx = \int_\Theta (1 + |\theta|) \left[\int_{\mathbb{R}} g(x - \theta)\, dx\right] \mu(d\theta) = \omega(\mu).$$

The predictive moment bound follows from $|x| \le |\theta| + |\varepsilon_0|$.

(b) Suppose $W_1(\mu_n, \mu) \to 0$ and $x_n \to x$. As a consequence of (17), for every continuous $\varphi\colon \Theta \to \mathbb{R}$ with at most linear growth,

$$\int_\Theta \varphi(\theta) g(x_n - \theta)\, \mu_n(d\theta) \longrightarrow \int_\Theta \varphi(\theta) g(x - \theta)\, \mu(d\theta).$$

At fixed x, this follows because $\varphi(\theta)g(x-\theta)$ is continuous with at most linear growth. The change from x to x_n is controlled by uniform continuity of g and boundedness of $\int_\Theta |\varphi|\, d\mu_n$. Taking $\varphi = 1$ proves $f_{\mu_n}(x_n) \to f_\mu(x) > 0$. Bayes' rule and $f_{\mu_n}(x_n) \to f_\mu(x) > 0$ imply

$$\int_\Theta \varphi(\theta)\, \mu_n^{x_n}(d\theta) \longrightarrow \int_\Theta \varphi(\theta)\, \mu^x(d\theta)$$

for every continuous $\varphi\colon \Theta \to \mathbb{R}$ with at most linear growth. Applying this convergence to all bounded continuous φ gives $\mu_n^{x_n} \Rightarrow \mu^x$; applying it to $\varphi(\theta) = |\theta|$ gives convergence of the posterior first absolute moments. Thus (17) yields $W_1(\mu_n^{x_n}, \mu^x) \to 0$.

(c) For $v \in C_\omega(\Theta)$,

$$|\max\{x, \beta v(\mu^x)\}| \le |x| + \beta\|v\|_\omega \omega(\mu^x).$$

Integrating and using the moment bounds gives absolute integrability and the asserted norm bound.

To prove continuity, suppose $W_1(\mu_n, \mu) \to 0$. Joint continuity of updating gives pointwise convergence of the Bellman integrands. Moreover,

$$|\max\{x, \beta v(\mu_n^x)\}| f_{\mu_n}(x) \le \big(|x| + \beta\|v\|_\omega \omega(\mu_n^x)\big) f_{\mu_n}(x).$$

The nonnegative bounds on the right converge pointwise to the corresponding bound at μ. Their integrals are

$$\int_\Theta \mathbb{E}|\theta + \varepsilon_0|\, \mu_n(d\theta) + \beta\|v\|_\omega \omega(\mu_n),$$

which converge to the same expression at μ: the function $\theta \mapsto \mathbb{E}|\theta + \varepsilon_0|$ is one-Lipschitz with at most linear growth. Fatou's lemma, applied to each bound plus and minus its Bellman integrand, therefore gives convergence of the Bellman integrals. Taking the maximum with zero preserves continuity, so $\mathcal{T}v \in C_\omega(\Theta)$.

(d) Fix any policy π and any date $t \ge 1$. Consider following π through date $t-1$, but quitting at the beginning of date t if search has not already ended. Since $v = \mathcal{T}v$, the outer maximum in the Bellman operator implies $v \ge 0$. Thus, at the terminal date t, the payoff from mandatory quitting satisfies $0 \le v(\mu_t)$. By induction, at every period $j < t$, the payoff from searching is bounded by

$$v(\mu_j) \ge -c + \int_{\mathbb{R}} \max\{x, \beta v(\mu_j^x)\} f_{\mu_j}(x)\, dx.$$

Hence, the expected payoff of the truncated policy is at most $v(\mu)$. The original and truncated policies agree before date t, so their payoffs differ only through terms at dates $j \ge t$. The difference is at most

$$\mathbb{E}_\mu\left[\sum_{j=t}^{\infty} \beta^j (c + |x_j|)\right] \le \frac{\beta^t}{1-\beta}\big(c + \mathbb{E}_\mu|\theta| + \mathbb{E}|\varepsilon_0|\big).$$

The bound is finite and goes to zero as $t \to \infty$. It follows that $\int_\Theta J_\pi \, d\mu \leq v(\mu)$. Since π is arbitrary, the inequality holds for every policy. Taking the supremum over $\pi \in \Pi$ in (4) gives $\bar{v}(\mu) \leq v(\mu)$.

Conversely, fix $\mu_0 = \mu$ and use the policy π that searches exactly when $v(\mu_t) > 0$ and accepts exactly when $x_t \geq \beta v(\mu_{t+1})$. Then it is not difficult to see that the policy π earns $v(\mu)$, which implies $\int_\Theta J_\pi \, d\mu = v(\mu)$, and thus $\bar{v}(\mu) \geq v(\mu)$. Together with the preceding upper bound, this proves $v(\mu) = \bar{v}(\mu)$ and policy attainment. □

Next, we show $\mathcal{T}$ is a β-contraction, as well as the cost and discount factor monotonicities.

Proof of Proposition 1. Contraction. By Lemma A.1, $\mathcal{T}$ maps the complete space $C_\omega(\Theta)$ into itself. For $v_1, v_2 \in C_\omega(\Theta)$, we have

$$|(\mathcal{T}v_1)(\mu) - (\mathcal{T}v_2)(\mu)| \leq \beta \int_{\mathbb{R}} |v_1(\mu^x) - v_2(\mu^x)| f_\mu(x) \, dx.$$

The weighted norm and the moment identity in Lemma A.1 bound the right-hand side by $\beta \|v_1 - v_2\|_\omega \omega(\mu)$. Dividing by $\omega(\mu)$ and taking the supremum over beliefs gives

$$\|\mathcal{T}v_1 - \mathcal{T}v_2\|_\omega \leq \beta \|v_1 - v_2\|_\omega.$$

Banach's fixed-point theorem yields a unique fixed point in $C_\omega(\Theta)$. By Lemma A.1, it is $\bar{v}$ and is attained by a policy. Being in $C_\omega(\Theta)$ gives continuity, while (4) gives convexity in beliefs as a supremum of affine functions.[12]

Cost monotonicity. Each fixed-policy expected payoff is affine and nonincreasing in c, so their supremum $\bar{v}_c(\mu)$ is convex and nonincreasing. If $0 \leq c_1 < c_2$ and $\bar{v}_{c_2}(\mu) > 0$, searching attains the Bellman maximum at cost c_2. Thus the Bellman equations and $\bar{v}_{c_1} \geq \bar{v}_{c_2}$ imply

$$\bar{v}_{c_1}(\mu) \geq \bar{v}_{c_2}(\mu) + (c_2 - c_1) > \bar{v}_{c_2}(\mu).$$

Discount factor monotonicity. Fix $c \geq 0$ and $0 < \beta_1 < \beta_2 < 1$, and take $\mathcal{T}$ at discount β_1. Since $\bar{v}_{\beta_2} \geq 0$,

$$0 \leq \mathcal{T}v \leq \mathcal{T}\bar{v}_{\beta_2} \leq \bar{v}_{\beta_2} \qquad \text{whenever } 0 \leq v \leq \bar{v}_{\beta_2}.$$

Hence $\mathcal{T}$ preserves the nonempty closed order interval $\{v \in C_\omega(\Theta) : 0 \leq v \leq \bar{v}_{\beta_2}\}$, and closure follows because weighted-norm convergence implies pointwise convergence. Corollary 3.2.1 in Stokey and Lucas (1989) therefore gives $\bar{v}_{\beta_1} \leq \bar{v}_{\beta_2}$. □

[12] See Nyarko (1994) for more details on the convexity of value functions.

The following proves the monotonicity of $\bar{v}$ in the likelihood ratio ordering.

Proof of Lemma 2. (a) For $x_1 < x_2$, strict positivity of g implies that μ^{x_1} and μ^{x_2} are mutually absolutely continuous. Bayes' rule gives

$$\frac{d\mu^{x_2}}{d\mu^{x_1}}(\theta) = \frac{f_\mu(x_1)}{f_\mu(x_2)} \frac{g(x_2 - \theta)}{g(x_1 - \theta)}.$$

Logconcavity of g implies

$$\frac{d}{d\theta} \log \frac{g(x_2 - \theta)}{g(x_1 - \theta)} = (\log g)'(x_1 - \theta) - (\log g)'(x_2 - \theta) \geq 0.$$

The Radon–Nikodym derivative is therefore nondecreasing, proving $\mu^{x_2} \succeq_{\text{lr}} \mu^{x_1}$.

(b) Since the set of nonnegative functions in $C_\omega(\Theta)$ that are nondecreasing in the likelihood-ratio order is a nonempty closed subset, by Proposition 1 and Corollary 3.2.1 in Stokey and Lucas (1989), it suffices to show that $\mathcal{T}$ maps this set to itself.

Take such a function v and suppose $\nu \succeq_{\text{lr}} \mu$, which implies ν first-order stochastically dominates μ. It then follows that f_ν first-order stochastically dominates f_μ. Moreover, by Bayes' rule,

$$\frac{d\nu^x}{d\mu^x}(\theta) = \frac{f_\mu(x)}{f_\nu(x)} \frac{d\nu}{d\mu}(\theta).$$

So, it is clear that $\nu \succeq_{\text{lr}} \mu$ also implies $\nu^x \succeq_{\text{lr}} \mu^x$. Since v is assumed to be nondecreasing in the likelihood-ratio order, we have $v(\nu^x) \geq v(\mu^x)$. And by part (a), $x \mapsto \max\{x, \beta v(\mu^x)\}$ is nondecreasing. Consequently,

$$(\mathcal{T}v)(\nu) \geq \max\left\{0, -c + \int_{\mathbb{R}} \max\{x, \beta v(\mu^x)\} f_\nu(x)\, dx\right\} \geq (\mathcal{T}v)(\mu) \geq 0.$$

The required preservation follows. □

B The Covariance Criterion: Proofs for Section 4

This appendix first specifies the norm and derivative conventions. For a finite signed Borel measure ν on Θ, define the weighted total-variation norm

$$\|\nu\|_{1,\text{TV}} := \int_\Theta (1 + |\theta|)\, |\nu|(d\theta), \tag{20}$$

where $|\nu|$ denotes the total variation measure of ν. The signed measures for which this norm is finite form a Banach space containing $\mathcal{P}_1(\Theta)$, and $\|\mu\|_{1,\text{TV}} = \omega(\mu)$ for every

$\mu \in \mathcal{P}_1(\Theta)$.

For a belief path $x \mapsto \nu_x \in \mathcal{P}_1(\Theta)$, its derivative exists and equals $\dot{\nu}_x$ if as $t \to 0$,

$$\left\| \frac{\nu_{x+t} - \nu_x}{t} - \dot{\nu}_x \right\|_{1,\mathrm{TV}} \longrightarrow 0.$$

In particular, $\dot{\nu}_x(\Theta) = 0$ and $\|\dot{\nu}_x\|_{1,\mathrm{TV}} < \infty$. The one-sided derivatives in Lemma 8 are defined by

$$\frac{d^{\pm}}{dx}\bar{v}(\nu_x) := \lim_{t \to 0^{\pm}} \frac{\bar{v}(\nu_{x+t}) - \bar{v}(\nu_x)}{t},$$

whenever the corresponding limit exists.

Lemma 7. *For every $\mu \in \mathcal{P}_1(\Theta)$, the map $x \mapsto \mu^x$ is locally absolutely continuous and is differentiable almost everywhere with derivative*

$$\frac{\partial}{\partial x}\mu^x(d\theta) = \big(s(\theta, x) - \mathbb{E}_{\mu^x}[s(\theta, x)]\big)\mu^x(d\theta). \tag{21}$$

A marginal increase in the offer raises posterior weight on states with above-average score, and lowers posterior weight on states with below-average score: the posterior is tilted in the direction of the score.

Proof of Lemma 7. Fix $\mu \in \mathcal{P}_1(\Theta)$. Let m be a mode of g. Logconcavity makes g unimodal with exponentially bounded tails. Integration by parts on either side of m gives

$$\int_{\mathbb{R}} |g'(\varepsilon)|\, d\varepsilon = 2g(m), \qquad \int_{\mathbb{R}} |\varepsilon - m|\, |g'(\varepsilon)|\, d\varepsilon = 1.$$

Consequently, g and g' satisfy

$$\int_{\mathbb{R}} (1 + |\varepsilon|)\big[g(\varepsilon) + |g'(\varepsilon)|\big]\, d\varepsilon < \infty. \tag{22}$$

For every $a < b$, Tonelli's theorem yields

$$\int_a^b \int_{\Theta} (1 + |\theta|)|g'(t - \theta)|\, \mu(d\theta)\, dt \le \omega(\mu) \int_{\mathbb{R}} |g'(\varepsilon)|\, d\varepsilon < \infty.$$

The fundamental theorem of calculus therefore gives the identity

$$g(b - \cdot)\mu - g(a - \cdot)\mu = \int_a^b g'(t - \cdot)\mu\, dt,$$

where the right-hand side is a Bochner integral in weighted total variation. This integral

representation shows that $x \mapsto g(x-\cdot)\mu$ is locally absolutely continuous with derivative $g'(x-\cdot)\mu$ almost everywhere.

Taking total masses gives $f'_\mu(x) = \int_\Theta g'(x-\theta)\,\mu(d\theta)$ almost everywhere and shows that f_μ is locally absolutely continuous. Since f_μ is continuous and strictly positive, it is bounded away from zero on every compact interval. Bayes' rule thus makes $x \mapsto \mu^x$ locally absolutely continuous, and the quotient rule gives, almost everywhere,

$$\frac{\partial}{\partial x}\mu^x(d\theta) = \frac{g'(x-\theta)}{f_\mu(x)}\,\mu(d\theta) - \frac{f'_\mu(x)}{f_\mu(x)}\,\mu^x(d\theta).$$

By (9), $s(\theta,x)g(x-\theta) = g'(x-\theta)$, so $\mathbb{E}_{\mu^x}[s(\theta,x)] = f'_\mu(x)/f_\mu(x)$ and (21) follows.

The preceding Tonelli bound and Bayes' rule also give, for almost every x,

$$\int_\Theta (1+|\theta|)|s(\theta,x)|\,\mu^x(d\theta) = \frac{\int_\Theta (1+|\theta|)|g'(x-\theta)|\,\mu(d\theta)}{f_\mu(x)} < \infty. \tag{23}$$

□

Before we move on, we fix some notations. Denote by $\Pi^*(\mu)$ the set of the optimal policies at μ, the argmax of (4). Put differently,

$$\Pi^*(\mu) := \Big\{\pi \in \Pi : \int_\Theta J_\pi\, d\mu = \bar{v}(\mu)\Big\}. \tag{24}$$

Let $\widetilde{\Pi}$ denote the set of mixed policies, the class of functions $\widetilde{\pi}$ with the same domains as (pure) policies but values in $[0,1]$, interpreted as search and acceptance probabilities. Define $J_{\widetilde{\pi}}$ by (2) and $\widetilde{\Pi}^*(\mu)$ similarly.

Lemma B.1. *Let $\mu \in \mathcal{P}_1(\Theta)$.*

(a) The sets $\{J_{\widetilde{\pi}} : \widetilde{\pi} \in \widetilde{\Pi}\}$ and $\{J_{\widetilde{\pi}} : \widetilde{\pi} \in \widetilde{\Pi}^*(\mu)\}$ are nonempty and compact in $L^1(\mu)$.

(b) For every finite signed Borel measure ν with $\|\nu\|_{1,\mathrm{TV}} < \infty$, the maximum and minimum of $\int_\Theta J_\pi\, d\nu$ over $\pi \in \Pi^*(\mu)$ are attained and equal the corresponding extrema over $\widetilde{\pi} \in \widetilde{\Pi}^*(\mu)$.

Proof. (a) Fix $\mu \in \mathcal{P}_1(\Theta)$. By (2), every policy satisfies

$$|J_\pi(\theta)| \le \mathbb{E}_\theta\left[\sum_{j=0}^{\infty} \beta^j(c+|x_j|)\right] \le \frac{c+|\theta|+\mathbb{E}|\varepsilon_0|}{1-\beta}. \tag{25}$$

Thus every payoff function lies in $L^1(\mu)$. The same bound holds for every mixed policy $\widetilde{\pi} \in \widetilde{\Pi}$ as well.

For a policy π, let $\pi_t^S(h_t)$ be the indicator of reaching period t and searching, and let $\pi_t^A(h_t, x_t)$ be the indicator of accepting at t.[13] Similarly, for any given mixed policy $\widetilde{\pi}$, let $\widetilde{\pi}_t^S(h_t)$ be the probability, over policy randomization, of reaching period t and searching, and let $\widetilde{\pi}_t^A(h_t, x_t)$ be the probability of accepting at t. These weights satisfy $0 \le \widetilde{\pi}_0^S \le 1$, $0 \le \widetilde{\pi}_t^A \le \widetilde{\pi}_t^S$, and $0 \le \widetilde{\pi}_{t+1}^S \le \widetilde{\pi}_t^S - \widetilde{\pi}_t^A$. For pure and mixed policies, in terms of the indicators or weights, we can express (2) as

$$J_{\widetilde{\pi}}(\theta) = \mathbb{E}_\theta\left[\sum_{t=0}^{\infty} \beta^t\{-c\widetilde{\pi}_t^S(h_t) + x_t\widetilde{\pi}_t^A(h_t, x_t)\}\right].$$

Take any sequence $(\widetilde{\pi}^n)$ of mixed policies and write $((\widetilde{\pi}^n)_t^S, (\widetilde{\pi}^n)_t^A)$ for their weights. For $t \ge 1$, $(\widetilde{\pi}^n)_t^S$ lies in the unit ball of $L^\infty(\mathbb{R}^t)$, and for $t \ge 0$, $(\widetilde{\pi}^n)_t^A$ lies in that of $L^\infty(\mathbb{R}^{t+1})$. Banach–Alaoglu, separability of the corresponding L^1 spaces, and diagonal extraction give a subsequence on which $(\widetilde{\pi}^n)_0^S \to \widetilde{\pi}_0^S$ and all remaining weights $(\widetilde{\pi}^n)_t^S, (\widetilde{\pi}^n)_t^A$ converge weak-star to $\widetilde{\pi}_t^S, \widetilde{\pi}_t^A$, respectively. For every $t \ge 0$, weak-star convergence gives

$$\mathbb{E}_\theta[-c(\widetilde{\pi}^n)_t^S + x_t(\widetilde{\pi}^n)_t^A] \longrightarrow \mathbb{E}_\theta[-c\widetilde{\pi}_t^S + x_t\widetilde{\pi}_t^A].$$

Since $\beta^t\mathbb{E}_\theta\big[|-c(\widetilde{\pi}^n)_t^S + x_t(\widetilde{\pi}^n)_t^A|\big] \le \beta^t\big(c + |\theta| + \mathbb{E}|\varepsilon_0|\big)$, the dominated convergence theorem then yields $J_{\widetilde{\pi}^n}(\theta) \to J_{\widetilde{\pi}}(\theta)$ for every $\theta \in \Theta$. The common bound (25) is integrable under μ, so again by the dominated convergence theorem,

$$\int_\Theta |J_{\widetilde{\pi}^n}(\theta) - J_{\widetilde{\pi}}(\theta)|\,\mu(d\theta) \longrightarrow 0. \tag{26}$$

Thus every sequence of mixed-policy payoff functions has a subsequence converging in $L^1(\mu)$, which proves the compactness of $\{J_{\widetilde{\pi}} : \widetilde{\pi} \in \widetilde{\Pi}\}$.

By Proposition 1, $\Pi^*(\mu)$ is nonempty, and so is $\widetilde{\Pi}^*(\mu)$. Since $J \mapsto \int_\Theta J\,d\mu$ is continuous on $L^1(\mu)$, the optimal-payoff set $\{J_{\widetilde{\pi}} : \int_\Theta J_{\widetilde{\pi}}\,d\mu = \bar{v}(\mu)\}$ is a closed subset of $\{J_{\widetilde{\pi}} : \widetilde{\pi} \in \widetilde{\Pi}\}$, and is therefore compact.

(b) Fix a mixed policy $\widetilde{\pi}$. Let $U = (U_t^s, U_t^a)_{t\ge 0}$ have independent components, each uniform on $(0, 1)$, and be independent of the state and offers. For each fixed realization

[13] The rules π_t^s and π_t^a specify the decisions at the respective histories, whereas π_t^S and π_t^A also account for whether search has already ended. Thus

$$\pi_t^S(h_t) = \left(\prod_{j=0}^{t-1} \pi_j^s(h_j)\big[1 - \pi_j^a(h_j, x_j)\big]\right)\pi_t^s(h_t), \qquad \pi_t^A(h_t, x_t) = \pi_t^S(h_t)\pi_t^a(h_t, x_t).$$

u, define the (pure) policy $\pi^u \in \Pi$ by

$$(\pi^u)^s_t(h_t) = \mathbf{1}\{u^s_t \le \widetilde{\pi}^s_t(h_t)\}, \qquad (\pi^u)^a_t(h_t, x_t) = \mathbf{1}\{u^a_t \le \widetilde{\pi}^a_t(h_t, x_t)\}.$$

Independence and (2) imply

$$J_{\widetilde{\pi}}(\theta) = \mathbb{E}_U[J_{\pi^U}(\theta)].$$

Moreover, if $\widetilde{\pi} \in \widetilde{\Pi}^*(\mu)$, then

$$\mathbb{E}_U\left[\bar{v}(\mu) - \int_\Theta J_{\pi^U}\, d\mu\right] = 0,$$

which implies $\pi^U \in \Pi^*(\mu)$ for almost every U.

Apply part (a) to $(\mu + |\nu|)/(1 + |\nu|(\Theta)) \in \mathcal{P}_1(\Theta)$. The mixed-policy payoff set is compact in $L^1(\mu + |\nu|)$. So both extrema over $\widetilde{\Pi}^*(\mu)$ are attained. If $\widetilde{\pi}$ attains the maximum, then

$$\int_\Theta J_{\widetilde{\pi}}\, d\nu = \mathbb{E}_U\left[\int_\Theta J_{\pi^U}\, d\nu\right].$$

Almost every π^U is optimal at μ and has its integral at most this maximum. Its average equals the maximum, so almost every such policy attains it. Since $\Pi^*(\mu) \subseteq \widetilde{\Pi}^*(\mu)$, the two maxima coincide and the maximum over policies is attained. Applying the same argument to $-\nu$ proves the assertion for the minimum. □

Lemma 8 (Derivatives of the value along a belief path)**.** *Let $\nu : \mathbb{R} \to \mathcal{P}_1(\Theta)$ be a locally absolutely continuous belief path, with derivative $\dot{\nu}_x = \frac{\partial}{\partial x}\nu_x$ at every x where it is differentiable. Then $x \mapsto \bar{v}(\nu_x)$ is locally absolutely continuous with right and left derivatives*

$$\frac{d^+}{dx}\bar{v}(\nu_x) = \max_{\pi \in \Pi^*(\nu_x)} \int_\Theta J_\pi\, d\dot{\nu}_x, \qquad \frac{d^-}{dx}\bar{v}(\nu_x) = \min_{\pi \in \Pi^*(\nu_x)} \int_\Theta J_\pi\, d\dot{\nu}_x. \tag{27}$$

Applying Lemma 8 to the posterior path $x \mapsto \mu^x$ in Lemma 7 yields the covariance formula.

Proof of Lemma 8. The policy-payoff bound (25) implies, for all beliefs $\mu, \nu \in \mathcal{P}_1(\Theta)$,

$$|\bar{v}(\mu) - \bar{v}(\nu)| \le \frac{c + 1 + \mathbb{E}|\varepsilon_0|}{1 - \beta}\|\mu - \nu\|_{1,\mathrm{TV}}.$$

Thus local absolute continuity of the belief path implies local absolute continuity of its value. Now fix an x at which the belief path is differentiable. Since $(\nu_x + |\dot{\nu}_x|)/(1 + |\dot{\nu}_x|(\Theta)) \in \mathcal{P}_1(\Theta)$, Lemma B.1(a) makes the mixed-policy payoff set com-

pact in $L^1(\nu_x + |\dot{\nu}_x|)$, in which both $J \mapsto \int_\Theta J\, d\nu_x$ and $J \mapsto \int_\Theta J\, d\dot{\nu}_x$ are continuous. By part (b) of Lemma B.1, both extrema in Lemma 8 are attained by policies and equal the corresponding mixed-policy extrema. Since the belief path is continuously differentiable, it is locally Lipschitz, and so is $x \mapsto \bar{v}(\nu_x)$.

Choose $\pi^* \in \Pi^*(\nu_x)$ maximizing $\int_\Theta J_\pi\, d\dot{\nu}_x$. For any sequence $t_n \downarrow 0$, choose $\pi^n \in \Pi^*(\nu_{x+t_n})$, so $\bar{v}(\nu_{x+t_n}) = \int_\Theta J_{\pi^n}\, d\nu_{x+t_n}$. By compactness, along a subsequence $J_{\pi^n} \to J_{\widetilde{\pi}}$ in $L^1(\nu_x + |\dot{\nu}_x|)$ for some $\widetilde{\pi} \in \widetilde{\Pi}$. Note that

$$\int_\Theta J_{\pi^n}\, d\nu_x = \bar{v}(\nu_{x+t_n}) - \int_\Theta J_{\pi^n}\, d(\nu_{x+t_n} - \nu_x) \longrightarrow \bar{v}(\nu_x).$$

Thus $\widetilde{\pi} \in \widetilde{\Pi}^*(\nu_x)$ and $\int_\Theta J_{\pi^n}\, d\dot{\nu}_x \to \int_\Theta J_{\widetilde{\pi}}\, d\dot{\nu}_x$. Since π^* is optimal at ν_x and π^n at ν_{x+t_n}, we have

$$\int_\Theta J_{\pi^*}\, d\left(\frac{\nu_{x+t_n} - \nu_x}{t_n}\right) \leq \frac{\bar{v}(\nu_{x+t_n}) - \bar{v}(\nu_x)}{t_n} \leq \int_\Theta J_{\pi^n}\, d\left(\frac{\nu_{x+t_n} - \nu_x}{t_n}\right).$$

The lower and upper bounds converge to $\int_\Theta J_{\pi^*}\, d\dot{\nu}_x$ and $\int_\Theta J_{\widetilde{\pi}}\, d\dot{\nu}_x$, respectively. By Lemma B.1(b), the latter cannot exceed the former, because $\widetilde{\pi}$ is an optimal auxiliary rule and π^* maximizes the same functional over policies. Hence the right derivative exists and satisfies (27).

Applying this result to the reversed path $t \mapsto \nu_{x-t}$, whose derivative at zero is $-\dot{\nu}_x$, gives the left-derivative formula in (27). □

Lemma 9 (The payoff-score covariance)**.** *Fix $\mu \in \mathcal{P}_1(\Theta)$. The function Γ is locally absolutely continuous. For almost every x,*

$$\Gamma'(x) = 1 - \beta \operatorname{Cov}_{\mu^x}(J_\pi, s(\cdot, x)) \qquad \textit{for every } \pi \in \Pi^*(\mu^x). \tag{28}$$

Proof of Lemma 9. By Lemma 7, $x \mapsto \mu^x$ is locally absolutely continuous in weighted total variation and is differentiable almost everywhere. Applying Lemma 8 shows that $x \mapsto \bar{v}(\mu^x)$ and Γ are locally absolutely continuous.

At almost every x, both the posterior path and $x \mapsto \bar{v}(\mu^x)$ are differentiable, and (21) and (23) hold. Fix such an x and any $\pi \in \Pi^*(\mu^x)$. The payoff bound (25) makes $t \mapsto \int_\Theta J_\pi\, d\mu^{x+t}$ differentiable at zero. Optimality gives

$$\bar{v}(\mu^{x+t}) \geq \int_\Theta J_\pi\, d\mu^{x+t},$$

with equality at $t = 0$. Both sides are differentiable there, so their derivatives coincide. By (21) and (23), the common derivative is $\operatorname{Cov}_{\mu^x}(J_\pi, s(\cdot, x))$. This proves (28),

simultaneously for every optimal policy at every such offer. □

Lemma B.2. *Fix a belief $\mu \in \mathcal{P}_1(\Theta)$ and a nondecreasing Borel function φ satisfying*

$$\int_\Theta (1 + |\theta|)|\varphi(\theta)|\, \mu(d\theta) < \infty. \tag{29}$$

Then

$$\mathrm{Cov}_\mu(J_\mu, \varphi) \leq \mathrm{Cov}_\mu(J_\pi, \varphi) \qquad \textit{for every } \pi \in \Pi^*(\mu). \tag{30}$$

Proof. Compare π and $\pi^*(\mu)$ under the same state and sequence of potential offers. At almost every history reached by π, its action and remaining policy are optimal for the corresponding Bellman problem. Hence the two policies agree at every reached strict decision, apart from null histories. At their first disagreement, $\pi^*(\mu)$ therefore stops and π continues at indifference.

If this disagreement occurs before search, let ν be the current posterior and $\hat{\pi}$ the remaining policy under π. Then $\hat{\pi} \in \Pi^*(\nu)$ and $\bar{v}(\nu) = 0$. The conditional expected payoff difference from searching instead of quitting, weighted by $\varphi(\theta)$, is

$$\int_\Theta \varphi(\theta) J_{\hat{\pi}}(\theta)\, \nu(d\theta) = \mathrm{Cov}_\nu(J_{\hat{\pi}}, \varphi) \geq 0.$$

If the disagreement occurs after an offer x, let ν be the posterior after this offer and $\hat{\pi}$ the policy followed under π starting next period. Then $\hat{\pi} \in \Pi^*(\nu)$ and $x = \beta\bar{v}(\nu)$. The common search cost cancels, and the conditional expected payoff difference from rejecting instead of accepting, weighted by $\varphi(\theta)$, is

$$\int_\Theta \varphi(\theta)\big[\beta J_{\hat{\pi}}(\theta) - x\big]\, \nu(d\theta) = \beta\, \mathrm{Cov}_\nu(J_{\hat{\pi}}, \varphi) \geq 0.$$

Both inequalities follow from Lemma 5. The conditional covariances are finite: the posterior at each finite history has a bounded density relative to μ, so (29) and (25) apply.

It follows that $\int_\Theta \varphi(J_\pi - J_\mu)\, d\mu \geq 0$. Since both policies have expected payoff $\bar{v}(\mu)$, this yields (30). □

Proof of Lemma 4. Part (a) follows from Lemma 9: since $\Gamma' \geq 0$ almost everywhere and Γ is locally absolutely continuous, the fundamental theorem of calculus makes it nondecreasing. Therefore its nonnegative set $A(\mu) = \{x \in \mathbb{R} : \Gamma(x) \geq 0\}$ is an upper interval.

(b) Fix an indifferent offer x satisfying the covariance inequality in the statement. For this proof, truncate the state argument of the likelihood by setting $\varphi_n(\theta) :=$

$\max\{-n, \min\{\theta, n\}\}$ for each integer $n \geq 1$. The statistic $s(\varphi_n(\theta), x)$ is bounded and nondecreasing in θ, converges pointwise to $s(\theta, x)$, and satisfies

$$|s(\varphi_n(\theta), x)| \leq |s(\theta, x)| + |(\log g)'(x)|.$$

Here the truncated score is evaluated using its expression $(\log g)'(x - \varphi_n(\theta))$. Since the covariance in the hypothesis is finite and J_{μ^x} is integrable, dominated convergence implies convergence of the truncated covariances to $\mathrm{Cov}_{\mu^x}(J_{\mu^x}, s(\cdot, x))$. Fix n large enough that $\beta \,\mathrm{Cov}_{\mu^x}(J_{\mu^x}, s(\varphi_n(\cdot), x)) > 1$.

Consider the auxiliary belief path

$$\nu_t(d\theta) = \frac{[g(x + t - \varphi_n(\theta))/g(x - \varphi_n(\theta))]\, \mu^x(d\theta)}{\int_\Theta [g(x + t - \varphi_n(\vartheta))/g(x - \varphi_n(\vartheta))]\, \mu^x(d\vartheta)}.$$

Every ν_t is a belief in $\mathcal{P}_1(\Theta)$, and $\nu_0 = \mu^x$. Since φ_n takes values in $[-n, n]$, positivity and continuous differentiability of g bound the likelihood ratio and its derivative uniformly in the state when t ranges over a compact interval. Dominated convergence and the quotient rule make this path continuously differentiable in weighted total variation, with

$$\dot{\nu}_0(d\theta) = \big[s(\varphi_n(\theta), x) - \mathbb{E}_{\mu^x} s(\varphi_n(\theta), x)\big] \mu^x(d\theta).$$

By Lemma B.2, the selected policy minimizes the covariance with this bounded increasing statistic among optimal policies. Thus Lemma 8 gives

$$\frac{d^-}{dt} \bar{v}(\nu_t)\bigg|_{t=0} = \mathrm{Cov}_{\mu^x}(J_{\mu^x}, s(\varphi_n(\cdot), x)) > \frac{1}{\beta}, \tag{31}$$

while the right derivative is the maximum over optimal policies and is at least as large.

Logconcavity of g makes $\theta \mapsto g(x + t - \theta)/g(x - \theta)$ nondecreasing when $t > 0$ and nonincreasing when $t < 0$. Up to a positive normalizing constant,

$$\frac{d\mu^{x+t}}{d\nu_t}(\theta) \propto \frac{g(x + t - \theta)/g(x - \theta)}{g(x + t - \varphi_n(\theta))/g(x - \varphi_n(\theta))}.$$

The ratio on the right equals one on $[-n, n]$ and has the same monotonicity on each tail, with both tails meeting this constant interval. Hence

$$\mu^{x+t} \succeq_{\mathrm{lr}} \nu_t \quad (t > 0), \qquad \nu_t \succeq_{\mathrm{lr}} \mu^{x+t} \quad (t < 0).$$

All these beliefs are mutually absolutely continuous. By Lemma 2(b), division by t gives the same direction of inequality for both signs of t. Together with the one-sided

derivative bounds above, this yields, for all sufficiently small nonzero t,

$$\frac{\bar{v}(\mu^{x+t}) - \bar{v}(\mu^x)}{t} \geq \frac{\bar{v}(\nu_t) - \bar{v}(\nu_0)}{t} > \frac{1}{\beta}.$$

Since $x = \beta\bar{v}(\mu^x)$, it follows that

$$\frac{\Gamma(x+t)}{t} = 1 - \beta\frac{\bar{v}(\mu^{x+t}) - \bar{v}(\mu^x)}{t} < 0.$$

Therefore $\Gamma(x-k) > 0$ and $\Gamma(x+k) < 0$ for all sufficiently small $k > 0$, as claimed. □

Proof of Lemma 5. Here we show a more general result that

$$\mathrm{Cov}_\mu\big(J_\pi(\theta), \varphi(\theta)\big) \geq 0 \qquad \text{for every } \pi \in \Pi^*(\mu) \tag{32}$$

whenever the covariance exists. Fix $\pi \in \Pi^*(\mu)$. First suppose φ is bounded. Define the belief

$$\nu(d\theta) := \big[1 - t\big(\varphi(\theta) - \mathbb{E}_\mu\varphi\big)\big]\mu(d\theta)$$

for sufficiently small $t > 0$, so that its density relative to μ is strictly positive, bounded, nonincreasing μ-almost everywhere, and has integral one. Hence $\nu \in \mathcal{P}_1(\Theta)$ and $\mu \succeq_{\mathrm{lr}} \nu$. By Lemma 2 and optimality of π at μ,

$$\bar{v}(\mu) \geq \bar{v}(\nu) \geq \int_\Theta J_\pi \, d\nu = \bar{v}(\mu) - t\,\mathrm{Cov}_\mu(J_\pi, \varphi),$$

where the last equality holds by the definition of ν. Since $t > 0$, this proves (32) for every optimal π when φ is bounded.

For an unbounded φ, set $\varphi_n := \max\{-n, \min\{\varphi, n\}\}$. Each φ_n is bounded and nondecreasing μ-almost everywhere, so $\mathrm{Cov}_\mu(J_\pi, \varphi_n) \geq 0$. Whenever the covariance in the statement exists, φ and $J_\pi\varphi$ are μ-integrable. Since $|\varphi_n| \leq |\varphi|$ and $|J_\pi\varphi_n| \leq |J_\pi\varphi|$, dominated convergence yields $\mathrm{Cov}_\mu(J_\pi, \varphi_n) \to \mathrm{Cov}_\mu(J_\pi, \varphi)$, as desired. □

C Covariance Decomposition: Proof of Lemma 6

For a belief μ, we define the (general) boundary penalty for any policy π (not only for the policy $\pi^*(\mu)$ as in the main body) by

$$B_\mu(J_\pi) := p(\underline{\theta})\big[J_\pi(\underline{\theta}) - \bar{v}(\mu)\big] + p(\overline{\theta})\big[\bar{v}(\mu) - J_\pi(\overline{\theta})\big], \tag{33}$$

with an endpoint term omitted when that endpoint is infinite.

Here, we show a result that is more general: For every $\pi \in \Pi^*(\mu^x)$, we have

$$\mathrm{Cov}_{\mu^x}\big(J_\pi(\theta), s(\theta,x)\big) = \mathbb{E}_{\mu^x}[J'_\pi(\theta)] + \mathrm{Cov}_{\mu^x}\big(J_\pi(\theta), (\log p)'(\theta)\big) + B_{\mu^x}(J_\pi), \tag{34}$$

with $0 \le \mathbb{E}_{\mu^x}[J'_\pi(\theta)] \le 1$, $B_{\mu^x}(J_\pi) \le 0$, and $\mathrm{Cov}_{\mu^x}\big(J_\pi(\theta), s(\theta,x)\big) \ge 0$.

We first show an auxiliary lemma.

Lemma C.1 (Expected marginal payoff and boundary penalty)**.** *Let $\mu \in \mathcal{P}_1(\Theta)$ be a belief that admits a density p satisfying the assumptions in Lemma 6.*

(a) For every $\pi \in \Pi$, the payoff J_π is continuously differentiable on $\mathbb{R}$ and satisfies

$$|J'_\pi(\theta)| \le C(1+|\theta|) \tag{35}$$

for some finite constant C depending only on β, c, g.

(b) $0 \le \mathbb{E}_\mu[J'_\pi(\theta)] \le 1$ for every $\pi \in \Pi^*(\mu)$;

(c) $B_\mu(J_\pi) \le 0$ for every $\pi \in \Pi^*(\mu)$.

Proof. (a) By (22), $g, g' \in L^1(\mathbb{R}, (1+|x|)\,dx)$, in which translations are continuous. It follows that $\theta \mapsto g(\cdot - \theta)$ is continuously differentiable in the weighted L^1 norm, with derivative $-g'(\cdot - \theta)$.

Fix a policy π and use the indicators π_t^S, π_t^A from the proof of Lemma B.1. For each finite t, the product history density $\prod_{j=0}^t g(x_j - \theta)$ is continuously differentiable in L^1 with weight $1+|x_t|$. Hence its pairing with $-c\pi_t^S + x_t\pi_t^A$ is continuously differentiable. Since $0 \le \pi_t^S, \pi_t^A \le 1$, for a constant C independent of π, t, θ,

$$\left|\frac{d}{d\theta}\mathbb{E}_\theta[-c\pi_t^S + x_t\pi_t^A]\right| \le C(t+1)(1+|\theta|). \tag{36}$$

Since $|\mathbb{E}_\theta[-c\pi_t^S + x_t\pi_t^A]| \le c + |\theta| + \mathbb{E}|\varepsilon_0|$, the payoff series and its derivative series converge uniformly on compact state intervals, using (36) and $\sum_{t\ge 0}\beta^t(t+1) = (1-\beta)^{-2}$. Termwise differentiation proves that J_π is continuously differentiable with $J'_\pi(\theta) = \sum_{t=0}^\infty \beta^t \frac{d}{d\theta}\mathbb{E}_\theta[-c\pi_t^S + x_t\pi_t^A]$, and gives (35), after enlarging C.

(b) Fix $\pi \in \Pi^*(\mu)$. For $b \in \mathbb{R}$, define the shifted policy π^b by

$$(\pi^b)_t^s(h_t) := \pi_t^s(h_t + b), \qquad (\pi^b)_t^a(h_t, x_t) := \pi_t^a(h_t + b, x_t + b), \tag{37}$$

where $h_t + b$ adds b to each component. Under the realization path, π^b at state θ makes exactly the same decisions as π at state $\theta + b$. Search costs and quitting payoffs

agree, but the accepted offer under π^b is lower by b. At most one offer is accepted, so $J_\pi(\theta+b) - J_{\pi^b}(\theta) \leq \max\{b,0\}$ for both signs of b. By the optimality of π at the belief μ, we have

$$\int_\Theta J_\pi(\theta+b)\,\mu(d\theta) \leq \int_\Theta J_{\pi^b}(\theta)\,\mu(d\theta) + \max\{b,0\} \leq \int_\Theta J_\pi(\theta)\,\mu(d\theta) + \max\{b,0\},$$

which implies

$$\int_\Theta [J_\pi(\theta+b) - J_\pi(\theta)]\,\mu(d\theta) \leq \max\{b,0\}. \tag{38}$$

By (35), dominated convergence permits differentiation under μ at $b=0$. Dividing (38) by $b>0$ and letting $b \downarrow 0$ gives $\mathbb{E}_\mu[J'_\pi] \leq 1$, while dividing by $b<0$ and letting $b \uparrow 0$ gives $\mathbb{E}_\mu[J'_\pi] \geq 0$.

(c) For any $y \in (\underline{\theta}, \overline{\theta})$, apply Lemma 5 to the nondecreasing statistics $-\mathbf{1}\{\theta < y\}$ and $\mathbf{1}\{\theta > y\}$. Both conditioning events have positive probability, so the resulting covariance inequalities give

$$\mathbb{E}_\mu[J_\pi(\theta) \mid \theta < y] \leq \bar{v}(\mu) \leq \mathbb{E}_\mu[J_\pi(\theta) \mid \theta > y]. \tag{39}$$

At a finite lower endpoint, let $y \downarrow \underline{\theta}$, so continuity of J_π gives $J_\pi(\underline{\theta}) \leq \bar{v}(\mu)$. At a finite upper endpoint, let $y \uparrow \overline{\theta}$ to obtain $\bar{v}(\mu) \leq J_\pi(\overline{\theta})$. The endpoint density traces are nonnegative, so (33) yields $B_\mu(J_\pi) \leq 0$. □

Proof of Lemma 6. Fix $x \in \mathbb{R}$ and $\pi \in \Pi^*(\mu^x)$, and write

$$p^x(\theta) := \frac{p(\theta)g(x-\theta)}{f_\mu(x)} \tag{40}$$

for the density of μ^x on $(\underline{\theta}, \overline{\theta})$. It is strictly positive and locally absolutely continuous, with

$$\frac{(p^x)'(\theta)}{p^x(\theta)} = (\log p)'(\theta) - s(\theta, x) \tag{41}$$

almost everywhere. Finite total variation makes p bounded. By (22),

$$\int_\Theta (1+|\theta|)p(\theta)|g'(x-\theta)|\,d\theta \leq \|p\|_\infty \int_{\mathbb{R}} (1+|x-\varepsilon|)|g'(\varepsilon)|\,d\varepsilon < \infty.$$

Moreover, $(1+|\theta|)g(x-\theta)$ is bounded because g has exponential tails, so the finite total variation of p makes $(1+|\theta|)|p'(\theta)|g(x-\theta)$ integrable. The product rule therefore

gives

$$\int_{\underline{\theta}}^{\overline{\theta}} (1+|\theta|)|(p^x)'(\theta)|\,d\theta < \infty. \tag{42}$$

The same bounds show that $(\log p)'$ and $s(\cdot, x)$ are integrable with respect to $(1+|\theta|)\mu^x(d\theta)$. Together with (25), this proves that both covariances in (34) are finite. By Lemma 1 and (35), $\mathbb{E}_{\mu^x}|J'_\pi| < \infty$ as well. The finite total variation of p and Conditions (42) and (25) make sure that $[J_\pi(\theta) - \bar{v}(\mu^x)]p^x(\theta)$ vanishes at each infinite endpoint and converges at a finite endpoint.

Using (33) and integration by parts yield

$$\int_{\underline{\theta}}^{\overline{\theta}} [J_\pi(\theta) - \bar{v}(\mu^x)](p^x)'(\theta)\,d\theta = -B_{\mu^x}(J_\pi) - \mathbb{E}_{\mu^x}[J'_\pi(\theta)]. \tag{43}$$

By (41) and $\mathbb{E}_{\mu^x} J_\pi = \bar{v}(\mu^x)$, the left-hand side is $\mathrm{Cov}_{\mu^x}(J_\pi, (\log p)') - \mathrm{Cov}_{\mu^x}(J_\pi, s(\cdot, x))$. Rearranging proves (34).

Finally, Lemma C.1 applied to μ^x gives $0 \le \mathbb{E}_{\mu^x}[J'_\pi] \le 1$ and $B_{\mu^x}(J_\pi) \le 0$. Since $s(\cdot, x)$ is nondecreasing by logconcavity of g, Lemma 5 gives $\mathrm{Cov}_{\mu^x}(J_\pi, s(\cdot, x)) \ge 0$. □

D Reservation Failure for Heavy-Tailed Beliefs: Proof of Proposition 3

We first fix some notations. For any θ, define *optionality*:

$$\Omega_\theta(\bar{x}) := \int_{\bar{x}}^{\infty} (x - \bar{x}) F(dx|\theta) = \mathbb{E}_\theta[(x-\bar{x})^+] \tag{44}$$

Then given state θ, when $\beta = 1$ and $c > 0$, we can write the Bellman equation for search without learning and quitting as $c = \Omega_\theta(\bar{x})$.

More generally, for $\beta \in (0,1)$ and $c \ge 0$, the Bellman equation for search with full information and without quitting is

$$(1-\beta)\bar{x} + c = \mathbb{E}_\theta\big[(x - \beta\bar{x})^+\big]. \tag{45}$$

Let $\bar{x}_\theta(\beta, c)$ be the solution of (45). Then it is well-defined when (a) $\beta = 1$ and $c > 0$ (in which case $\bar{x}_\theta(1, c) = \Omega_\theta^{-1}(c)$) or (b) $\beta \in (0,1)$ and $c \ge 0$. With the option to quit, the value of the known-state search problem becomes $\bar{v}_{\beta,c}(\delta_\theta) = \big[\bar{x}_\theta(\beta, c)\big]^+$, where δ_θ is the Dirac measure on state θ.

Lemma D.1 (The posterior distribution is "asymptotically fiducial")**.** *Let p be assumed as in Proposition 3.*

(a) As $\theta \to \infty$, we have $(\log p)'(\theta) \longrightarrow 0$ and $(\log p)''(\theta) \longrightarrow 0$. And for all sufficiently large θ,

$$(\log p)'(\theta) < 0, \qquad (\log p)''(\theta) > 0. \tag{46}$$

(b) The density p has a long tail, i.e., for every $0 \leq M < \infty$, as $\theta \to \infty$

$$\sup_{|z| \leq M} \left| \frac{p(\theta - z)}{p(\theta)} - 1 \right| \longrightarrow 0. \tag{47}$$

(c) Let $\mathcal{L}(\varepsilon_0)$ be the distribution (law) of ε_0 and $\mathcal{L}_{\mu^x}(x - \theta)$ of $x - \theta$ where x is a constant and θ has distribution μ^x. Then as $x \to \infty$,

$$\mathcal{L}_{\mu^x}(x - \theta) \longrightarrow \mathcal{L}(\varepsilon_0) \tag{48}$$

in W_1, and

$$x - \mathbb{E}_{\mu^x}[\theta] \longrightarrow \mathbb{E}\varepsilon_0, \qquad \mathbb{E}_{\mu^x}[(x - \theta)^2] \longrightarrow \mathbb{E}[\varepsilon_0^2]. \tag{49}$$

Proof. (a) Choose $\tilde{\theta} > \underline{\theta}$ so that $(\log p)'' \geq 0$ and $(\log p)''' \leq 0$ on $[\tilde{\theta}, \infty)$. Then $(\log p)'$ is nondecreasing there and must be strictly negative, for otherwise p would be nondecreasing on an infinite interval, contradicting integrability. Its limit therefore exists and is nonpositive. A strictly negative limit would give an exponential upper bound on the right tail of p, making $\mathbb{E}_\mu[e^{t\theta}]$ finite for some $t > 0$, contrary to the heavy-tail assumption. Hence $(\log p)' \to 0$.

Since $(\log p)''$ is nonnegative and nonincreasing on $[\tilde{\theta}, \infty)$, it also has a limit. This limit must be zero, since a positive limit would force $(\log p)'$ eventually to be positive. Moreover, if $(\log p)''$ vanished at a point in $[\tilde{\theta}, \infty)$, it would vanish thereafter. Then $(\log p)'$ would be constant on that tail, and its zero limit would make p a positive constant there, again contradicting integrability. This proves (46).

(b) Fix $0 \leq M < \infty$. For all sufficiently large θ, the fundamental theorem of calculus and the monotonicity of $(\log p)'$ give

$$\sup_{|z| \leq M} \left| \log \frac{p(\theta - z)}{p(\theta)} \right| \leq -M(\log p)'(\theta - M) \longrightarrow 0.$$

Exponentiating proves (47).

(c) We adapt the dominated-convergence argument underlying Dawid (1973, Result 2). Finite total variation makes p bounded, and we extend it by zero outside

its support. Logconcavity and integrability of g allow us to choose $b > 0$ so that $(1+\varepsilon^2)e^{b|\varepsilon|}g(\varepsilon)$ is integrable. By part (a), increase $\tilde{\theta}$ until $-b \leq (\log p)' < 0$ above $\tilde{\theta}$. For $x \geq \tilde{\theta}$, integration gives $p(x) \geq p(\tilde{\theta})e^{-b(x-\tilde{\theta})}$ and so, for some $C \geq 1$ and all $\varepsilon \in \mathbb{R}$,

$$0 \leq \frac{p(x-\varepsilon)}{p(x)} \leq Ce^{b\varepsilon^+}.$$

The ratio tends to one for each fixed ε by part (b), so dominated convergence gives

$$\lim_{x\to\infty} \frac{f_\mu(x)}{p(x)} = \lim_{x\to\infty} \int_{\mathbb{R}} \frac{p(x-\varepsilon)}{p(x)} g(\varepsilon)\, d\varepsilon = \int_{\mathbb{R}} \lim_{x\to\infty} \frac{p(x-\varepsilon)}{p(x)} g(\varepsilon)\, d\varepsilon = \int_{\mathbb{R}} g(\varepsilon)\, d\varepsilon = 1.$$

For all sufficiently large x, therefore, $p(x)/f_\mu(x) \leq 2$. The integrable bound $(1+\varepsilon^2)(2Ce^{b\varepsilon^+}+1)g(\varepsilon)$ then permits another application of dominated convergence, yielding

$$\int_{\mathbb{R}} (1+\varepsilon^2) \left| \frac{p(x-\varepsilon)}{f_\mu(x)} - 1 \right| g(\varepsilon)\, d\varepsilon \longrightarrow 0.$$

By Bayes' rule, $p(x-\varepsilon)g(\varepsilon)/f_\mu(x)$ is the density of $\mathcal{L}_{\mu^x}(x-\theta)$. Note that

$$\mathbb{E}_{\mu^x}[\varphi(x-\theta)] - \mathbb{E}[\varphi(\varepsilon_0)] = \int_{\mathbb{R}} \varphi(\varepsilon) \left(\frac{p(x-\varepsilon)}{f_\mu(x)} - 1 \right) g(\varepsilon)\, d\varepsilon.$$

The weighted convergence above implies weak convergence (take bounded continuous φ and use $|\varphi(\varepsilon)| \leq \|\varphi\|_\infty(1+\varepsilon^2)$) and convergence of the first absolute moments (take $\varphi(\varepsilon) = |\varepsilon|$), so (17) gives (48). It also gives the mean limit in (49), while the quadratic weight $\varphi(\varepsilon) = \varepsilon^2$ yields the second-moment limit in that equation. □

Lemma D.2 (High-patience value limit)**.** *Let $c > 0$ and $\mu \in \mathcal{P}_1(\Theta)$. For all discount factor $\beta \in (0,1)$, the value satisfies*

$$0 \leq \bar{v}_{\beta,c}(\mu) \leq \mathbb{E}_\mu[(\theta + \bar{x}_0(1,c))^+] < \infty. \tag{50}$$

Consequently, the limit $\bar{v}_{1,c}(\mu) := \lim_{\beta\uparrow 1} \bar{v}_{\beta,c}(\mu)$ exists and is finite.

Proof. By the location structure, $\Omega_\theta(\bar{x}) = \Omega_0(\bar{x}-\theta)$. Hence the defining equation $\Omega_\theta(\bar{x}_\theta(1,c)) = c$ gives $\bar{x}_\theta(1,c) = \theta + \bar{x}_0(1,c)$, which is the reservation offer for the known-state search problem without quitting. With quitting, the value becomes $\bar{v}_{1,c}(\delta_\theta) = [\theta + \bar{x}_0(1,c)]^+$. For general discount factor $\beta \in (0,1)$, it is clear that the value satisfies $\bar{v}_{\beta,c}(\delta_\theta) \leq [\theta + \bar{x}_0(1,c)]^+$. Take any policy π available to the searcher with belief μ. A searcher who is told the true state θ can always ignore that information and follow the same policy. Therefore, $J_\pi(\theta) \leq \bar{v}_{\beta,c}(\delta_\theta) \leq [\theta + \bar{x}_0(1,c)]^+$. Integrating with respect to

μ gives

$$\int_\Theta J_\pi(\theta)\,\mu(d\theta) \le \mathbb{E}_\mu\big[(\theta + \bar{x}_0(1,c))^+\big].$$

The upper bound is finite because $\mathbb{E}_\mu[(\theta + \bar{x}_0(1,c))^+] \le \mathbb{E}_\mu|\theta| + |\bar{x}_0(1,c)| < \infty$. Taking the supremum over policies yields the desired bound (50), which, together with the discount factor monotonicity of $\bar{v}$ in Proposition 1, implies the existence and finiteness of the limit. □

Lemma D.3 (Uniform duration and vanishing-discount bounds at positive costs)**.** *Fix $0 < c_0 < c_1 < \infty$. There exist $\beta_0 \in (0,1)$ and $C, b > 0$ such that the following statements hold for every $\mu \in \mathcal{P}_1(\Theta)$, $\beta \in [\beta_0, 1)$, $c \in [c_0, c_1]$, and $\pi \in \Pi^*_{\beta,c}(\mu)$.*

(a) The search duration N_π (the number of searches made before acceptance or quitting) satisfies

$$\Pr_\mu(N_\pi > n) \le Ce^{-bn} \tag{51}$$

for $n \ge 0$. Moreover, for $c' \in [c_0, c_1]$,

$$|\bar{v}_{\beta,c}(\mu) - \bar{v}_{\beta,c'}(\mu)| \le C|c - c'|. \tag{52}$$

(b) The expected marginal payoff equals the expected discounted acceptance probability, and

$$\mathbb{E}_\mu[J'_\pi] = \sum_{t\ge 0} \beta^t \mathbb{E}_\mu[\pi^A_t] \ge 1 - C(1-\beta) - \Pr_\mu(\pi \text{ quits}). \tag{53}$$

(c) If $\mathbb{E}_\mu\theta^2 < \infty$, then

$$0 \le \bar{v}_{1,c}(\mu) - \bar{v}_{\beta,c}(\mu) \le C(1-\beta)\big(1 + (\mathbb{E}_\mu\theta^2)^{1/2}\big). \tag{54}$$

Proof. Note that for any policy π, we have $\pi^S_t = \mathbf{1}\{N_\pi > t\}$ and $\pi^A_t = \mathbf{1}\{\pi \text{ accepts at } t\} \le \mathbf{1}\{N_\pi > t\}$.

(a) Searching once and accepting only nonnegative offers gives

$$\bar{v}_{\beta,c}(\mu) \ge \mathbb{E}_\mu[(\theta + \varepsilon_0)^+] - c \ge \mathbb{E}_\mu[(\theta + \mathbb{E}\varepsilon_0)^+] - c. \tag{55}$$

Evaluating π at cost $d = c_0/2$, applying (50) and (55), and using the fact that $x \mapsto x^+$ is 1-Lipschitz yield

$$(c-d)\sum_{t\ge 0}\beta^t\mathbb{E}_\mu[\pi^S_t] \le \bar{v}_{\beta,d}(\mu) - \bar{v}_{\beta,c}(\mu) \le c_1 + (\bar{x}_0(1,d) - \mathbb{E}\varepsilon_0)^+,$$

which implies $\sum_{t\geq 0}\beta^t\mathbb{E}_\mu[\pi_t^S] \leq [c_1 + (\bar{x}_0(1,d) - \mathbb{E}\varepsilon_0)^+]/d$. Choosing an integer H greater than four times this upper bound and $\beta_0 < 1$ sufficiently close to one yields $\Pr_\mu(N_\pi > kH) \leq 2^{-k}$ for $k = 0, 1, 2, \ldots$ by induction, and hence

$$\Pr_\mu(N_\pi > n) \leq 2^{-\lfloor n/H\rfloor} \leq 2\exp\bigl(-n\log(2)/H\bigr).$$

Next, take $C = [c_1 + (\bar{x}_0(1,d) - \mathbb{E}\varepsilon_0)^+]/d$. Evaluating a policy optimal at cost c at the higher cost c' gives

$$\bar{v}_{\beta,c'}(\mu) \geq \bar{v}_{\beta,c}(\mu) - (c'-c)\sum_{t\geq 0}\beta^t\mathbb{E}_\mu[\pi_t^S] \geq \bar{v}_{\beta,c}(\mu) - C(c'-c).$$

Cost monotonicity of $\bar{v}$ in Proposition 1 then proves (52).

(b) Using the shifted policy π^b in (37), we have

$$\int_\Theta J_{\pi^b}(\theta)\,\mu(d\theta) = \int_\Theta J_\pi(\theta+b)\,\mu(d\theta) - b\sum_{t\geq 0}\beta^t\int_\Theta \mathbb{E}_{\theta+b}[\pi_t^A]\,\mu(d\theta).$$

Optimality of π makes the right-hand side maximal at $b = 0$. Differentiation there gives the equality in (53). The first integral can be differentiated by (35), whose proof does not use a belief density.

By part (a), $\mathbb{E}_\mu[N_\pi] = \sum_{n\geq 0}\Pr_\mu(N_\pi > n) \leq C$ after enlarging the constant C if necessary, and search ends almost surely. If acceptance occurs at date t, then $t = N_\pi - 1$ and $1 - \beta^t \leq (1-\beta)t \leq (1-\beta)N_\pi$. Note that $\sum_{t\geq 0}\mathbb{E}_\mu[\pi_t^A] = \Pr_\mu(\pi \text{ accepts}) = 1 - \Pr_\mu(\pi \text{ quits})$, which implies

$$1 - \Pr_\mu(\pi \text{ quits}) - \sum_{t\geq 0}\beta^t\mathbb{E}_\mu[\pi_t^A] = \sum_{t\geq 0}(1-\beta^t)\Pr_\mu(\pi \text{ accepts at } t).$$

Applying the preceding bounds and taking expectations prove the inequality.

(c) For $\beta \leq \hat{\beta} < 1$, evaluate at discount factor β a policy $\hat{\pi}$ optimal at $\hat{\beta}$. Since $(\hat{\beta})^t - \beta^t \leq t(1-\beta)$, we have

$$\bar{v}_{\hat{\beta},c}(\mu) - \bar{v}_{\beta,c}(\mu) \leq (1-\beta)\sum_{t\geq 0} t\,\mathbb{E}_\mu[x_t^+\hat{\pi}_t^A].$$

Since $x_t^+ \leq |\theta| + |\varepsilon_t|$, we can separate the state and noise contribution to the sum on the right-hand side. The state contribution to the sum is at most $\mathbb{E}_\mu[|\theta|N_{\hat{\pi}}] \leq C(\mathbb{E}_\mu\theta^2)^{1/2}$

by Cauchy–Schwarz and part (a). For the noise contribution, $\hat{\pi}_t^A \le \mathbf{1}\{N_{\hat{\pi}} > t\}$ gives

$$\sum_{t\ge 0} t\,\mathbb{E}_\mu[|\varepsilon_t|\hat{\pi}_t^A] \le (\mathbb{E}\varepsilon_0^2)^{1/2} \sum_{t\ge 0} t\,\Pr_\mu(N_{\hat{\pi}} > t)^{1/2} \le C.$$

Letting $\hat{\beta} \uparrow 1$ and using Lemma D.2 prove (54). □

Proof of Proposition 3. Jensen's inequality gives $\tilde{\tilde{c}}_g \ge D_g > \tilde{c}_g > 0$. Apply Lemma D.3 with $[c_0, c_1] = [\tilde{c}_g, \tilde{\tilde{c}}_g]$. Below, $C, b > 0$ may change from line to line but are independent of $\beta \in [\beta_0, 1)$, $c \in [\tilde{c}_g, \tilde{\tilde{c}}_g]$, and the optimal policy. Write

$$\Gamma_{\beta,c}(x) := x - \beta \bar{v}_{\beta,c}(\mu^x), \qquad \Gamma_{1,c}(x) := x - \bar{v}_{1,c}(\mu^x). \tag{56}$$

Step 1: Tail estimates and remote log-curvature. By Lemma D.1(c), $\mathbb{E}_{\mu^x}\theta^2 \le C(1+x^2)$ for all large x. Also, $(\log p)' \to 0$ gives $\log p(x) = o(x)$, so $f_\mu(y)^{-1} = e^{o(x)}$ uniformly for $|y - x| \le 1$. Finite total variation and exponential decay of g imply

$$\sup_{|y-x|\le 1} \int_{\Theta\setminus[x/2,2x]} \frac{(1+x+|\theta|)^2 g(y-\theta)}{f_\mu(y)} [p(\theta) + |p'(\theta)|]\, d\theta \le Ce^{-bx} \tag{57}$$

as $|y-\theta| \ge x/2 - 1$ on the integration region, exponential decay absorbs the polynomial factor, and the remaining integral of $p + |p'|$ is finite. By equations (50) and (54), we also have

$$0 \le \Gamma_{\beta,c}(x) - \Gamma_{1,c}(x) \le C(1+x)(1-\beta). \tag{58}$$

For any policy $\pi \in \Pi^*(\mu^x)$, quitting at a reached history requires $\bar{v}_{\beta,c}(\mu_t) = 0$. The feasible one-search payoff then implies $\mathbb{E}_{\mu_t}\theta \le c_1 - \mathbb{E}\varepsilon_0$. Thus the nonnegative martingale $\mathbb{E}_{\mu^x}[(x/2-\theta)^+ \mid h_t]$ reaches at least $x/2 - c_1 + \mathbb{E}\varepsilon_0$. Its maximal inequality and (57) give

$$\Pr_{\mu^x}(\pi \text{ quits}) \le \frac{\mathbb{E}_{\mu^x}[(x/2-\theta)^+]}{x/2 - c_1 + \mathbb{E}\varepsilon_0} \le Ce^{-bx}.$$

The maximal inequality follows by stopping at the first crossing before a finite date and then increasing that date. Consequently, (53) yields

$$\mathbb{E}_{\mu^x}[J'_\pi] \ge 1 - C(1-\beta) - Ce^{-bx}. \tag{59}$$

Choose $\tilde{\theta} > \max\{0, \underline{\theta}\}$ above which (46) holds. Since $(\log p)' \to 0$ and $p(x) \to 0$, Tonelli's theorem gives

$$\int_{\tilde{\theta}}^\infty (t - \tilde{\theta})(\log p)''(t)\, dt = \int_{\tilde{\theta}}^\infty -(\log p)'(\theta)\, d\theta = \infty.$$

Hence $(\log p)''$ cannot be bounded above by e^{-at} eventually for any $a > 0$. Its positivity and convergence to zero let us choose $x_n \uparrow \infty$ such that

$$\delta_n := \frac{(\log p)''(2x_n)}{2\Lambda}, \qquad 0 < \delta_n < \tfrac{1}{2}, \qquad \log(1/\delta_n) = o(x_n). \tag{60}$$

Monotonicity gives $(\log p)'' \geq 2\Lambda\delta_n$ on $[x_n/2, 2x_n]$, and $e^{-bx_n} = o(\delta_n)$ for every fixed $b > 0$.

Step 2: A negative slope at suitably chosen discount factors. Fix $|x - x_n| \leq 1$ and $\pi \in \Pi^*(\mu^x)$. Put $\varphi(\theta) := (\log p)'(\theta) - \delta_n s(\theta, x)$. On $[x_n/2, 2x_n]$,

$$\varphi'(\theta) = (\log p)''(\theta) + \delta_n(\log g)''(x - \theta) \geq \Lambda\delta_n > 0.$$

Let $\widetilde{\varphi}$ equal φ on this interval and be constant on either side, taking the corresponding endpoint value. It is bounded and nondecreasing, so Lemma 5 gives $\mathrm{Cov}_{\mu^x}(J_\pi, \widetilde{\varphi}) \geq 0$.

The score grows at most linearly, and $(\log p)'$ is bounded on the right tail. Off $[x_n/2, 2x_n]$,

$$p(\theta)|\varphi(\theta) - \widetilde{\varphi}(\theta)| \leq |p'(\theta)| + C(1 + x_n + |\theta|)p(\theta).$$

Using (25), (50), and (57) therefore gives

$$\mathrm{Cov}_{\mu^x}(J_\pi, \varphi) \geq -\frac{Ce^{-bx_n}}{1-\beta}, \qquad B_{\mu^x}(J_\pi) \geq -\frac{Ce^{-bx_n}}{1-\beta}. \tag{61}$$

For the boundary bound, the only possible term is at a finite $\underline{\theta}$, where $p(\underline{\theta}) < \infty$ and $g(x - \underline{\theta})/f_\mu(x)$ is exponentially small. If $\underline{\theta} = -\infty$, the boundary penalty is zero.

Apply the exact decomposition (34) with $(\log p)' = \varphi + \delta_n s(\cdot, x)$. Equations (59) and (61) yield

$$(1 - \delta_n)\,\mathrm{Cov}_{\mu^x}(J_\pi, s(\cdot, x)) \geq 1 - C(1 - \beta) - \frac{Ce^{-bx_n}}{1-\beta}. \tag{62}$$

Choose $\beta_n := 1 - e^{-bx_n/2}$, using the b in this bound. Then $\beta_n \geq \beta_0$ eventually, and both error terms, as well as $1 - \beta_n$, are $o(\delta_n)$. Thus, uniformly over these offers, costs, and policies, for all large n,

$$\beta_n\,\mathrm{Cov}_{\mu^x}(J_\pi, s(\cdot, x)) \geq \frac{1 - \delta_n/2}{1 - \delta_n} \geq 1 + \delta_n/2.$$

By (28), $\Gamma'_{\beta_n,c}(x) \leq -\delta_n/2$ for almost every $x \in [x_n - 1, x_n + 1]$. Local absolute

continuity and the fundamental theorem of calculus give

$$\Gamma_{\beta_n,c}(y) - \Gamma_{\beta_n,c}(x) \le -\tfrac{1}{2}\delta_n(y-x) \qquad (x_n - 1 \le x < y \le x_n + 1). \tag{63}$$

Moreover, (58) and (60) imply, uniformly over the same offers and costs,

$$0 \le \Gamma_{\beta_n,c}(x) - \Gamma_{1,c}(x) \le C(1+x_n)e^{-bx_n/2} = o(\delta_n). \tag{64}$$

Step 3: Locate an indifference cost and two strict witnesses. By (52), $c \mapsto \Gamma_{1,c}(x)$ is Lipschitz on $[\tilde{c}_g, \tilde{\tilde{c}}_g]$. At posterior μ^x, let X_1, X_2 be the next two potential offers. Consider searching once, accepting X_1 if $X_1 \ge x$, and otherwise searching once more and accepting X_2 unconditionally. Taking $\beta \uparrow 1$ for this finite policy gives

$$\bar{v}_{1,c}(\mu^x) - x \ge \mathbb{E}_{\mu^x}\big[(X_1 - x)\mathbf{1}\{X_1 \ge x\} + (X_2 - x)\mathbf{1}\{X_1 < x\}\big] - c\big[1 + \Pr_{\mu^x}(X_1 < x)\big].$$

By Lemma D.1(c), the first expectation tends to $D_g/4$, while $\Pr_{\mu^x}(X_1 < x) \to 1/2$. Indeed, write $X_i - x = \varepsilon_i - (x - \theta)$ with independent noises. Conditional on $x - \theta$, the expectations are continuous functions with at most linear growth, so W_1 convergence applies. Exchangeability gives

$$\mathbb{E}[(\varepsilon_1 - \varepsilon_0)^+] = D_g/2, \qquad \mathbb{E}[(\varepsilon_2 - \varepsilon_0)\mathbf{1}\{\varepsilon_1 < \varepsilon_0\}] = -D_g/4,$$

where ε_2 is another independent draw from g. For the second identity, $\mathbb{E}[\varepsilon_0\mathbf{1}\{\varepsilon_1 < \varepsilon_0\}] = \frac{1}{2}\mathbb{E}\max\{\varepsilon_0, \varepsilon_1\} = \frac{1}{2}\mathbb{E}\varepsilon_0 + D_g/4$. Consequently,

$$\limsup_{x\to\infty} \Gamma_{1,\tilde{c}_g}(x) \le -D_g/4 + \tfrac{3}{2}\tilde{c}_g = -D_g/8 < 0. \tag{65}$$

At the upper cost, $\bar{x}_0(1, \tilde{\tilde{c}}_g) = \mathbb{E}\varepsilon_0 - D_g$. Letting $\beta \uparrow 1$ in (50) gives

$$\Gamma_{1,\tilde{\tilde{c}}_g}(x) \ge x - \mathbb{E}_{\mu^x}[(\theta + \mathbb{E}\varepsilon_0 - D_g)^+] \longrightarrow D_g > 0.$$

The expression on the right converges by (49) as its lower-tail correction vanishes by (57).

Choose one sufficiently large n so that both cost signs hold at x_n, (63) holds, and the error in (64) is at most $\delta_n/8$. Cost continuity gives $c_* \in (\tilde{c}_g, \tilde{\tilde{c}}_g)$ with $\Gamma_{1,c_*}(x_n) = 0$. Put $x^A = x_n - 1/2$ and $x^R = x_n + 1/2$. Equations (63) and (64) yield

$$\Gamma_{1,c_*}(x^A) \ge \delta_n/8 > 0, \qquad \Gamma_{1,c_*}(x^R) \le -\delta_n/8 < 0. \tag{66}$$

Step 4: Obtain a fixed cost interval and one patience threshold. By cost continuity, choose $\tilde{c}_g < \underline{c} < c_* < \bar{c} < \tilde{\tilde{c}}_g$ sufficiently close to c_* that

$$\Gamma_{1,\underline{c}}(x^A) > 0, \qquad \Gamma_{1,\bar{c}}(x^R) < 0.$$

Since $\bar{v}_{\beta,\underline{c}} \geq 0$ is nondecreasing in β, $\Gamma_{\beta,\underline{c}}(x^A) \geq \Gamma_{1,\underline{c}}(x^A) > 0$ for every $\beta < 1$. At the other endpoint, convergence to $\Gamma_{1,\bar{c}}(x^R)$ supplies $\bar{\beta} \in (0,1)$ such that $\Gamma_{\beta,\bar{c}}(x^R) < 0$ whenever $\beta > \bar{\beta}$. Finally, $\Gamma_{\beta,c}$ is nondecreasing in c. Therefore, simultaneously for every $c \in [\underline{c}, \bar{c}]$ and $\beta \in (\bar{\beta}, 1)$,

$$\Gamma_{\beta,c}(x^A) \geq \Gamma_{\beta,\underline{c}}(x^A) > 0, \qquad \Gamma_{\beta,c}(x^R) \leq \Gamma_{\beta,\bar{c}}(x^R) < 0. \tag{67}$$

Thus $x^A < x^R$ are strictly accepted and strictly rejected, respectively, throughout the claimed cost interval and patience range. □

E Reservation Failure for Discrete Beliefs: Proof of Proposition 5

Lemma E.1. *For every $\mu \in \mathcal{P}_1(\Theta)$,*

$$\mathbb{E}_\mu \theta + \mathbb{E}\varepsilon_0 - c \leq \bar{v}(\mu) \leq \mathbb{E}_\mu[\theta^+] + \frac{\mathbb{E}[\varepsilon_0^+]}{1-\beta}. \tag{68}$$

Proof. The lower bound follows by searching once and accepting the offer. For the upper bound, use the acceptance indicators π_t^A from the proof of Lemma B.1. At most one offer is accepted, so $\sum_{t\geq 0} \beta^t \mathbb{E}_\theta[\pi_t^A] \leq 1$. Dropping the nonnegative search costs and using $0 \leq \pi_t^A \leq 1$ gives

$$J_\pi(\theta) \leq \theta^+ \sum_{t\geq 0} \beta^t \mathbb{E}_\theta[\pi_t^A] + \sum_{t\geq 0} \beta^t \mathbb{E}_\theta[\varepsilon_t^+ \pi_t^A] \leq \theta^+ + \frac{\mathbb{E}[\varepsilon_0^+]}{1-\beta}.$$

Integrating with respect to μ and taking the supremum over policies proves the upper bound. □

Fix $\beta \in (0,1)$ and $c \geq 0$, and let $\eta > 0$ satisfy $\eta \leq -(\log g)''$. Write $\Delta := \theta_H - \theta_L$ and set $\hat{q} := 1/2$. The lower value bound and the pointwise upper bound in the proof of Lemma E.1, together with $\theta_L \geq 0$, give

$$\bar{v}(\mu_{\hat{q}}) - J_{\mu_{\hat{q}}}(\theta_L) \geq \frac{\Delta}{2} + \mathbb{E}\varepsilon_0 - c - \frac{\mathbb{E}[\varepsilon_0^+]}{1-\beta}. \tag{69}$$

Moreover, strong logconcavity implies, for every $x \in \mathbb{R}$,

$$s(\theta_H, x) - s(\theta_L, x) = -\int_{x-\theta_H}^{x-\theta_L} (\log g)''(\varepsilon)\, d\varepsilon \geq \eta\Delta. \tag{70}$$

Set $x^\star := \beta\bar{v}(\mu_{\hat{q}})$ and choose the prior probability

$$q^\star := \frac{g(x^\star - \theta_L)}{g(x^\star - \theta_L) + g(x^\star - \theta_H)} \in (0,1) \tag{71}$$

so that $\mu_{q^\star}^{x^\star} = \mu_{\hat{q}}$. Thus $x^\star$ is an indifferent offer under prior $\mu_{q^\star}$.

Optimality of $\pi^*(\mu_{\hat{q}})$ gives $\bar{v}(\mu_{\hat{q}}) - J_{\mu_{\hat{q}}}(\theta_L) = \hat{q}[J_{\mu_{\hat{q}}}(\theta_H) - J_{\mu_{\hat{q}}}(\theta_L)]$. Applying (16) at the prior $\mu_{q^\star}$ and offer $x^\star$ therefore yields

$$\mathrm{Cov}_{\mu_{\hat{q}}}(J_{\mu_{\hat{q}}}, s(\cdot, x^\star)) = (1-\hat{q})\big[\bar{v}(\mu_{\hat{q}}) - J_{\mu_{\hat{q}}}(\theta_L)\big]\big[s(\theta_H, x^\star) - s(\theta_L, x^\star)\big]. \tag{72}$$

For all sufficiently large Δ, the right-hand side of (69) is positive. Applying the identity (72) with $\hat{q} = 1/2$, and using (69) and (70), therefore yields

$$\mathrm{Cov}_{\mu_{\hat{q}}}(J_{\mu_{\hat{q}}}, s(\cdot, x^\star)) \geq \frac{\eta\Delta}{2}\left[\frac{\Delta}{2} + \mathbb{E}\varepsilon_0 - c - \frac{\mathbb{E}[\varepsilon_0^+]}{1-\beta}\right]. \tag{73}$$

The lower bound tends to ∞ as $\Delta \to \infty$, uniformly over nonnegative state levels and optimal policies. Choose $\Delta_0 > 0$ large enough that

$$\beta\, \mathrm{Cov}_{\mu_{\hat{q}}}(J_{\mu_{\hat{q}}}, s(\cdot, x^\star)) > 1 \tag{74}$$

holds whenever $\Delta \geq \Delta_0$.

By Lemma 4(b), for some $k > 0$, the offer $x^\star - k$ is strictly accepted whereas $x^\star + k$ is strictly rejected at prior $\mu_{q^\star}$. For each of these two fixed offers, the posterior depends continuously on q in W_1, and Proposition 1 gives continuity of $\bar{v}$. Both strict inequalities therefore persist on a nonempty open interval $Q(\theta_L, \theta_H) \subset (0,1)$ containing $q^\star$.

Corollary 2 (Reservation failure with finitely many states)**.** *Fix $\beta \in (0,1)$, $c \geq 0$, and a strongly logconcave noise density g. There exists $\Delta_0 > 0$ such that, for every finite state space $\Theta \subset [0,\infty)$ satisfying $\max\Theta - \min\Theta \geq \Delta_0$, a reservation policy fails under some belief assigning positive probability to every state.*

Proof. Let $\theta_L := \min\Theta$ and $\theta_H := \max\Theta$, and take Δ_0 from the proof of Proposition 5. For any Θ satisfying the stated conditions, that proposition and its proof supply a two-

state belief $\mu_q = (1-q)\delta_{\theta_L} + q\delta_{\theta_H}$ with $q \in (0,1)$ and offers $x^A < x^R$ such that

$$x^A - \beta\bar{v}(\mu_q^{x^A}) > 0, \qquad x^R - \beta\bar{v}(\mu_q^{x^R}) < 0.$$

Viewing μ_q as a belief on Θ leaves these inequalities unchanged: the intermediate states have zero probability, so all policy values and posterior beliefs coincide with those in the two-state problem.

For either fixed offer, Bayes updating is continuous in the prior by Lemma A.1(b), and the optimal value is continuous by Proposition 1. Both strict inequalities therefore persist on a relatively open neighborhood of μ_q in $\mathcal{P}_1(\Theta)$. This neighborhood contains full-support beliefs: if ν assigns positive probability to every state, then $(1-t)\mu_q + t\nu$ belongs to the neighborhood and has full support for all sufficiently small $t > 0$. Because Θ is finite, its full-support beliefs form a relatively open subset of $\mathcal{P}_1(\Theta)$. Intersecting that subset with the neighborhood yields a nonempty relatively open set of full-support beliefs. At every belief in this set, x^A is strictly accepted while x^R is strictly rejected, so a reservation policy fails to be optimal. □

References

Adam, K. (2001): "Learning While Searching for the Best Alternative," *Journal of Economic Theory*, 101(1), 252–280.

Albright, S. C. (1977): "A Bayesian Approach to a Generalized House Selling Problem," *Management Science*, 24(4), 432–440.

Bertsekas, D. P., and S. E. Shreve (1978): *Stochastic Optimal Control: The Discrete-Time Case*. Academic Press.

Bickel, P. J., and D. A. Freedman (1981): "Some Asymptotic Theory for the Bootstrap," *The Annals of Statistics*, 9(6), 1196–1217.

Bikhchandani, S., and S. Sharma (1996): "Optimal Search with Learning," *Journal of Economic Dynamics and Control*, 20(1-3), 333–359.

Boyd, III, J. H. (1990): "Recursive Utility and the Ramsey Problem," *Journal of Economic Theory*, 50(2), 326–345.

Burdett, K., and T. Vishwanath (1988): "Declining Reservation Wages and Learning," *The Review of Economic Studies*, 55(4), 655–665.

Cox, J. C., and R. L. Oaxaca (2000): "Good News and Bad News: Search from Unknown Wage Offer Distributions," *Experimental Economics*, 2(3), 197–225.

Dawid, A. P. (1973): "Posterior Expectations for Large Observations," *Biometrika*, 60(3), 664–667.

DeGroot, M. H. (1968): "Some Problems of Optimal Stopping," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 30(1), 108–122.

Gershkov, A., and B. Moldovanu (2012): "Optimal Search, Learning and Implementation," *Journal of Economic Theory*, 147(3), 881–909.

Karlin, S. (1962): "Stochastic Models and Optimal Policy for Selling an Asset," in *Studies in Applied Probability and Management Science*, ed. by K. J. Arrow, S. Karlin, and H. Scarf, chap. 9, pp. 148–158. Stanford University Press.

McCall, J. J. (1970): "Economics of Information and Job Search," *The Quarterly Journal of Economics*, 84(1), 113–126.

Milgrom, P. R. (1981): "Good News and Bad News: Representation Theorems and Applications," *The Bell Journal of Economics*, pp. 380–391.

Morris, S., and M. Yildiz (2019): "Crises: Equilibrium Shifts and Large Shocks," *American Economic Review*, 109(8), 2823–2854.

Moscarini, G., and L. Smith (2001): "The Optimal Level of Experimentation," *Econometrica*, 69, 1629–1644.

Nyarko, Y. (1994): "On the Convexity of the Value Function in Bayesian Optimal Control Problems," *Economic Theory*, 4(2), 303–309.

Prékopa, A. (1973): "On logarithmic concave measures and functions," *Acta Sci. Math.*, 34, 335–343.

Rosenfield, D. B., and R. D. Shapiro (1981): "Optimal Adaptive Price Search," *Journal of Economic Theory*, 25(1), 1–20.

Rothschild, M. (1974): "Searching for the Lowest Price When the Distribution of Prices Is Unknown," *Journal of Political Economy*, 82(4), 689–711.

Seierstad, A. (1992): "Reservation Prices in Optimal Stopping," *Operations Research*, 40(2), 409–415.

Stokey, N. L., and R. E. Lucas (1989): *Recursive Methods in Economic Dynamics*. Harvard University Press.

Weitzman, M. L. (2007): "Subjective Expectations and Asset-Return Puzzles," *American Economic Review*, 97(4), 1102–1130.

Whitt, W. (1979): "A Note on the Influence of the Sample on the Posterior Distribution," *Journal of the American Statistical Association*, 74(366a), 424–426.