Learning in Job Search:
Bayesian Beliefs, Reservation Wages, and Fast Policy Computation

Hazel Chen

September 23, 2025

Abstract

Unemployed workers typically do not know the true offer distribution; instead, they learn from observed offers while deciding when to accept a job. This project studies job search with Bayesian learning about the offer distribution and develops a fast algorithm to compute the reservation-wage policy. We (i) formalize the dynamic programming problem with beliefs as a state, (ii) derive a one-dimensional fixed-point equation for the reservation wage as a function of belief, and (iii) compare a standard two-dimensional value-function iteration (VFI) with a contraction on the reservation-wage map that is orders of magnitude faster. We provide a validation plan, comparative-statics experiments, and an aggregate regime-change simulation. Implementation follows the QuantEcon "Job Search VII (Search with Learning)" lecture and the companion code structure (e.g., `SearchProblem`, `operator_factory`, `Q_factory`).

1 Motivation and Questions

Workers observe a sequence of wage offers and choose to accept or reject. When the offer-generating process is unknown, beliefs evolve via Bayes' rule, creating a state variable that interacts with stopping decisions. We ask:

1. How do prior beliefs and signal informativeness shape the reservation-wage function $\bar{w}(\pi)$ and unemployment durations?

2. How do policy parameters (e.g., unemployment compensation c) shift optimal stopping under learning?

3. What are the computational trade-offs between classic VFI and a one-dimensional reservation-wage fixed-point iteration?

2 Model

2.1 Known Distribution (McCall Model)

Time is discrete, $\beta \in (0, 1)$, per-period unemployment compensation $c \geq 0$. Each period a wage w is drawn from a known density q on $[0, \bar{w}_{\max}]$. If the worker accepts, she receives w forever; if she rejects, she receives c and continues. Let $v(w)$ be the value upon seeing w:

$$v(w) = \max\left\{\frac{w}{1-\beta},\ c + \beta\,\mathbb{E}\big[v(w')\big]\right\}, \tag{1}$$

which implies a *constant* reservation wage $\bar{w}$.

2.2 Unknown Distribution with Bayesian Learning

Now the worker is uncertain between two candidate densities f and g (on the same support). Let $\pi_t = \Pr(q = f \mid \mathcal{F}_t)$ be the belief that f generates offers at date t. Given prior π_t and observed offer w_{t+1}, the belief updates via Bayes' rule

$$\kappa(w, \pi) \;=\; \frac{\pi f(w)}{\pi f(w) + (1-\pi) g(w)}. \tag{2}$$

Define the mixture density $q_\pi(w) := \pi f(w) + (1-\pi) g(w)$. The Bellman equation with state (w, π) is

$$v(w, \pi) \;=\; \max\left\{ \frac{w}{1-\beta},\; c + \beta \int v\big(w', \kappa(w', \pi)\big)\, q_\pi(w')\, dw' \right\}. \tag{3}$$

By standard arguments, the optimal policy is a cutoff: accept iff $w \geq \bar{w}(\pi)$, where the *reservation-wage function* $\bar{w}(\pi)$ depends on the belief π.[1]

2.3 Reservation-Wage Functional Equation (RWFE)

At the reservation wage, the worker is indifferent between accepting and rejecting. Setting $w = \bar{w}(\pi)$ in (3) and simplifying yields the *reservation-wage functional equation*

$$\bar{w}(\pi) \;=\; (1-\beta)c + \beta \int \max\big\{w',\, \bar{w}\big(\kappa(w', \pi)\big)\big\}\; q_\pi(w')\, dw'. \tag{4}$$

Define the operator Q on bounded functions $\omega : [0,1] \to \mathbb{R}$ by

$$(Q\omega)(\pi) := (1-\beta)c + \beta \int \max\big\{w',\, \omega\big(\kappa(w', \pi)\big)\big\}\, q_\pi(w')\, dw'. \tag{5}$$

Proposition 1 (Contraction)**.** *Q is a contraction on $(\mathcal{B}([0,1]), \|\cdot\|_\infty)$ with modulus β. Hence a unique fixed point $\bar{w}$ exists and is obtained by iterates $\omega_{k+1} = Q\omega_k$.*

Proof sketch. For ω_1, ω_2,

$$\big|(Q\omega_1)(\pi) - (Q\omega_2)(\pi)\big| \leq \beta \int \big|\max\{w', \omega_1(\kappa)\} - \max\{w', \omega_2(\kappa)\}\big|\, q_\pi(w')\, dw' \leq \beta \|\omega_1 - \omega_2\|_\infty.$$

Taking the supremum over π gives the result.

3 Algorithms

We compare two solution strategies.

[1] If f is stochastically dominated by g, then $\bar{w}(\pi)$ is typically decreasing in π, reflecting faster acceptance when the worker believes the environment is worse.

3.1 Algorithm A: 2D Value Function Iteration (VFI)

State grid. Discretize wages $w \in [0, \bar{w}_{\max}]$ and beliefs $\pi \in [0,1]$.
Step. Apply the Bellman operator to $v(w_i, \pi_j)$:

$$(Tv)(w,\pi) = \max\left\{ \frac{w}{1-\beta},\ c + \beta\, \mathbb{E}\big[v(w', \kappa(w', \pi))\big] \right\},$$

approximating the integral via Monte Carlo (MC) draws from f and g mixed by π. Use interpolation for off-grid (w', π').

Algorithm 1 VFI with Belief State

Initialize $v_0(w_i, \pi_j)$ on a tensor grid
repeat
    for each π_j **do**
        for each w_i **do**
            Draw $\{w_f^m\}_{m=1}^M \sim f$, $\{w_g^m\}_{m=1}^M \sim g$
            $\pi_f^m \leftarrow \kappa(w_f^m, \pi_j), \quad \pi_g^m \leftarrow \kappa(w_g^m, \pi_j)$
            $\text{cont} \leftarrow \beta\Big[\pi_j \cdot \frac{1}{M}\sum_m v_k(w_f^m, \pi_f^m) + (1-\pi_j) \cdot \frac{1}{M}\sum_m v_k(w_g^m, \pi_g^m)\Big] + c$
            $v_{k+1}(w_i, \pi_j) \leftarrow \max\Big\{\frac{w_i}{1-\beta},\ \text{cont}\Big\}$
until $\|v_{k+1} - v_k\|_\infty < \text{tol}$
Policy: accept iff $\frac{w}{1-\beta} \geq c + \beta\, \mathbb{E}[v(w', \pi')]$.

Complexity. $O(|\mathcal{W}| \cdot |\Pi| \cdot M)$ per iteration; requires 2D interpolation and policy extraction.

3.2 Algorithm B: 1D Reservation-Wage Fixed Point (RWFE)

Iterate on the belief-only object $\omega(\pi)$ using the contraction Q in (5). No value grid, no policy extraction: the policy is $w \geq \bar{w}(\pi)$ with $\bar{w}$ returned by the fixed point.

Algorithm 2 Reservation-Wage Fixed Point via Q

Choose belief grid $\{\pi_j\}_{j=1}^J$; initialize $\omega_0(\pi_j)$
repeat
    for each π_j **do**
        Draw $\{w_f^m\}_{m=1}^M \sim f$, $\{w_g^m\}_{m=1}^M \sim g$
        $\pi_f^m \leftarrow \kappa(w_f^m, \pi_j), \quad \pi_g^m \leftarrow \kappa(w_g^m, \pi_j)$
        $\text{avg} \leftarrow \pi_j \cdot \frac{1}{M}\sum_m \max\{w_f^m, \omega_k(\pi_f^m)\} + (1-\pi_j) \cdot \frac{1}{M}\sum_m \max\{w_g^m, \omega_k(\pi_g^m)\}$
        $\omega_{k+1}(\pi_j) \leftarrow (1-\beta)c + \beta \cdot \text{avg}$
until $\|\omega_{k+1} - \omega_k\|_\infty < \text{tol}$
Return $\bar{w}(\pi_j) = \omega_{k+1}(\pi_j)$ (interpolate for off-grid π).

Complexity. $O(|\Pi| \cdot M)$ per iteration; only 1D interpolation in π. In practice this is substantially faster than VFI on a comparable accuracy target.

4 Implementation Plan

We adopt the code architecture in the companion notebook / lecture:

- `SearchProblem`: parameters, grids, MC pre-draws (w_f, w_g), density evaluators, and a jitted Bayes map κ.
- `operator_factory`: builds the Bellman operator T and a greedy policy extractor (for VFI).
- `Q_factory`: builds the Q-operator for RWFE. We use `numpy` for vectorization, `Numba` for JIT, and `np.interp` for 1D interpolation.

We standardize on RWFE for baseline runs and retain VFI for validation.

5 Validation and Experiments

5.1 Internal Checks

1. **Degenerate learning.** If $f = g$, learning is irrelevant; $\bar{w}(\pi)$ collapses to a constant equal to the McCall reservation wage.
2. **Monotonicity.** If f is stochastically dominated by g, verify $\bar{w}(\pi)$ decreases with π.
3. **Speed/accuracy.** Compare RWFE and VFI on a shared grid: sup-norm difference in induced policies and wall-clock runtimes.

5.2 Comparative Statics and Dynamics

- **Benefits c.** Evaluate $c \in \{0.1, 0.3, 0.8\}$ (or policy-relevant values) and trace shifts in $\bar{w}(\pi)$ and unemployment durations.
- **Informativeness.** Vary (f, g) to change likelihood-ratio variability (e.g., Beta families with matched means but different dispersion).
- **Beliefs at acceptance.** Plot the ECDF of π at stopping and the ECDF of durations under baseline vs. counterfactuals.

5.3 Aggregate Regime-Change Simulation

Simulate a large population; at a known date the true generator switches (e.g., from g to f). Document the temporary unemployment spike caused by slow learning and policy dependence (via c).

6 Data and Estimation

With micro data on offers and accept/reject decisions, estimate parameters of (f, g) and priors by MLE or SMM using the fixed-point policy $\bar{w}(\pi)$. Investigate offline policy evaluation under alternative c.

7 Expected Contributions

1. A clean characterization of search with learning via the 1D RWFE fixed point.
2. A fast and reproducible computation pipeline (RWFE) suitable for embedding in estimation or ABM contexts.
3. Quantified effects of benefits and informativeness on stopping, durations, and aggregate responses to regime changes.

References

- QuantEcon. "Job Search VII: Search with Learning." `https://python.quantecon.org/odu.html`.
- McCall, J. (1970). "Economics of Information and Job Search." *Quarterly Journal of Economics*, 84(1), 113–126.

Appendix

Full Bellman Construction of Baseline McCall Model

1. Environment and timing

Time is discrete, $t = 0, 1, 2, \ldots$, with discount factor $\beta \in (0, 1)$. Each period an unemployed worker observes a wage offer w_t drawn i.i.d. from a known offer distribution q (discrete or continuous). If the worker *accepts* an offer w, employment is permanent at wage w; if the worker *rejects*, she receives unemployment compensation $c \geq 0$ for the period and draws a new offer next period.

2. Value functions and the Bellman equation

Let $v^*(w)$ denote the value to an unemployed worker *after* seeing the current offer w.

Accept. If the worker accepts w today and employment is permanent, the present value is

$$V^A(w) = \frac{w}{1-\beta}.$$

Reject. If the worker rejects w today, she receives c now and draws $W' \sim q$ next period. Let $v^*(W')$ be the optimally continued value next period. Then

$$V^R \;=\; c + \beta\, \mathbb{E}\big[v^*(W')\big].$$

Bellman equation. Taking the maximum over the two actions yields

$$v^*(w) \;=\; \max\left\{\frac{w}{1-\beta},\; c + \beta\, \mathbb{E}\big[v^*(W')\big]\right\}. \tag{6}$$

When q is discrete on $\mathbb{W}$ with pmf $q(w)$, the expectation is $\mathbb{E}[v^*(W')] = \sum_{w' \in \mathbb{W}} v^*(w')\, q(w')$; in the continuous case with density also denoted q, $\mathbb{E}[v^*(W')] = \int v^*(w')\, q(w')\, dw'$.

3. Continuation value and reservation-wage policy

Define the *continuation value*

$$h \;:=\; c + \beta\, \mathbb{E}\big[v^*(W')\big]. \tag{7}$$

Then (6) becomes $v^*(w) = \max\{\frac{w}{1-\beta},\, h\}$. Hence the optimal policy is *threshold*: accept iff

$$\frac{w}{1-\beta} \geq h \qquad \Longleftrightarrow \qquad w \;\geq\; \bar{w}, \quad \text{where } \bar{w} := (1-\beta)\, h.$$

Thus the policy is summarized by the single number $\bar{w}$, the *reservation wage*.

4. Two fixed-point forms

(A) Scalar fixed point in h. Using $v^*(w) = \max\{\frac{w}{1-\beta}, h\}$ inside the definition (7) gives the one-dimensional contraction

$$h \;=\; c + \beta\, \mathbb{E}\left[\max\left\{\frac{W}{1-\beta},\, h\right\}\right]. \tag{8}$$

Define the operator $\Phi(h) := c + \beta\, \mathbb{E}[\max\{\frac{W}{1-\beta}, h\}]$. Then $|\Phi(h_1) - \Phi(h_2)| \leq \beta |h_1 - h_2|$, so h is the unique fixed point, obtainable by simple iteration $h^{(k+1)} = \Phi(h^{(k)})$.

(B) Scalar fixed point in $\bar{w}$. Equivalently, substituting $\bar{w} = (1-\beta)h$ into (8),

$$\bar{w} \;=\; (1-\beta)c \;+\; \beta\,\mathbb{E}\big[\max\{W,\,\bar{w}\}\big]. \tag{9}$$

Using $\max\{W, \bar{w}\} = \bar{w}\,\mathbf{1}\{W < \bar{w}\} + W\,\mathbf{1}\{W \geq \bar{w}\}$,

$$\bar{w}\,\big[1 - \beta F(\bar{w})\big] \;=\; (1-\beta)c \;+\; \beta\,\mathbb{E}\big[\,W \cdot \mathbf{1}\{W \geq \bar{w}\}\big]\,, \tag{10}$$

where F is the CDF of W. In the continuous case, $\mathbb{E}[W\,\mathbf{1}\{W \geq \bar{w}\}] = \int_{\bar{w}}^{\infty} w\,q(w)\,dw$.

5. Unemployment value and spell length

The value of being unemployed *before* seeing the offer equals the continuation value:

$$V_U \;=\; h \;=\; \frac{\bar{w}}{1-\beta}.$$

Under the reservation rule, the acceptance probability is $p^* = \Pr\{W \geq \bar{w}\} = 1 - F(\bar{w})$. With i.i.d. draws each period, the unemployment spell is geometric with mean $\mathbb{E}[\text{duration}] = 1/p^*$.

6. Computation (discrete grid version)

1. Choose β, c and a finite wage grid $\mathbb{W}$ with pmf $q(w)$.
2. Initialize $h^{(0)}$ (e.g. $h^{(0)} = \frac{c}{1-\beta}$).
3. Iterate the scalar map

$$h^{(k+1)} \;\leftarrow\; c + \beta \sum_{w \in \mathbb{W}} \max\left\{\frac{w}{1-\beta},\, h^{(k)}\right\} q(w)$$

   until $|h^{(k+1)} - h^{(k)}| < \varepsilon$.

4. Return $\bar{w} = (1-\beta)\,h^{(k+1)}$; accept iff $w \geq \bar{w}$.

Continuous case. Replace the sum by an integral, or approximate the integral by Monte Carlo samples from q.