Robust Sequential Search

Karl H. Schlag
University of Vienna†

Andriy Zapechelnyuk
University of Glasgow‡

February 13, 2016

Abstract

We consider the problem of sequential search with free recall and discounting. Performance of a search rule for a given prior is measured as the fraction of the maximal payoff under this prior. A search rule is robust if it ~~has~~ a high performance under every prior. Bayesian search rules are not robust as, while being optimal under one prior, they can perform very poorly under a different prior. We present a search rule that performs well after any search history under any prior with a given support. In each round the rule stops searching with a probability that is linear in the best previous offer.

1 Introduction

The paper revisits the problem of sequential search with free recall. Search may be for the lowest price, the best match for an employee or the best job offer. For example, consider an individual who receives job offers that arrive one by one. Each offer has a value that can be compared to the others. In each round the individual has to decide:

†Department of Economics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria. *E-mail:* karl.schlag@univie.ac.at.

‡*Corresponding author*. Adam Smith Business School, University of Glasgow, University Avenue, Glasgow G12 8QQ, UK. *E-mail:* azapech@gmail.com

to stop and select one of the previous offers, or to continue the search. Offers are discounted over time, which makes waiting for better offers costly. The opportunity cost of not accepting an offer is the only kind of search cost in our model.

In the classical formulation of the search problem with i.i.d. values the decision maker knows the environment, identified by the number of offers to come and the probability distribution according to which each offer drawn. This makes the solution simple, the optimal search rule is a stationary strategy that stipulates to wait for an offer above a certain cutoff (e.g., McCall 1970). At the same time, this formulation is quite unrealistic, as search is typically accompanied by uncertainty about what outcomes are even possible. This concern has been addressed in the classic Bayesian approach in which a decision maker forms beliefs about possible environments (e.g., Rothschild 1974). This approach is often unrealistic as well. Numerous questions arise. Imagine a shopper visiting one shop after another in search for a particular product. How could the shopper know the probability that the next shop will have the desired product? How would she form a prior about such probabilities? Even if a prior is formed, will the shopper be able to derive an optimal rule? Is she willing to derive the optimal rule for each product from scratch? It would be nice if she could reuse rules, if a rule that is optimal for a specific prior also performs well for other priors. However, we show that quite the opposite is true. A Bayesian optimal rule for one prior can perform very badly when evaluated under a different prior.

In this paper we consider a decision maker who does not know the true environment and who does not want to base the choice of her rule on specific beliefs. So we do not search for a rule that is best for a specific prior. Instead we search for a rule that does well for any prior. In this sense it will be robust. One might say that it is universal as it does not have to be adapted to a particular environment or to a new information as it arrives. Such a rule can be proposed as a compromise among Bayesians who have different priors. It is a short cut to avoid cumbersome calculations of the Bayesian optimal rule. As a universal rule it is a useful benchmark for empirical studies. Understanding how it performs in the least favorable environment will provide a bound on how valuable it is to gather more information.

All we assume about the environment is that outcomes belong to a specified interval. This interval of outcomes results from the description of the problem. For instance, the set of prices at which a good can be offered is usually bounded. The only other input for designing the rule, besides the interval of possible outcomes, is the discount factor of a decision maker. The discount factor can model the degree of the decision maker's impatience. It can capture a constant decay of offers that are not accepted. It can capture an exogenous probability that the search terminates with no previous offers to recall.

The performance of a candidate rule will be measured as under the competitive ratio (Sleator and Tarjan 1985). One imagines a Bayesian who knocks on the door, presents her prior and her optimal expected payoff under that prior, and then demands to know how well the candidate rule performs relative to it. The competitive ratio of the rule is the lower bound on this relative performance across all priors. It is the fraction of the Bayesian optimal payoff that the rule guarantees for every Bayesian prior. As this comparison includes a Bayesian who knows the true environment, the competitive ratio measures the maximal relative loss of not knowing the truth.

There is an additional innovative aspect to our approach. In the spirit of Bayesian decision making we not only consider relative performance ex-ante, before any new information arrives, but also evaluate it after each additional bit of information has been gathered. In other words, the rule is compared to the Bayesian optimal performance for every prior and at every stage of the search.

The value of deriving a rule based on a worst case analysis with so little assumptions only makes sense if the resulting bounds are not too large. The reader should judge for oneself. We present a rule such that, for instance, if the best previous offer is at least $1/6$ of the range of possible offers, then this rule guarantees a payoff that is at least $2/3$ of any Bayesian, and hence of the maximal possible expected payoff in the true environment. Just for comparison, for every Bayesian rule without outside option, the guaranteed fraction to any other Bayesian is zero.

We derive the optimal rule under the condition that the initial value or the best previous offer is above $1/6$. This rather complex rule guarantees in each round the highest performance bound, anticipating that the same rule is also used in all future rounds. We also identify a simple linear rule that is almost as good. A key insight is to only accept with certainty those offers that exceed the discount factor. Randomize for all other lower offers, the acceptance probability being increasing in the best previous offer.

This randomization allows the decision maker not to miss out on lost opportunities, such as stopping to search too early, and thus to avoid a low performance ratio.

There is an extensive literature on decision making without priors, sometimes also referred to as robust control. Almost all of the papers consider a decision maker who is able to commit to a rule. The specific application of this paper, search without priors, has been investigated by Bergemann and Schlag (2011) and Parakhonyak and Sobolev (2015), albeit under commitment. Papers that analyze dynamically consistent behavior under multiple priors include Baliga, Hanany and Klibanoff (2013), Schlag and Zapechelnyuk (2015), and Beissner, Lin and Riedel (2016). The methodology for decision making with multiple priors used in this paper is taken from Schlag and Zapechelnyuk (2015). There are several advantages of this method as compared to the others used in the literature. It is easy to explain. It is simple to evaluate (note that the complications in this paper stem from the fact that the set of priors is extremely rich so that one cannot proceed by computing the Bayesian optimal solution for each prior). It was designed to be as close as possible to the Bayesian (i.e., subjective expected utility) framework. In particular, it provides a solution for a Bayesian where optimal solutions are too cumbersome to compute. Moreover, it reveals a bound on the value to a Bayesian of gathering more information.

2 Model

2.1 Preliminaries

An individual searches for a job. Offers with i.i.d. values arrive one by one in discrete times $t = 1, 2, ...$ At each stage the individual decides whether to stop the search or to wait for another offer. With every offer all values decay by a discount factor. There is *free recall*: when the search is stopped, the individual picks the highest-valued offer arrived thus far.

The individual has an outside option with value $a > 0$ that can be chosen any time. Let $x_1, x_2, ..., x_t$ denote the realizations of t offers. The value of the best available offer at each $t = 0, 1, ...$ will be called the *search value at stage* t and denoted by y_t:

$$y_t = \max\{a, x_1, ..., x_t\}.$$

If the individual stops after t draws, then her payoff is the discounted value, $\delta^t y_t$, where $\delta \in (0,1)$ is a discount factor. The search ends after the individual decides to stop, or after all offers have been received. The total number of offers is denoted by n.

The individual knows her discount factor δ and the outside option value a. We assume that

$$0 < a < \delta < 1.$$

A positive outside option, $a > 0$, together with impatience of the individual, $\delta < 1$, implies that search is costly, as every new offer costs the individual at least the delayed consumption of the value a. Also, we assume $a < \delta$, as otherwise it makes no sense for anyone to search, the search problem then being trivial.

The individual knows neither the number of offers n, nor the distribution of their values F. A pair (n, F) is called the *environment*. We consider the class of environments, denoted by $\mathcal{E}$, whose distributions have a bounded support (known to the individual),

$$\mathcal{E} = \{(n, F) \ : \ n \in \mathbb{N} \cup \{\infty\}, \ supp(F) \subset [0,1]\}.$$

Environments with $n = \infty$ describe infinite i.i.d sequences. The normalization of the support bounds to $[0,1]$ interval is without loss of generality.

A *search rule* p prescribes for each stage $t = 0, 1, 2, ...$ and each history of draws $h_t = (x_1, ..., x_t)$ the probability $p(h_t)$ of stopping at that stage. Since the individual is not aware of the environment $E = (n, F)$, a search rule of an individual must not depend on either n or F, but it may depend on the individual's beliefs about (n, F).

Note that even though the individual does not know n, the search automatically stops if all n offers are received. This is as if the individual discovers after n draws that there are no more offers left, and hence stops the search.[1]

2.2 Robustness to Bayesian evaluation

A Bayesian model specifies some prior beliefs over environments in $\mathcal{E}$ at the outset. At each stage a Bayesian decision maker updates these beliefs by Bayes' rule and takes the

[1]The results do not change if the individual discovers n some periods in advance, since the results are driven by the case of $n = \infty$.

optimal decision w.r.t. the updated beliefs: stop the search if and only if the current search value is at least as high as the expected continuation payoff under these beliefs.

There are well known problems of the Bayesian approach to search. First, it may be difficult to justify a choice of prior beliefs. How should one form beliefs about the distribution of values of a chandelier in the next antique shop? Second, Bayesian updating can be tedious and even intractable when dimensionality of the model is large. Third, in an event which is impossible under one's beliefs, posterior beliefs are indeterminate and Bayesian approach gives no prescription what to do. Finally, it may be difficult for one Bayesian individual to justify her actions to another Bayesian who has different beliefs.

In this paper we take a different approach: we seek simple search rules that do not require formation and updating of beliefs, and yet perform relatively well in the eyes of every Bayesian, no matter what beliefs he or she has. In that sense, we are interested in *robust* rules that can be justified to Bayesian decision makers with various priors.

Before specifying the criterion of performance evaluation of a search rule, we need to formally define the performance of Bayesian models which will serve as benchmarks.

A Bayesian prior μ is a discrete probability measure on $\mathcal{E}$, that is, its support $supp(\mu)$ is at most countable collection of environments such that $\sum_{E\in supp(\mu)} \mu(E) = 1$. Denote by $\mathcal{M}$ the class of such discrete probability measures. For every history of draws h_t denote by $\mu(\cdot|h_t)$ the posterior after h_t. Note that the posteriors, whenever exist, are also discrete probability measures in $\mathcal{M}$.

For every environment $E = (n, F) \in \mathcal{E}$, every period $t < n$, and every history of draws $h_t = (x_1, ..., x_t)$, we denote by $U_t^p(E, h_t)$ the expected payoff from a given search rule p conditional on history of draws h_t. For every finite n it is defined by backward induction, $U_n^p(E, h_n) = y_n$, and for every $t < n$

$$U_t^p(E, h_t) = p(h_t)y_t + (1 - p(h_t))\delta \mathrm{E}_F\left[U_{t+1}^p(E, h_{t+1})|h_t\right],$$

where $\mathrm{E}_F\left[\,\cdot\,|h_t\right]$ denotes the expectation in the next stage under F conditional on current history h_t. The payoff of rule p under a prior μ at stage t and history h_t is given by

$$U_t^p(\mu, h_t) = \sum_{E\in supp(\mu(\cdot|h_t))} U_t^p(E, h_t)\mu(E|h_t).$$

The Bayesian optimal payoff under a prior μ is the maximal payoff:

$$V_t(\mu, h_t) = \sup_p U_t^p(\mu, h_t).$$

A measure μ is called *Dirac measure* if it places the unit mass on a single environment $E = (n, F)$ in which F has full support on $[0, 1]$. The subclass of Dirac measures describes Bayesian decision makers who are certain about the true environment. We will refer to such Bayesians as *experts*. The expert's optimal rule stipulates to stop with certainty whenever search value y_t is at least as large as some cutoff $\bar{c}$ such that the expert is indifferent between stopping and grabbing that search value and continuing. The cutoff $\bar{c}$ is the unique solution of the equation

$$\bar{c} = \delta \left(F(\bar{c})\bar{c} + \int_{\bar{c}}^{1} x \mathrm{d}F(x) \right), \tag{1}$$

and it is independent of n (see, e.g., McCall 1970). The expert's payoff,

$$V_t(E, h_t) = \sup_p U^p(E, h_t),$$

is computable by backward induction for $n < \infty$ and via Bellman equation for $n = \infty$.

2.3 Performance criterion

The performance of a search rule p in the eyes of a Bayesian with prior μ is evaluated as the ratio of p's payoff to the Bayesian optimal payoff in this environment, $\frac{U_t^p(\mu, h_t)}{V_t(\mu, h_t)}$.

We are interested in rules that perform well in the eyes of all Bayesians. The minimal performance ratio across all Bayesians is called the *competitive ratio of decision rule p*,

$$R_t^p(h_t) = \inf_{\mu \in \mathcal{M}} \frac{U_t^p(\mu, h_t)}{V_t(\mu, h_t)}.$$

In fact, since

$$\frac{U_t^p(\mu, h_t)}{V_t(\mu, h_t)} \geq \inf_{E \in supp(\mu(\cdot|h_t))} \frac{U_t^p(E, h_t)}{V_t(E, h_t)},$$

it is sufficient to consider only Dirac measures,

$$R_t^p(h_t) = \inf_{E \in \mathcal{E}(h_t)} \frac{U_t^p(E, h_t)}{V_t(E, h_t)},$$

where $\mathcal{E}(h_t)$ denotes the set of environments whose distributions include the historical realizations $\{x_1, ..., x_t\}$ in their support:

$$\mathcal{E}(h_t) = \{(n - t, F) \in \mathcal{E} \,:\, n > t,\, \{x_1, ..., x_t\} \subset supp(F)\}.$$

When we talk about *optimal performance*, we use the concept of *sequential optimality* (or subgame perfection) which requires the individual to optimize the competitive ratio at every stage of the search while expecting this optimization to occur at all future stages.

Formally, we say that *search rule q is an improvement over search rule p at history h_t* if

$$R_t^q(h_t) > R_t^p(h_t)$$

and

$$R_k^q(h_k) \geq R_k^p(h_k) \ \text{ for all } k > t \text{ and all } h_k \text{ consistent with } h_t.$$

A search rule p is said to be *sequentially optimal* if there does not exist an improvement over that rule for any history.

3 Learning and Randomization

One can guarantee the *trivial ratio*, a/δ, by simply taking the outside option a, where the worst environment deterministically gives the upper bound value 1 in the first offer. The question of interest is if one can attain a better competitive ratio.

We now establish two important properties of search under the competitive ratio performance criterion.

First, learning from the history (i.e., inference from past observations) is useless. As a result, without any loss of optimality one can restrict attention to search rules that ignore all past information except the search value.

Second, randomness of decisions is essential to obtain any competitive ratio better than the trivial one. Deterministic search rules (which prescribe to either surely stop or surely continue) can guarantee only the trivial ratio.

3.1 Learning

Stationary search rules are those that ignore the history and do not learn from past observations. Here we show restriction to stationary search rules is without loss of optimality, and thus learning is useless in this problem.

A search rule p is called *stationary* if it is a function of the search value only. That is, $p(h_s) = p(\hat{h}_t)$ for any two histories $h_s = (x_1, ..., x_s)$ and $\hat{h}_t = (\hat{x}_1, ..., \hat{x}_t)$ with the same search value, $\max\{a, x_1, ..., x_s\} = \max\{a, \hat{x}_1, ..., \hat{x}_t\}$.

We say that two search rules p and q are *payoff-equivalent* if they have the same competitive ratio for all histories:

$$R_t^p(h_t) = R_t^q(h_t) \text{ for every stage } t \text{ and every history } h_t.$$

Proposition 1. *There exists a sequentially optimal search rule which is stationary. All sequentially optimal search rules are payoff-equivalent to that rule.*

The proof is in the Appendix. Here we outline its intuition. At the outset, the individual only knows that she is facing an environment in the class $\mathcal{E}$: any number of alternatives in $\mathbb{N}$ that are i.i.d. distributed on $[0, 1]$. After having observed t realizations, $h_t = (x_1, ..., x_t)$, the number of remaining alternatives is still any number in $\mathbb{N}$. Moreover, all i.i.d. distributions are ex-ante possible, including those under which the observed history h_t is an arbitrarily unlikely event. The closure of the set of distributions which generate history h_t with a positive probability is the same as the ex-ante set of distributions. Consequently, there is no history after which the individual can narrow down the set of environments. After every history she faces the original set of environments, $\mathcal{E}$. This feature of the problem implies that nothing can be learnt from the past. Restriction to decision rules that ignore the history and do not learn from past observations is without loss of optimality.

3.2 Randomization

Here we show that deterministic search rules can guarantee at most the trivial competitive ratio, a/δ. Consequently, randomization is essential for attaining any nontrivial performance. A search rule that attains a better competitive ratio must randomize between stopping and continuing in some contingencies.

The *guaranteed competitive ratio of decision rule p* is the minimal competitive ratio across all stages and all histories, including the empty history at the outset:

$$\bar{R}^p = \inf_{t,h_t} R_t^p(h_t).$$

A search rule p is called *deterministic* if $p(h_t) \in \{0,1\}$ for every t and every history h_t.

Proposition 2. *For every deterministic search rule p,*

$$\bar{R}^p \leq \frac{a}{\delta}.$$

The intuition behind the result is straightforward. A deterministic behavior pattern can be exploited against the individual. The worst-case environment will be the one where the individual surely stops when she should have continued if she knew the environment, or vice versa, the individual surely keeps searching when she should have stopped. The complete proof is in the Appendix.

3.3 Comparing Bayesian Models

Is there a Bayesian model with some beliefs that can be justified in front of Bayesian decision makers with different beliefs?

A search rule p is called *Bayesian* if it specifies some prior beliefs over environments in $\mathcal{E}$ at the outset, and prescribes the individual to update these beliefs after each stage by Bayes' rule and to take the optimal decision w.r.t. the updated beliefs: to stop the search if and only if the current search value is at least as high as the expected continuation payoff under these beliefs.

Observe that Bayesian rules feature *learning* and *determinism*. An individual "learns" from past observations by making inference and adjusting behavior. Also, every Bayesian rule is deterministic: either the search value is strictly smaller than the expected continuation payoff (so the individual surely continues) or not (so the individual surely stops).

Thus, as concerns performance evaluation of Bayesian rules in the eyes of other Bayesians, learning is useless by Proposition 1 and determinism prohibits attainment of any competitive ratio above the trivial one, a/δ, by Proposition 2. The following corollary is immediate.

Corollary 1. *For every Bayesian search rule p,*

$$\bar{R}^p \leq \frac{a}{\delta}.$$

4 Optimal Performance

4.1 Worst-case environments.

Let us narrow down the set of environments that may constitute the worst cases for search rules.

Proposition 1 allows us to focus on stationary search rules that depend on the search value only. In what follows, we consider only stationary rules.

With this restriction we can simplify notations as follows. We replace a history h_t by the correspondent search value y_t. In the new notations, for a given search value y, $p(y_t)$ denotes the probability of stopping and accepting value y_t,

$$V_t(E, y_t) \text{ and } U_t^p(E, y_t)$$

denote, respectively, the expert's payoff and the payoff of rule p under environment E at stage t, and the competitive ratio is

$$R^p(y_t) = \inf_{E \in \mathcal{E}} \frac{U_t^p(E, y_t)}{V_t(E, y_t)}.$$

Note that since the problem is stationary, the competitive ratio is independent of t. It depends only on the search value y_t.

Denote by $\mathcal{E}_2$ the binary-valued environments with infinite number of alternatives, $n = \infty$, whose values are i.i.d. lotteries over a pair $(w, z) \in [0, 1]^2$, $w \leq z$. The following lemma shows that the class of binary-valued environments $\mathcal{E}_2$ contains the worst-case environments for every stationary search rule.

Lemma 1. *Let p be a stationary search rule. Then, for every period t and every search value $y_t > 0$,*

$$\frac{U_t^p(E, y_t)}{V_t(E, y_t)} = R^p(y_t) \quad \textit{implies} \quad E \in \mathcal{E}_2.$$

4.2 Sequentially optimal performance.

A sequentially optimal rule is very complex, and finding it is a tedious task. In this paper we focus instead on simple search rules (presented in the next subsection) whose competitive ratio approximates the sequentially optimal one.

Nevertheless, we find a partial solution for the sequentially optimal performance, to which the performance of simple search rules will be compared.

Here we present two optimal performance results. First, we establish a lower bound on the best competitive ratio. This lower bound is tight and equal to $\frac{1}{4}$ as $y \to 0$, so the individual has virtually nothing at her hand.

Proposition 3.

$$\lim_{y \to 0} \left\{ \sup_{p \in \mathcal{P}_{stat}} R^p(y) \right\} = \frac{1}{4}.$$

It is straightforward to verify that the search rule $\bar{p}(x) = \frac{1-\delta}{2-\delta}$ for all $x \in [0, 1]$ delivers the competitive ratio $R^{\bar{p}}(y) \geq \frac{1}{4}$ for all $y > 0$. The proof of the converse, that $\lim_{y \to 0} R^p(y) \leq \frac{1}{4}$ for every stationary strategy p, is more involved. We find the best rule p against environments that randomize between two values, y and z, with some probabilities $1 - \sigma$ and σ, respectively. We consider the limit of both y and z approaching zero, while the ratio z/y gets arbitrarily large. If the individual stops after observing y, then she might forgo a likely z in the next period, with the performance ratio approaching zero as $z/y \to \infty$. Otherwise, if the individual continues after observing y, then it might be the case that z is extremely unlikely to appear, so the expert stops and grabs y while the individual keeps waiting, again, performance ratio being arbitrarily small. Thus, randomization between stopping and continuing is necessary. The optimal limit probability of stopping, as $y \to 0$, turns out to be $\frac{1-\delta}{2-\delta}$, which delivers the limit competitive ratio 1/4. The complete proof is in the Appendix.

Next, we derive a sequentially optimal rule and a corresponding competitive ratio under the assumption that the outside value satisfies $a \geq \frac{1}{6}$.

Proposition 4. *For every subgame with the search value at least $\frac{1}{6}$, the rule*

$$p^*(x) = \begin{cases} \frac{2(1-\delta)}{4-2\delta+x-\sqrt{x(x+8)}}, & \text{if } x \leq \frac{\delta^2}{2-\delta} \\ \frac{1-K(\delta,x)+\sqrt{K^2(\delta,x)-1}}{1-x/\delta}, & \text{if } \frac{\delta^2}{2-\delta} < x < \delta \\ 1, & \text{if } x \geq \delta, \end{cases} \tag{2}$$

is sequentially optimal, where $K(\delta, x) = \frac{1}{\delta}\left(1 - \frac{(1-\delta)x}{2\delta}\right)$. The competitive ratio of this rule is equal to

$$R^*(y) = \begin{cases} \frac{1}{2} + \frac{1}{8}\left(y + \sqrt{y(y+8)}\right), & if \quad \frac{1}{6} \leq y \leq \frac{\delta^2}{2-\delta}, \\ K(\delta, y) - \sqrt{K^2(\delta, y) - 1}, & if \quad \frac{\delta^2}{2-\delta} < y < \delta, \\ 1, & if \quad y \geq \delta. \end{cases}$$

The proof of Proposition 4 proceeds as follows. In simple environments it is simple to find the tight upper bound on the ratio that the decision maker can guarantee if she can commit to a rule. The simple environment we consider is one where in each period nature randomizes equally likely between the best current offer y and the maximal payoff 1. This determines the acceptance probability for y. We postulate this solution for all values of the best offer. So, following Lemma 1, the worst case scenario is that nature puts weight only on two values, w and z. Part of the proof then involves showing that $w = y$ and $z = 1$. We are then able to show that $z = 1$ holds for this worst case when $y \geq 1/6$.

Note that the need to differentiate the two cases, whether y is larger or smaller than $\frac{\delta^2}{2-\delta}$, has the following reason. When $y > \frac{\delta^2}{2-\delta}$ then the constraint that nature can put at most mass 1 on $z = 1$ is binding.

The complete proof of Proposition (4) is in the Appendix.

We thus know that the best competitive ratio approaches $1/4$ as the search value tends to zero, and it is $R^*(y)$ for the range of search values $[1/6, 1]$. We do not know what the best competitive ratio is for $y \in (0, 1/6)$. However, condition $y \geq 1/6$ is sufficient for $R^*(y)$ being optimal, but not necessary. Denote by $\bar{y}(\delta)$ the exact lower bound on y at which $R^*(y)$ becomes optimal. Numerical evaluations shows that $\bar{y}(\delta)$ is strictly below 1/6 and is decreasing with the discount factor, and $\bar{y}(\delta) \to 1/6$ as $\delta \to 1$. The table below shows the numerical evaluations of $\bar{y}(\delta)$ for some values of δ.

δ	0.9	0.95	0.99	0.999
$\bar{y}(\delta)$	0.137	0.151	0.163	0.166

Quite surprisingly, the value of the competitive ratio is independent of the discount factor as long as the search value is not too large, $y \leq \frac{\delta^2}{2-\delta}$. Intuitively, since we evaluate the rule's performance by the fraction of the maximal payoff, under the sequentially optimal rule, a more patient individual simply waits longer in expectation, exactly to the extent that offsets the benefit of a cheaper waiting time.

The table below illustrates some values of the competitive ratio for the optimal rule when the individual is not too impatient. It shows for each best offer y what fraction of the maximal payoff can be guaranteed under any prior, provided $y \leq \frac{\delta^2}{2-\delta}$ (note that otherwise the competitive ratio is even larger).

Best offer y	1/6	1/5	1/4	1/3	1/2
Competitive ratio $R^*(y)$	0.666	0.685	0.71	0.75	0.82

4.3 Linear Decision Rules

In this section we consider the performance of linear rules. We wish to investigate how well they can perform as compared to our benchmark in Proposition 4. In particular we will gather information on guaranteed performance when the current search value is low, so for $y < 1/6$.

We note that the functional form of our sequentially optimal rule is in some sense simple. It is history dependent apart from depending on the best past offer. On the other hand the specific form looks complex, looking at the graph it does not seem to be easy to approximate with simpler functions. In this section we seek to investigate how simpler, linear rules perform. In a linear rule, the probability of stopping the search, is linear in the search value (and truncated at 1). We compare these rules with the above benchmark and investigate the percentage of the competitive ratio that is lost due to this simpler form. We will also making this comparison for small search values, $y < 1/6$, where we do not know how tight the bound given in Proposition 4 is.

The linear rules we investigate are denoted by $p_L(y|b)$,

$$p_L(y|b) = \min\left\{\tfrac{1-\delta}{2-\delta} + by, 1\right\},$$

where $b > 0$ is a parameter to be determined. Note that the intercept $\frac{1-\delta}{2-\delta}$ is taken from our analysis of the case of $y \to 0$ in Proposition 3. It is the acceptance probability of a job when the individual's outside option value tends to zero. By choosing this intercept we approximate the best competitive ratio for very small search values.

For investigating the competitive ratio of such a linear rule, by Lemma 1 we can limit attention to nature generating i.i.d. outcomes with only two point masses w and z, and using derivatives we show that the ratio is minimal for a given z when $w = y$ where y is the search value.

Given these two properties, for each value of b and y we can derive the minimal ratio by searching for the probability σ put on mass z and the value of z. We also restrict attention to the range of search values for which we know the sequentially optimal competitive ratio, $y \geq \bar{y}(\delta)$.

For each value of δ and b we evaluate

$$\phi(b) = \max_{y \in [\bar{y}(\delta),1]} \frac{R^*(y) - R^{p_L(\cdot|b)}(y)}{R^*(y)}$$

and search for the value b^* that maximizes the above expression,

$$\phi^* = \phi(b^*) = \max_{b \geq 0} \phi(b).$$

So we look for linear rules that generate the smallest maximal loss as compared to our sequentially optimal rule. The value of ϕ^* measures maximum performance loss relative to the sequentially optimal rule,

$$R^{p_L(\cdot|b^*)}(y) \geq (1-\phi^*)R^*(y) \quad \text{for all } y \geq \bar{y}(\delta).$$

We then investigate how small y can be such that this bound remains to be true. A numerical evaluation of this bound will be denoted by $\underline{y}(\delta)$.

In the table below we present for various values of δ the numerically best linear rule coefficient b^*, the bound on performance loss of that rule relative to $R^*(y)$, and the lower bound $\underline{y}(\delta)$ on the interval search values for which this performance loss is not exceeded.

δ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95	0.99	0.999
b^*	4.68	2.48	1.75	1.38	1.19	1.05	0.93	0.81	0.6	0.39	0.1	0.01
ϕ^*	4.9%	4.7%	4.6%	4.5%	4.8%	5%	4.8%	4.4%	5.5%	6.6%	8.1%	8.3%
$\underline{y}(\delta)$	0.01	0.01	0.01	0.02	0.05	0.06	0.05	0.03	0.02	0.02	0.01	0.01

Note that the value of b^* is such that $p_L(\delta|b^*) = \frac{1-\delta}{2-\delta} + b^* \cdot \delta \approx 0.95$. Note also that we have no way of understanding how good the performance of these linear rules is for extremely low but strictly positive search values. Given our choice of the constant we know that they cannot be beaten as $y \to 0$, and we know that they only lose minimally as compared to the upper bound when y is not too small. So there is only an extremely small part of the space of outcomes where we have no benchmark to compare their performance too.

One may not be satisfied by the performance of the linear rules when δ is large. For large δ there is a different simple rule that performs almost as well as the sequentially optimal rule. Let

$$p_s(y|\alpha) = \min\left\{\sqrt{\frac{\alpha(1-\delta)y}{1-y}}, 1\right\}$$

for $y < \delta$ and $p_s(y|\alpha) = 1$ for $y \geq \delta$, where $\alpha > 0$ is a parameter. Again we are searching for the parameter α^* that minimizes the maximal relative loss in efficiency, ϕ^*. Here we limit attention to search values y above $\bar{y}(\delta)$. We list the values of α^* in the table below.

δ	0.9	0.95	0.99	0.999
α^*	1.35	0.8	0.22	0.024
ϕ^*	2.8%	1.6%	2.5%	3%
$\bar{y}(\delta)$	0.137	0.151	0.163	0.166

Appendix

Proof of Proposition 1. Fix any history $h_t = (x_1, ..., x_t)$, and suppose that there remain alternatives to draw, $n > t$. The set of environments $\mathcal{E}(h_t)$ that the individual faces from stage $t+1$ on is obtained by elimination from $\mathcal{E}$ the environments whose distributions do not support the historical values $x_1, ..., x_t$. (Note that the set of possible numbers of remaining alternatives after period t is the same as at the start, since $\{n - t : n \in \mathbb{N}, n > t\} = \mathbb{N}$.) Consequently, $Closure(\mathcal{E}(h_t)) = \mathcal{E}$. The supremum of the performance ratio is the same whether the domain is $\mathcal{E}(h_t)$ or $Closure(\mathcal{E}(h_t)) = \mathcal{E}$.

We thus have obtained that the set of environments is irreducible after any history of observations. At every stage t the only payoff relevant variable is the search value y_t, any other information about past observations is payoff irrelevant. We show now that the sequentially optimal performance ratio at any stage is history independent, it depends only on the search value at that stage.

Let p be a sequentially optimal decision rule. Consider two histories, $h_s = (x_1, ..., x_s)$ and $h_t = (\hat{x}_1, ..., \hat{x}_t)$ with the same search value, $y_s = y_t = \hat{y}$. That is, $\hat{y} = \max\{a, x_1, ..., x_s\} = \max\{a, \hat{x}_1, ..., \hat{x}_t\}$. Suppose that $R^p(h_s) < R^p(h_t)$. Then we can construct an improvement over p at history h_s by decision rule $\hat{p}$ identical to p in all subgames except the one following history h_s. In that subgame we define $\hat{p}$ equal to p in the subgame following history h_t. This contradicts the assumption of sequential optimality of p.

As performance at each stage t can depend only on the search value y_t, without loss it can be attained by a stationary decision rule.

Finally, we show that all sequentially optimal rules are payoff-identical. Suppose that there exist two sequentially optimal rules p and q whose performance ratio differs at some history h_t, say, $R^q(h_t) > R^p_t(h_t)$. Then we can construct an improvement over p at history h_t by decision rule $\hat{p}$ identical to p in all subgames except the one following history h_t. In that subgame we define $\hat{p}$ equal to q in the same subgame. This contradicts the assumption of sequential optimality of p. ∎

Proof of Proposition 2. Denote by E_σ the environment with infinite number of alternatives, $n = \infty$, whose values are i.i.d. lotteries between 0 and 1 with probabilities

$1 - \sigma$ and σ, respectively. Let $\hat{\mathcal{E}}$ be the set of such environments with nondegenerate lotteries:

$$\hat{\mathcal{E}} = \{E_\sigma : \sigma \in (0,1)\}.$$

Let p be a deterministic rule. Suppose that p stipulates to stop and take the outside alternative after k zero-valued observations, $h_k^0 = (x_1, ..., x_k) = (0, ..., 0)$, for some $k \geq 0$, and hence

$$U_k^p(E, h_k^0) = a.$$

Every environment in $\hat{\mathcal{E}}$ is consistent with this history, as h_k^0 has a nonzero probability of realization under every $E_\sigma \in \hat{\mathcal{E}}$. The first-best payoff under $E_\sigma \in \hat{\mathcal{E}}$ with σ close enough to 1 is

$$V_k(E_\sigma, h_k^0) = \delta(\sigma + (1-\sigma)V_{k+1}(h_{k+1}^0)) = ... = \frac{\delta\sigma}{1 - \delta(1-\sigma)}.$$

Hence,

$$R_k^p(h_k) \leq \inf_{E_\sigma \in \hat{\mathcal{E}}} \frac{U_k^p(E_\sigma, h_k^0)}{V_k(E_\sigma, h_k^0)} \leq \lim_{\sigma \to 1} \frac{a}{\frac{\delta\sigma}{1-\delta(1-\sigma)}} = \frac{a}{\delta}.$$

Now consider the complementary case, where p stipulates to continue after k zero-valued observations for all $k = 0, 1, 2, ...$ The first-best rule under $E_\sigma \in \hat{\mathcal{E}}$ with σ close enough to 0 would stipulate to stop immediately and yields at $t = 0$

$$V_0(E_\sigma) = a.$$

On the other hand, the payoff of p at $t = 0$ satisfies

$$U_0(E_\sigma) \leq \delta(\sigma + (1-\sigma)U_1(E_\sigma, h_1^0)) \leq ... \leq \frac{\delta\sigma}{1 - \delta(1-\sigma)}.$$

Hence,

$$R_0^p \leq \inf_{E_\sigma \in \hat{\mathcal{E}}} \frac{U_0^p(E_\sigma)}{V_0(E_\sigma)} \leq \lim_{\sigma \to 0} \frac{\frac{\delta\sigma}{1-\delta(1-\sigma)}}{a} = 0 \leq \frac{a}{\delta}.$$

■

Proof of Lemma 1. Consider any stationary rule p. Fix a period t and a search value $y > 0$ in that period.

Recall that the expert's optimal strategy is to stop whenever the search value exceeds some cutoff $\bar{c}$ and to continue otherwise. By (1), cutoff $\bar{c}$ must satisfy $\bar{c} = \delta\left(F(\bar{c})\bar{c} + \int_{\bar{c}}^{1} x\mathrm{d}F(x)\right)$, where F is the distribution. Using integration by parts, we can rewrite this condition as

$$\bar{c} = \delta\left(F(\bar{c})\bar{c} + \left(1 - \bar{c}F(\bar{c}) - \int_{\bar{c}}^{1} F(x)\mathrm{d}x\right)\right) = \delta\left(1 - \int_{\bar{c}}^{1} F(x)\mathrm{d}x\right),$$

or equivalently,

$$\int_{\bar{c}}^{1} F(x)\mathrm{d}x = 1 - \frac{\bar{c}}{\delta}. \tag{3}$$

Fix a cutoff $\bar{c} \in (0, \delta)$ and denote by $\mathcal{E}_{\bar{c}}$ the set of environments whose distributions have this cutoff,

$$\mathcal{E}_{\bar{c}} = \{(n, F) \in \mathcal{E} : n > t \text{ and } F \text{ satisfies (3)}\}.$$

Let $E = (n, F)$ be an environment in $\mathcal{E}_{\bar{c}}$. We now construct an environment $E' = (\infty, G) \in \mathcal{E}_2$ with the same cutoff $\bar{c}$, such that the performance ratio is smaller, $\frac{U_t^p(E,y)}{V_t(E,y)} \geq \frac{U_t^p(E',y)}{V_t(E',y)}$.

First, suppose that $y \geq \bar{c}$, so the expert's payoff is y. The worst-case payoff for rule p is attained if there is an infinite sequence of alternatives whose values never exceed y, in particular, the degenerate distribution that places probability 1 on value $\bar{c}$. This environment is in $\mathcal{E}_2$.

Next, suppose that $y < \bar{c}$, and hence the expert continues searching until finds a value above $\bar{c}$. Observe that the expert's payoff satisfies

$$\begin{aligned}
V_t(E, y) &= \delta\left(\int_{\bar{c}}^{1} x\mathrm{d}F(x) + F(\bar{c})V_{t+1}(E, y)\right) \\
&= \delta(1 + \delta F(\bar{c}) + ... + (\delta F(\bar{c}))^{n-t-1}) \int_{\bar{c}}^{1} x\mathrm{d}F(x) + (\delta F(\bar{c}))^{n-t} y \\
&\leq \delta \sum_{k=0}^{\infty} \delta^k F^k(\bar{c})) \int_{\bar{c}}^{1} x\mathrm{d}F(x) = \frac{\delta}{1 - \delta F(\bar{c})} \int_{\bar{c}}^{1} x\mathrm{d}F(x) = \bar{c}.
\end{aligned}$$

The individual's payoff is given by

$$U_t^p(E, y) = p(y)y + (1 - p(y))\delta \int_0^1 U_{t+1}^p(E, \max\{y, x\})\mathrm{d}F(x).$$

Imagine that the individual plays a zero-sum game with Nature. The individual's objective is to maximize her payoff, while Nature strives to minimize it. The individual chooses a strategy p, to which Nature responds by a choice of environment E in the set $\mathcal{E}_{\bar{c}}$.

Let us now make the individual's payoff smaller by having the individual to choose strategy p once and for all, but allowing Nature to choose a new environment at every stage, with the constraint that chosen environments must be in the set $\mathcal{E}_{\bar{c}}$. In this game, the individual's payoff at stage $t+1$ from strategy p and value $x_{t+1} = x$ is

$$\hat{U}^p(x) = \inf_{\hat{E} \in \mathcal{E}_{\bar{c}}} U^p_{t+1}(\hat{E}, \max\{y, x\}).$$

Observe that $\hat{U}^p(x)$ is independent of t, since p is stationary and $\mathcal{E}_{\bar{c}}$ does not change with time. By construction $U^p_{t+1}(E, x) \geq \hat{U}^p(x)$, hence

$$\begin{aligned} \int_0^1 U^p_{t+1}(E, \max\{y, x\}) \mathrm{d}F(x) &\geq \int_0^1 \hat{U}^p(x) \mathrm{d}F(x) \\ &= \int_0^{\bar{c}} \hat{U}^p(x) \mathrm{d}F(x) + \int_{\bar{c}}^1 \hat{U}^p(x) \mathrm{d}F(x). \end{aligned}$$

We now find a distribution that minimizes the right-hand side expression of the above inequality subject to (3) so the cutoff $\bar{c}$ remains constant,

$$\begin{aligned} &\min_G \int_0^{\bar{c}} \hat{U}^p(x) \mathrm{d}G(x) + \int_{\bar{c}}^1 \hat{U}^p(x) \mathrm{d}G(x) \\ &\text{s.t.} \int_{\bar{c}}^1 G(x) \mathrm{d}x = 1 - \frac{\bar{c}}{\delta}. \end{aligned}$$

A solution to the above problem is a distribution with two-point support $\{w, z\}$ and probabilities $1 - \sigma$ and σ, respectively, where w satisfies

$$w \in \arg\min_{x \in [0, \bar{c}]} \hat{U}^p(x),$$

and (σ, z) solve

$$\begin{aligned} &\min_{\sigma \in [0,1],\ z \in [\bar{c}, 1]} \left\{ (1-\sigma)\hat{U}^p(w) + \sigma \hat{U}^p(z) \right\} \\ &\text{s.t. } \bar{c} = \delta(\sigma z + (1-\sigma)\bar{c}), \end{aligned}$$

where the constraint is by (3).

Define environment $E' = (\infty, G)$ and observe that the minimum payoff in the above problem is attained under the environment E'. Also, notice that

$$V(E', y) = \bar{c},$$

since E' is the environment with cutoff $\bar{c}$ and infinite alternatives. We thus have $\frac{U_t^p(E,y)}{V_t(E,y)} \geq \frac{U_t^p(E',y)}{V_t(E',y)}$, and environment E' is in $\mathcal{E}_2$.

Thus we conclude that $\mathcal{E}_2$ must contain an environment that minimizes the ratio, $\frac{U_t^p(E,y)}{V_t(E,y)}$, among all environments in $\mathcal{E}$. ■

Proof of Proposition 3. First we show that $\lim_{y\to 0} R^p(y) \leq \frac{1}{4}$ for all stationary p. Denote

$$q = \limsup_{y\to 0} p(y)$$

and consider two decreasing sequences, $\{y_1, y_2, ...\}$ and $\{z_1, z_2, ...\}$, that converge to zero and satisfy

$$\lim_{k\to\infty} p(y_k) = \lim_{k\to\infty} p(z_k) = q \quad \text{and} \quad \lim_{k\to\infty} \frac{y_k}{z_k} = 0.$$

Fix $\sigma < 1/2$. For each k consider the environment that randomizes between y_k and z_k with probabilities $1-\sigma$ and σ, respectively. Denote such an environment by (y_k, z_k, σ). For every large enough k the ratio y_k/z_k is sufficiently small, so that $y_k < \frac{\delta\sigma z_k}{1-\delta(1-\sigma)}$, so that the expert waits for z_k to realize and obtains

$$V(y_k, z_k, \sigma) = \frac{\delta\sigma z_k}{1-\delta(1-\sigma)}.$$

The individual's payoff is

$$U^p(y_k, z_k, \sigma) = \frac{p(y_k)y_k + (1-p(y_k))\delta\sigma z_k \frac{p(z_k)}{1-\delta(1-p(z_k))}}{1-\delta(1-\sigma)(1-p(y_k))}$$

As we take the limit, $k \to \infty$, the ratio y_k/z_k approaches zero, while both $p(y_k)$ and $p(z_k)$ approach q. Hence we have

$$\frac{U^p(y_k, z_k, \sigma)}{V(y_k, z_k, \sigma)} = \frac{\left(p(y_k)\frac{y_k}{z_k} + (1-p(y_k))\delta\sigma\frac{p(z_k)}{1-\delta(1-p(z_k))}\right)(1-\delta(1-\sigma))}{\delta\sigma(1-\delta(1-\sigma)(1-p(y_k)))}$$

$$\to \frac{(1-q)\frac{q}{1-\delta(1-q)}(1-\delta(1-\sigma))}{1-\delta(1-\sigma)(1-q))} \quad \text{as } k \to \infty.$$

The above expression is increasing in σ, and hence

$$\inf_{\sigma\in[0,1]}\lim_{k\to\infty}\frac{U^p(y_k,z_k,\sigma)}{V(y_k,z_k,\sigma)}=\left.\frac{(1-q)\frac{q}{1-\delta(1-q)}(1-\delta(1-\sigma))}{1-\delta(1-\sigma)(1-q))}\right|_{\sigma=0}=\frac{q(1-q)(1-\delta)}{(1-\delta(1-q))^2}.$$

Thus we have obtained

$$\lim_{y\to0}R^p(y)\leq\lim_{k\to\infty}R^p(y_k)\leq\frac{q(1-q)(1-\delta)}{(1-\delta(1-q))^2}.$$

Since for $\delta\in(0,1)$

$$\max_{q\in[0,1]}\frac{q(1-q)(1-\delta)}{(1-\delta(1-q))^2}=\frac{1}{4},$$

it is immediate that $\lim_{y\to0}R^p(y)\leq\frac{1}{4}$.

We now prove the converse: that there exists a stationary rule p such that $\lim_{y\to0}R^p(y)\geq\frac{1}{4}$. Let

$$p(x)=q=\frac{1-\delta}{2-\delta},$$

independent of x. Observe that this strategy satisfies conditions of Lemma 1, and hence we only need to consider binary environments with $n=\infty$ that randomize between values w and z with probabilities $1-\sigma$ and σ, respectively. Denote any such environment by (w,z,σ).

First, suppose that the first-best rule dictates to stop immediately and obtain y. The individual's payoff in this case is increasing in both w and σ, hence consider $w=y$ and $\sigma=0$. Thus,

$$U^p(y,z,0)=\frac{qy}{1-\delta(1-q)}=\frac{\frac{1-\delta}{2-\delta}y}{1-\delta(1-\frac{1-\delta}{2-\delta})}=\frac{y}{2},$$

and the ratio is

$$\frac{U^p(y,z,0)}{V(y,z,0)}=\frac{y/2}{y}=\frac{1}{2}>\frac{1}{4}.$$

Second, suppose that the first-best rule dictates to continue until z realizes, so

$$V(w,z,\sigma)=\frac{\delta\sigma z}{1-\delta(1-\sigma)}.$$

The ratio is then increasing in w and decreasing in z, hence consider $w=y$ and $z=1$. Thus,

$$U^p(y,1,\sigma)=\frac{qy+(1-q)\frac{q\delta\sigma}{1-\delta(1-q)}}{1-\delta(1-q)(1-\sigma)}.$$

As y can be arbitrarily small, we have

$$\lim_{y\to 0}\frac{U^{p}(y,1,\sigma)}{V(y,1,\sigma)}=\frac{q(1-q)(1-\delta(1-\sigma))}{(1-\delta(1-q))(1-\delta(1-q)(1-\sigma))}.$$

This expression is increasing in σ, hence substituting $\sigma=0$ and $q=\frac{1-\delta}{2-\delta}$ we have

$$\lim_{y\to 0}\frac{U^{p}(y,1,\sigma)}{V(y,1,\sigma)}\geq\frac{q(1-q)(1-\delta)}{(1-\delta(1-q))^{2}}=\frac{1}{4}.$$

■

Proof of Proposition 4.

$$p^{*}(x)=\begin{cases}\frac{2(1-\delta)}{4-2\delta+x-\sqrt{x(x+8)}}, & \text{if } x\leq\frac{\delta^{2}}{2-\delta}\\ \frac{1-K(\delta,x)+\sqrt{K^{2}(\delta,x)-1}}{1-x/\delta}, & \text{if } \frac{\delta^{2}}{2-\delta}<x<\delta\\ 1, & \text{if } x\geq\delta,\end{cases}\tag{4}$$

is sequentially optimal, where $K(\delta,x)=\frac{1}{\delta}\left(1-\frac{(1-\delta)x}{2\delta}\right)$. The competitive ratio of this rule is equal to

$$R^{*}(y)=\begin{cases}\frac{1}{2}+\frac{1}{8}\left(y+\sqrt{y(y+8)}\right), & if\ \ \frac{1}{6}\leq y\leq\frac{\delta^{2}}{2-\delta},\\ K(\delta,y)-\sqrt{K^{2}(\delta,y)-1}, & if\ \ \frac{\delta^{2}}{2-\delta}<y<\delta,\\ 1, & if\ \ y\geq\delta.\end{cases}$$

The proof consists of two steps. On Step 1 we restrict attention to a very small set of environments and then find a sequentially optimal rule, p, w.r.t. this restricted set. On Step 2 we show the environments in the restricted set are actually the relevant worst-case environments within the general class of environments, $\mathcal{E}$, and hence the rule p is sequentially optimal on the general class.

Step 1. Fix a period t and a search value $y>1$ in that period. To simplify notations, we omit the reference to t.

Denote by E_{σ} the environment with $n=\infty$ that randomizes between two values, 0 and 1, with probabilities $1-\sigma$ and σ, respectively. Suppose that the set of feasible environments consists of only two environments: E_{0} and $E_{\sigma^{*}}$, where

$$\sigma^{*}=\begin{cases}\frac{3y+\sqrt{y(y+8)}}{2(1-y)}\frac{1-\delta}{\delta}, & \text{if } y\leq\frac{\delta^{2}}{2-\delta},\\ 1, & \text{if } y>\frac{\delta^{2}}{2-\delta}.\end{cases}$$

Note that σ^* is continuous w.r.t. y.

Let us find a search rule p that maximizes the competitive ratio w.r.t. these two environments. Notice that under these environments the search value will remain equal to y in all periods until the first realization of 1. Thus, there are only two possible values to consider: y and 1. Set

$$p(1) = 1 \quad \text{and} \quad p(y) = q.$$

That is, there is a single parameter $q \in [0,1]$ of the search rule that needs to be optimized.

If $y \geq \delta$, then the highest attainable payoff is y. The individual stops immediately, $q = 1$, and obtains the first-best payoff y.

In what follows we assume $y < \delta$.

Under environment E_0, the search value remains y forever. Thus the expert's payoff is $V(E_0, y) = y$, and the individual's payoff is

$$U^p(E_0, y) = qy + (1-q)\delta U^p(E_0, y) = \frac{qy}{1 - \delta(1-q)}.$$

Under environment E_{σ^*}, assuming $y < \delta$, the expert waits until the realization of 1 and obtains

$$V(E_{\sigma^*}, y) = \delta(\sigma^* + (1-\sigma^*)V(E_{\sigma^*}, y)) = \frac{\delta\sigma^*}{1 - \delta(1-\sigma^*)} > y.$$

The individual's payoff is equal to

$$U^p(E_{\sigma^*}, y) = qy + (1-q)\delta(\sigma^* + (1-\sigma^*)U^p(E_{\sigma^*}, y)) = \frac{qy + (1-q)\delta\sigma^*}{1 - \delta(1-q)(1-\sigma^*)}.$$

Hence

$$\min_{\sigma \in \{0, \sigma^*\}} \frac{U^p(E_\sigma, y)}{V(E_\sigma, y)} = \min\left\{ \frac{q}{1 - \delta(1-q)}, \frac{\frac{qy + (1-q)\delta\sigma^*}{1 - \delta(1-q)(1-\sigma^*)}}{\frac{\delta\sigma^*}{1 - \delta(1-\sigma^*)}} \right\}.$$

Observe that the two expressions in the curly brackets are monotonic in q in the opposite directions. Hence q must satisfy

$$\frac{q}{1 - \delta(1-q)} = \frac{\frac{qy + (1-q)\delta\sigma^*}{1 - \delta(1-q)(1-\sigma^*)}}{\frac{\delta\sigma^*}{1 - \delta(1-\sigma^*)}}. \tag{5}$$

Tedious but straightforward analysis, carried out separately for the two cases, $y \leq \frac{\delta^2}{2-\delta}$ and $\frac{\delta^2}{2-\delta} < y < \delta$, yields the unique solution on the domain $[0,1]$:

$$q^* = \begin{cases} \frac{2(1-\delta)}{4-2\delta+y-\sqrt{y(y+8)}}, & \text{if } y \leq \frac{\delta^2}{2-\delta} \\ \frac{1-K(\delta,y)+\sqrt{K^2(\delta,y)-1}}{1-y/\delta}, & \text{if } \frac{\delta^2}{2-\delta} < y < \delta, \end{cases}$$

where $K(y) = 1 - \frac{(1-\delta)y}{2\delta}$. Also recall that $q = 1$ for $y \geq \delta$. Substituting this solution into the ratio expression $\frac{q}{1-\delta(1-q)}$ yields

$$\min_{\sigma \in \{0,\sigma^*\}} \frac{U^p(E_\sigma, y)}{V(E_\sigma, y)} = \begin{cases} \frac{1}{2} + \frac{y}{8} + \sqrt{\frac{y}{8}\left(1+\frac{y}{8}\right)}, & \text{if } \frac{1}{6} \leq y \leq \frac{\delta^2}{2-\delta}, \\ \frac{1}{\delta}\left(K(y) - \sqrt{K^2(y)-\delta}\right), & \text{if } \frac{\delta^2}{2-\delta} < y < \delta, \\ 1, & \text{if } y \geq \delta. \end{cases}$$

Thus we have obtained that rule p^* defined by (4) attains the above competitive ratio w.r.t. the restricted set of environments $\{E_0, E_{\sigma^*}\}$.

Step 2. We now show that for every $y \geq \frac{1}{6}$ the pair of environments $\{E_0, E_{\sigma^*}\}$ are in fact the worst-case environments for the rule p^* defined by (4). That is, expanding the set of environments to $\mathcal{E}$ will not change the competitive ratio for the rule p^*.

Consider search rule

$$p^*(x) = \begin{cases} \frac{2(1-\delta)}{4-2\delta+x-\sqrt{x(x+8)}}, & \text{if } x \leq \frac{\delta^2}{2-\delta} \\ \frac{1-K(\delta,x)+\sqrt{K^2(\delta,x)-1}}{1-x/\delta}, & \text{if } \frac{\delta^2}{2-\delta} < x < \delta \\ 1, & \text{if } x \geq \delta, \end{cases}$$

We will prove that for every environment $E \in \mathcal{E}$ and $y \geq \frac{1}{6}$ this rule attains the ratio at least

$$R^*(y) = \begin{cases} \frac{1}{2} + \frac{1}{8}\left(y + \sqrt{y(y+8)}\right), & if \quad \frac{1}{6} \leq y \leq \frac{\delta^2}{2-\delta}, \\ K(\delta,y) - \sqrt{K^2(\delta,y)-1}, & if \quad \frac{\delta^2}{2-\delta} < y < \delta, \\ 1, & if \quad y \geq \delta. \end{cases}$$

By Lemma 1, without loss we consider the environments with $n = \infty$ that randomize between two values, w and z, $0 \leq w < z \leq 1$. Denote by $1-\sigma$ and σ the probabilities assigned to w and z, respectively. For short, denote any such environment by (w, z, σ).

First, consider any environment (w, z, σ) such that the expert prefers to stop immediately. The expert's payoff y, and it is the same as under the environment E_0:

$$V((w, z, \sigma), y) = V(E_0, y) = y.$$

Under this assumption, the individual's payoff is increasing in σ, thus minimized at $\sigma = 0$. When $\sigma = 0$, the payoff is independent of z, so we can set $z = 1$. Moreover, for $w \geq y$,

$$\begin{aligned}
U^{p^*}((w, 1, 0), y) &= p^*(y)y + (1 - p^*(y))\delta U^{p^*}((w, 1, 0), \max\{w, y\}) \\
&= p^*(y)y + (1 - p^*(y))\delta(p^*(w)w + (1 - p^*(w))\delta U^{p^*}((w, 1, 0), w) \\
&= p^*(y)y + (1 - p^*(y))\delta \frac{p^*(w)w}{1 - \delta(1 - p^*(w))} \\
&= p^*(y)y + (1 - p^*(y))\delta w R^*(w),
\end{aligned}$$

which is increasing in w. As the continuation payoff is the function of $\max\{w, y\}$, the payoff is minimized at $w = 0$. The environment $(w, z, \sigma) = (0, 1, 0)$ is the same as E_0. Thus we have

$$U^{p^*}((w, z, \sigma), y) \geq U^{p^*}(E_0, y).$$

So environment E_0 minimizes the ratio $\frac{U^{p^*}(E_0,y)}{V(E_0,y)}$ among all environments in which the expert stops immediately.

Next we consider all environments (w, z, σ) such that the expert searches until the first draw of z and obtains

$$V((w, z, \sigma), y) = \delta(\sigma z + (1 - \sigma)V(w, z, \sigma)) = \frac{\delta \sigma z}{1 - \delta(1 - \sigma)} = \bar{c} > y.$$

This is independent of w. Thus $V((0, z, \sigma), y) = V((w, z, \sigma), y)$.

Next, the individual's payoff is

$$\begin{aligned}
U^{p^*}((w, z, \sigma), y) = p^*(y)y + (1 - p^*(y))\delta \Big(&(1 - \sigma)U^{p^*}((w, z, \sigma), \max\{w, y\}) \\
&+ \sigma U^{p^*}((w, z, \sigma), z) \Big)
\end{aligned}$$

Recall that $w \leq \bar{c} < z$ and evaluate the next-period payoffs conditional on search values z and w,

$$\begin{aligned}
U^{p^*}((w, z, \sigma), z) &= p^*(z)z + (1 - p^*(z))\delta U^{p^*}((w, z, \sigma), z) \\
&= \frac{zp^*(z)}{1 - \delta(1 - p^*(z))} = zR^*(z),
\end{aligned}$$

which is constant in w, and

$$
\begin{aligned}
U^{p^*}((w,z,\sigma),w) &= p^*(w)w + (1-p^*(w))\delta\big((1-\sigma)U^{p^*}((w,z,\sigma),w) \\
&\quad + \sigma U^{p^*}((w,z,\sigma),z)\big) \\
&= \frac{p^*(w)w + (1-p^*(w))\delta\sigma z R^*(z)}{1-\delta(1-p^*(w))(1-\sigma)}.
\end{aligned}
$$

This expression is generally convex in w for $w > y$. Let $w^*(\sigma, z)$ minimize the above expression on interval $[0, \bar{c}]$ for a given (σ, z). We will now consider the case $y \geq w*$, thus $U^{p^*}((w,z,\sigma),y)$ is minimized at $w = 0$. Later we will show that if y is below $w^*(\sigma, z)$, then under the optimal choice of (σ, z) the performance ratio will be increasing in w, so the ratio can be reduced by lowering w.

So we consider $w = 0$ and denote $q = p^*(y)$ to simplify notations. Thus,

$$
\begin{aligned}
U^{p^*}((0,z,\sigma),y) &= qy + (1-q)\delta\Big((1-\sigma)U^{p^*}((0,z,\sigma),y) + \sigma U^{p^*}((0,z,\sigma),z)\Big) \\
&= \frac{qy + (1-q)\delta\sigma z R^*(z)}{1-\delta(1-q)(1-\sigma)}.
\end{aligned}
$$

Hence

$$
\begin{aligned}
\frac{U^p((0,z,\sigma)),y}{V((0,z,\sigma),y)} &= \frac{\frac{qy+(1-q)\delta\sigma z\bar{R}(z)}{1-\delta(1-q)(1-\sigma)}}{\frac{\delta\sigma z}{1-\delta(1-\sigma)}} \\
&= \frac{1-\delta(1-\sigma)}{1-\delta(1-q)(1-\sigma)}\left(q\frac{y}{\delta\sigma z} + (1-q)R^*(z)\right).
\end{aligned}
$$

We need to minimize the above expression w.r.t. z and σ. Define

$$
\bar{R}(x) = \frac{1}{8}\left(4 + x + \sqrt{x(x+8)}\right).
$$

Observe that $\bar{R}(x) \leq R^*(x)$ for all x and $\bar{R}(1) = R^*(1) = 1$. We now replace $R^*(z)$ with $\bar{R}(z)$ in the ratio expression and show that it is minimized at $z = 1$, thus showing that the original ratio with $R^*(z)$ is also minimized at $z = 1$.

For a given z, the obtained expression

$$
\frac{1-\delta(1-\sigma)}{1-\delta(1-q)(1-\sigma)}\left(q\frac{y}{\delta\sigma z} + (1-q)\bar{R}(z)\right) \tag{6}
$$

is quasiconvex in σ and is minimized at

$$\bar{\sigma}(z) = \frac{\delta(1-\delta)(1-q)a + \sqrt{\delta(1-\delta)qa((1-\delta(1-q))b - \delta(1-q)a)}}{\delta(qb - \delta(1-q)a)},$$

where $a = \frac{qy}{\delta z}$ and $b = (1-q)\bar{R}(z)$. Straightforward, but tedious analysis shows that (6) evaluated at $\sigma = \max\{\bar{\sigma}(z), 1\}$ is strictly decreasing in z for all $z < 1$, provided

$$\frac{1}{6} \leq y \leq \frac{\delta\sigma z}{1 - \delta(1-\sigma)}.$$

Also, observe that $\max\{\bar{\sigma}(z), 1\}$ evaluated at $z = 1$ is equal to σ^*. Consequently, the environment $E_{\sigma^*} = (0, 1, \sigma^*)$ minimizes the performance ratio,

$$\min_{z,\sigma} \frac{U^p((0,z,\sigma),y)}{V((0,z,\sigma),y)} = \frac{U^p((0,1,\sigma^*),y)}{V((0,1,\sigma^*),y)} = R^*(y).$$

It remains to show that an environment (w, z, σ) with $w > y$ never minimizes the ratio. Let $w > y$. Notice that the minimization problem w.r.t. (z, σ) has not changed, hence $(z, \sigma) = (1, \sigma^*)$ (with w instead of y in the expression for σ^*. But then the expression $\frac{U^p((w,1,\sigma^*),y)}{V((w,1,\sigma^*),y)}$ is increasing in w and minimized at $w = 0$.

We thus have shown that E_{σ^*} is the worst environment among those where the expert waits for z, and E_0 is the worst environment among those where the expert stops immediately. ■

References

Baliga, Sandeep, Eran Hanany, and Peter Klibanoff, "Polarization and Ambiguity," *American Economic Review*, 2013, *103*, 3071–3083.

Beissner, Patrick, Qian Lin, and Frank Riedel, "Dynamically Consistent α-Maxmin Expected Utility," 2016. Mimeo.

Bergemann, D and K. H. Schlag, "Should I Stay or Should I Go? Search without Priors," 2011. Mimeo.

McCall, John J., "Economics of information and job search," *Quarterly Journal of Economics*, 1970, *84*, 113–126.

Parakhonyak, Alexei and Anton Sobolev, "Non-Reservation Price Equilibrium and Search without Priors," *Economic Journal*, 2015, *125*, 887–909.

Rothschild, Michael, "Searching for the Lowest Price When the Distribution of Prices is Unknown," *Journal of Political Economy*, 1974, *82*, 689–711.

Schlag, Karl and Andriy Zapechelnyuk, "Dynamic benchmark targeting," 2015. Mimeo.

Sleator, D. and R. Tarjan, "Amortized efficiency of list update and paging rules," *Communications of the ACM*, 1985, *28*, 202–208.