Sequential Search with Multi-Attribute Options

Churlzu Lim
Department of Systems and Industrial Engineering, University of Arizona, Tucson, Arizona 85721 clim@sie.arizona.edu,

J. Neil Bearden
Department of Management and Policy, University of Arizona, Tucson, Arizona 85721 jneilb@gmail.com,

J. Cole Smith
Department of Systems and Industrial Engineering, University of Arizona, Tucson, Arizona 85721 cole@sie.arizona.edu,

We describe a search problem in which a decision maker (DM) must select several sequentially-encountered options. Each option is described by multiple attributes, and the value of an option is given by a function of its attribute values. However, the attribute values are not known with certainty, and can only be ascertained in a predefined order, at some fixed cost. During the search the DM can choose to select an option, purchase information about an attribute value, reject (permanently) the current option and continue the search, or terminate the search and accept a status quo outcome. We introduce a Threshold Policy for this search process, and provide a class of option value functions for which the policy is optimal. We then furnish a dynamic programming procedure for prescribing an optimal policy for this problem. Finally, we derive analytic solutions to some special cases of the problem, and present a case study that demonstrates a possible use of the proposed approach.

Key words: multiattribute optimization; sequential search; dynamic programming

1. Introduction
In the current paper, we describe a sequential search problem in which the decision maker (DM) sequentially encounters options whose overall value is a function of multiple attributes. The DM encounters each option one at a time, and must decide whether to accept the option under consideration or permanently dismiss it. Furthermore, the DM must pay to learn the precise values of an option’s attributes. To illustrate this sequential search problem, consider the problem of buying a house. Typically one does not learn the total worth or utility of a house just by learning its price alone. Rather, one must expend resources and time to learn more about the integrity of the home (say by paying for the services of an appraiser), one’s potential future neighbors, plans for future development in the areas around the home, and so on. One might find a house that seems relatively well-priced, find out from an appraiser that the house is structurally sound, and then find out later that the next door neighbors are remarkably obnoxious. At any point during the information acquisition process for a given house, one can decide to purchase the house, pay to learn more about another aspect of the house, continue searching through houses, or terminate the search and accept the status quo (e.g., continuing to live in one’s current apartment). Importantly, when one decides to continue the search, the costly process of learning the worth of the next house must begin anew.

A number of sequential search problems have received considerable attention across a range of disciplines. The classical problem is as follows: A DM observes a sequence of $N$ options with the objective of selecting the maximum (or minimum) of the sequence. Typically, each option is represented by a random variable $X_n$ ($n = 1, \ldots, N$) whose distribution function is fully known to the DM. Further, the problems are often constrained so that once an option is rejected it cannot be chosen in the future. Problems with this structure have been referred to in the literature under different names; sometimes the problem is referred to as the house-selling problem, other times as...
the job-search problem. The basic theory of optimal stopping in classical search problems can be found in a number of sources (e.g. Chow et al. 1971, Fisher 1961, MacQueen and Miller 1960).

There also have been a number of extensions of the classical problem. These include allowing the recall of previously observed options, adding search costs, and the relaxation of the assumption that the DM knows the distribution from which the options are sampled. Most relevant to us here, some work has been done in multi-dimensional versions of the problem. In these, at each stage \( n \) in the search the DM observes a set of \( K \) options. Bruss and Ferguson (1997), for example, considered a problem in which the DM can sell any number of options at each stage, and in which the search continues until exactly \( K \) options are sold. The payoff that the DM obtains is equal to sum of the obtained selling prices minus a search cost, which is based on the total number of vectors observed before the final option is sold. Some related problems have been studied by Glickman (2004), Karlin (1962), Mazalov and Saario (2002), and Saario and Skaguchi (1990).

MacQueen (1964) described a problem similar to the one we propose here. In his study the \( n \)th option is represented by a two-dimensional vector \( X_n = (X_{n,1}, X_{n,2}) \), whose elements can be revealed at some cost to the DM. Specifically, the DM can pay a cost \( c_1 \) to observe \( X_{n,1} \) and can acquire further information by paying an additional cost \( c_2 \) to learn \( X_{n,2} \). Once \( X_{n,1} \) is known, the DM can accept the option immediately, pay \( c_2 \) to get a better estimate of the option’s worth by learning \( X_{n,2} \), or reject the current option and examine the next option. If the DM pays to examine \( X_{n,2} \), she then decides either to accept or reject the option. The variance in the DM’s estimate of the option’s value strictly decreases as more information is gathered. MacQueen showed that the optimal policy is a set of thresholds that determines for what values of \( X_{n,k} \) \( (k = 1, 2) \) the DM will exercise each of the available actions.

Stigler (1961) recognized the importance of search costs in a number of areas of economics. He pointed out that at any point in time—unless a market is centralized—no single buyer can know the prices of all the available sellers. Furthermore, given that there will be some degree of dispersion in the prices, a buyer may do well to search for good prices rather than simply selecting the first one encountered. However, search is costly. One may incur costs due to the necessity to travel to alternative sellers, or perhaps to the time it takes to search. This characteristic has been modelled in a variety of sequential search problems (e.g. Kennedy and Kertz 1990, 1991, 1992) and also in problems in which a DM must search within a single option (e.g. McCardle 1985).

Whereas Stigler (1961) was primarily concerned with costs associated with searching over options, McCardle (1985) looked at the effects of search costs within options. In particular, he examined the problem faced by a firm deciding whether to adopt a new technology. Typically, there is a great deal of uncertainty around the profitability of the adoption of some new innovation. Hence, it behooves a firm to reduce this uncertainty before making an adoption decision. McCardle modelled the information acquisition as a search problem in which the firm sequentially gathers information and updates its estimate of the innovation’s profitability in a Bayesian manner. At each stage in the information gathering process, the firm must decide whether to accept the change, gather more information, or reject the innovation. The basic search process can be described as follows. The firm starts with a prior estimate of the profitability of the innovation. Next it acquires information and updates its profitability estimate. Whenever the estimate crosses some acceptability threshold, the firm adopts the innovation; likewise, whenever the estimate goes below some lower threshold, the firm ceases its information search and rejects the innovation. In between these two bounds, the firm continues the search. (The threshold-based decision policy described in McCardle is related to the policies described in Monahan (1980) and Monahan (1982) for problems in machine replacement.) McCardle presented a dynamic programming procedure for determining the optimal thresholds for this search process and presented results on the effects of different search variables. He showed, for example, that as the cost of obtaining information increases, the firm will more quickly decide to accept or reject a new technology.
In our problem, the DM faces the dilemma of both how deeply to search within an option to learn its worth (cf. McCardle), and also how deeply across options to search (cf. Stigler). One important work that combines these concepts is presented by Lippman and McCardle (1991), who examine the problem of when to select an option out of an infinite pool of candidates, and how thoroughly to evaluate the worth of an option under examination. We build on the work of these researchers, and present a multi-attribute sequential search problem, where each attribute value for an option under examination can be learned with certainty in a single observation, and the value of an option is given by some function of these attributes.

The remainder of this paper is organized as follows. In Section 2, we introduce a policy framework that can be applied to our search problems, identify a class of value functions for which this framework is optimal, and present a dynamic programming (DP) procedure for finding optimal policies for those functions. Using the proposed DP procedure, we provide analytic solutions in Section 3 for some special cases, in which the value function is the summation of attributes. We demonstrate numerical results for a case study in Section 4, and conclude our work in Section 5.

2. Search Problem Properties and Algorithm
In this chapter we begin by formally introducing the Sequential Multi-Attribute Option Search Problem (SMOSP) in Section 2.1. We then state a threshold policy for the SMOSP in Section 2.2, along with sufficient conditions under which this policy is optimal. We synthesize this information in Section 2.3 with a dynamic programming algorithm for optimizing the SMOSP.

2.1. SMOSP Problem Statement
The SMOSP contains a finite set of $N$ options (or stages) to be considered in sequential fashion, from which the DM can select as many as $H \leq N$ options. Each option must be selected or rejected before continuing on to the next option. We can sequentially observe values of distinct attributes $1, \ldots, K$ of each option. For this problem, we assume that both the order in which the options are encountered and the order in which the attributes are investigated are fixed. (Although this is a limiting assumption, the problem in which the DM can either statically or dynamically determine this order seems to be significantly more difficult under our framework. We suggest an analysis of this extension for future research.)

The cost of ascertaining the value of attribute $k$ for option $n$ is given by $c_{n,k}$ for each $k = 1, \ldots, K$. In many applications, $c_{n,k} = c_{n',k}$ for all $1 \leq n < n' \leq N$. Additionally, selecting option $n$ without viewing any of its attributes incurs a cost of $c_{n,0}$. The attribute values $x_{n,k}$ represent the outcomes of the random variable $X_{n,k}$ for $n = 1, \ldots, N$ and $k = 1, \ldots, K$, where the possible outcomes for $X_{n,k}$ lie in a sample space $\Omega_{n,k}$. While we do not know these outcomes a priori, we are aware of the distributions of $X_{n,1}, \ldots, X_{n,K}$ for each $n = 1, \ldots, N$. We assume for the sake of model tractability that these distributions are independent.

The value function for an option is given by $f$; we let $E^n_f$ be the expected value of option $n = 1, \ldots, N$ without observing the outcome of any random variable, and let $E^n_f(x_{n,1}, \ldots, x_{n,k})$ be the expected value of option $n$ after having observed outcomes $x_{n,1}, \ldots, x_{n,k}$ for the first $k$ attributes. We also include a dummy stage $N + 1$, at which time we choose a status quo reward to account for any unselected options. The objective of the DM is to maximize the sum of the value functions for the selected options plus the status quo reward.

We also assume that the value function $f$ exhibits the property that for any $n = 1, \ldots, N$ and $k = 1, \ldots, K - 1$, if $E^n_f(x_{n,1}', \ldots, x_{n,k}') \leq E^n_f(x_{n,1}', \ldots, x_{n,k})$, then $E^n_f(x_{n,1}', \ldots, x_{n,k}' ; x_{n,k+1}) \leq E^n_f(x_{n,1}', \ldots, x_{n,k}, x_{n,k+1})$, for any $x_{n,i}' \in \Omega_{n,i}$ and $x_{n,i} \in \Omega_{n,i}$, $i \in \{1, \ldots, k\}$, and for any $x_{n,k+1} \in \Omega_{n,k+1}$. Intuitively, if the expected value of option $n$ under observations $x' = (x_{n,1}', \ldots, x_{n,k}')$ does not exceed the expected option value under observations $x'' = (x_{n,1}', \ldots, x_{n,k})$, then the realization of a common outcome $x_{n,k+1}$ will not make the set of outcomes $x'$ preferable to $x''$. We call such
functions expectation-monotonic functions. These functions include linear functions, or more generally, the summation of functions that are separable in attribute values. Note also that the order in which attribute values are observed plays an important role as to whether or not a function is expectation-monotonic (see Example 1).

2.2. Threshold Policy Definition
Consider a stage $n \in \{1, \ldots, N\}$ at which we have selected $h$ options so far, and define $v^h_n$ to be our optimal expected future objective value at stage $n$ given that $h$ options have been selected. Hence, recalling that stage $N + 1$ is the dummy termination stage, we take $v^h_{N+1}$ as the status quo reward associated with stopping the search after selecting $h$ alternatives. We define the following Threshold Policy (TP) rules. (This policy is illustrated in Figure 1.)

Threshold Policy (TP)
- **Rule 1:** Suppose that we have observed all attribute values $x_{n,1}, \ldots, x_{n,K}$ for option $n$. If $E^f_n(x_{n,1}, \ldots, x_{n,K}) \geq v^h_{n+1} - v^h_n$, then we select option $n$; else, we reject it.
- **Rule 2:** Suppose that we have observed the outcomes $x_{n,1}, \ldots, x_{n,k}$ for the first $k$ attributes, where $1 \leq k \leq K - 1$. If $E^f_n(x_{n,1}, \ldots, x_{n,k}) \geq \alpha^h_{n,k}$ for some scalar threshold $\alpha^h_{n,k}$, we select the current option and continue to stage $n + 1$ with $h + 1$ selected options. Else, if $E^f_n(x_{n,1}, \ldots, x_{n,k}) < \beta^h_{n,k}$ for a scalar threshold $\beta^h_{n,k} \leq \alpha^h_{n,k}$, we reject option $n$ and continue to the next stage $n + 1$ with $h$ selected options. Finally, when $E^f_n(x_{n,1}, \ldots, x_{n,k}) \in [\beta^h_{n,k}, \alpha^h_{n,k})$, we purchase information about attribute $k + 1$ to obtain more information on option $n$.
- **Rule 3:** If we have not yet observed any attribute values for option $n$, we decide whether or not to observe $X_{n,1}$ at a cost of $c_{n,1}$. If we do not view the value of this first attribute, we either select or reject option $n$ and continue to option $n + 1$, or stop and take the status quo reward $v^h_{n+1}$, whichever yields the maximum expected value. Consider the maximum reward among those obtained by accepting the status quo ($v^h_{N+1}$), rejecting the option outright ($v^h_{n+1}$), and accepting the option outright ($v^h_{n+1} + E^f_n - c_{n,0}$). If the expected reward for observing $X_{n,1}$ is greater than this maximum value, we inspect the value of $X_{n,1}$. Otherwise, we take the action that yields the maximum value.

While TP does not necessarily yield the optimal policy in general, Proposition 1 indicates that it is optimal for expectation-monotonic functions. We follow the statement of this proposition with an example showing that the TP rules are not necessarily optimal when $f$ is not expectation-monotonic.

**Proposition 1.** If $f$ is an expectation-monotonic function, then there exists an optimal policy for the SMOSEP that follows the TP rules.

**Proof.** See Appendix. □

**Example 1.** For functions that do not exhibit the expectation-monotonic property, the TP rule does not necessarily yield an optimal policy. To see an example for which the TP is suboptimal, suppose that $N = H = 1$ and $K = 2$, where $f(x_{1,1}, x_{1,2}) = x_{1,1}^4 (x_{1,2} - 3)^3$. Let $c_{1,0} = \infty$, $c_{1,1} = 0$, and $c_{1,2} = 90$, with $v^0_{N+1} = 120$. Let $x_{1,1}$ be uniformly distributed on the interval $[1, 3]$, and let $x_{1,2}$ be uniformly distributed on the interval $[1, 7]$. Note that this function is not expectation-monotonic, since although $E^f_1(1) < E^f_1(2)$, for instance, we have that $E^f_1((1, 1)) > E^f_1((2, 1))$. In fact, in this scenario, the optimal policy is to examine the first attribute, and if $E^f_1(x_{1,1}) < 119.945$ (or $x_{1,1} < 1.861$), reject option 1 and take the status quo. If $119.945 \leq E^f_1(x_{1,1}) < 433.537$ (or $1.861 \leq x_{1,1} < 2.566$), select the option without purchasing information about the second attribute. If $E^f_1(x_{1,1}) \geq 433.537$ (or $x_{1,1} \geq 2.566$), purchase information about $X_{1,2}$.

Observe that if these attributes were encountered in the reverse order, we would indeed have that the function is expectation-monotonic, and hence the TP rules are optimal. Suppose that $f(x_{1,1}, x_{1,2})$ is now defined as $x_{1,2}^4 (x_{1,1} - 3)^3$, with the same parameters above except that $X_{1,1}$ and
Figure 1 Illustration of TP at stage $n$ given $h$ options selected.

Stage $N+1$ given $h$

1: Select the current option.
2: Continue to the next stage given $h$ options selected.
3: Inspect the value of the next attribute.
4: Choose a status quo given $h$ options selected.

$X_{1,2}$ are now uniformly distributed on the intervals [1, 7] and [1, 3], respectively. Our optimal policy in this case would be to reject the option if $E_{1}^{f}(x_{1,1}) < 119.947$ and otherwise to accept the option. Reducing $c_{1,2}$ to 30 induces a policy in which we reject the option if $E_{1}^{f}(x_{1,1}) < 97.460$, purchase information about $X_{1,2}$ if $97.460 \leq E_{1}^{f}(x_{1,1}) < 245.918$, and select the option if $E_{1}^{f}(x_{1,1}) \geq 245.918$. Note that this policy is of the TP form.

2.3. Dynamic Programming Algorithm

We now describe a detailed dynamic programming procedure to find optimal TP strategies $\pi_{n}^{h} \equiv \{(\alpha_{n,1}^{h}, \beta_{n,1}^{h}), \ldots, (\alpha_{n,K-1}^{h}, \beta_{n,K-1}^{h})\}$, $\forall n, \forall h$, that maximize the expected value of an expectation-monotonic function for problem SMOSP. We denote the decision-making epoch at which we consider option $n \in \{1, \ldots, N\}$ as stage $n$. (At “stage $N+1$” we must choose the status quo reward of $v_{N+1}^{h}$ for $h = 0, \ldots, H$.) Suppose that we have found optimal TP strategies $\pi_{N}^{h}, \ldots, \pi_{n+1}^{h}$ and corresponding expected rewards $v_{N}^{h}, \ldots, v_{n+1}^{h}, \forall h$. Let $\pi_{n,k}^{h} = \{(\alpha_{n,k}^{h}, \beta_{n,k}^{h}), \ldots, (\alpha_{n,K-1}^{h}, \beta_{n,K-1}^{h})\}$ denote a partial policy at stage $n$ for some $k \in \{1, \ldots, K-1\}$. Furthermore, having observed values of
\( X_{n,i} = x_{n,i} \) for \( i = 1, \ldots, k-1 \), let \( w^h_{n,k}(x_{n,1}, \ldots, x_{n,k-1}; \pi^h_{n,k}) \) represent the optimal expected future reward when we observe \( X_{n,i} \) using the (partial) policy \( \pi^h_{n,k} \). Given \( X_{n,1}, \ldots, x_{n,k-1} \), define \( \Omega^c_{n,k}, \Omega^s_{n,k} \), and \( \Omega^k_{n,k} \) to be subsets of \( \Omega_{n,k} \) such that \( \Omega^c_{n,k} \equiv \{ x_{n,k} : E_n^f(x_{n,1}, \ldots, x_{n,k}) < \beta_{n,k} \} \), \( \Omega^f_{n,k} = \{ x_{n,k} : \beta_{n,k} \leq E_n^f(x_{n,1}, \ldots, x_{n,k}) < \alpha_{n,k} \} \), and \( \Omega^s_{n,k} \equiv \{ x_{n,k} : E_n^f(x_{n,1}, \ldots, x_{n,k}) \geq \alpha_{n,k} \} \) for \( k = 1, \ldots, K-1 \). Similarly, given \( x_{n,1}, \ldots, x_{n,K-1} \), define \( \Omega^c_{n,K} \) and \( \Omega^s_{n,K} \) to be subsets of \( \Omega_{n,K} \) such that \( \Omega^c_{n,K} \equiv \{ x_{n,K} : E_n^f(x_{n,1}, \ldots, x_{n,K}) < v_{n+1} \} \) and \( \Omega^s_{n,K} \equiv \Omega_{n,K} \setminus \Omega^c_{n,K} \). Then we have

i) for \( k = 1, \ldots, K-1 \),

\[
\begin{align*}
w^h_{n,k}(x_{n,1}, \ldots, x_{n,k-1}; \pi^h_{n,k}) &= -c_{n,k} + Pr[X_{n,k} \in \Omega^c_{n,k}] v^h_{n+1} \\
&\quad + Pr[X_{n,k} \in \Omega^f_{n,k}] E^h_{n,k+1}(x_{n,1}, \ldots, x_{n,k-1}, x_{n,k}, \pi^h_{n,k+1}) | X_{n,k} \in \Omega^f_{n,k}] \\
&\quad + Pr[X_{n,k} \in \Omega^s_{n,k}] \left[ E^f_{n}(x_{n,1}, \ldots, x_{n,k-1}) | X_{n,k} \in \Omega^s_{n,k} \right] + v^h_{n+1},
\end{align*}
\]  
(1a)

ii) for \( k = K \),

\[
\begin{align*}
w^h_{n,K}(x_1, \ldots, x_{K-1}; \pi^h_{n,k}) &= -c_{n,K} + Pr[X_{n,K} \in \Omega^c_{n,K}] v^h_{n+1} \\
&\quad + Pr[X_{n,K} \in \Omega^s_{n,K}] \left[ E^f_{n}(x_{n,1}, \ldots, x_{n,K-1}) | X_{n,K} \in \Omega^s_{n,K} \right] + v^h_{n+1},
\end{align*}
\]  
(1b)

where \( E^f_{n}(x_{n,1}, \ldots, x_{n,k-1}) | X_{n,k} \in G \) denotes the conditional expected value of option \( n \) given observed attribute values \( x_{n,1}, \ldots, x_{n,k-1} \) and conditioned on \( X_{n,k} \in Y \) for any set \( G \in \Omega_{n,k}, n \in \{1, \ldots, N\} \) and \( k \in \{1, \ldots, K\} \).

Using (1) and noting that \( \pi^h_{n} = \pi^h_{n,1} \), an optimal policy \( \pi^h_{n} \) at stage \( n \) given \( h \) is the solution to

\[
\max_{\pi^h_{n}} \max_{n} w^h_{n,1}(\pi^h_{n}).
\]  
(2)

Furthermore, the optimal expected reward at stage \( n \) for \( h \) is

\[
v^h_n = \max \left\{ v^h_{n+1}, v^h_{n+1} + E^f_n - c_{n,0}, w^h_{n,1}(\pi^h_{n}) \right\}.
\]  
(3)

Observe that we face \( K \) sequential decision-making steps given \( h \) selected options at stage \( n \). Furthermore, at step \( k \), optimal policies are already available with regard to the states following the current decision. Therefore, instead of directly solving the optimization problem in (2), we recommend a dynamic programming procedure to find an optimal policy at stage \( n \) for \( h \) as follows. The optimal policies \( (\beta^h_{n,K-1}, \alpha^h_{n,K-1}), \ldots, (\beta^h_{n,1}, \alpha^h_{n,1}) \) can be obtained by recursively solving

\[
w^h_{n,k}(x_1, \ldots, x_{k-1}; \pi^h_{n,k}) = \max_{(\alpha^h_{n,k}, \beta^h_{n,k})} w^h_{n,k}(x_1, \ldots, x_{k-1}, \pi^h_{n,k}) \}.
\]  
(4)

We summarize a dynamic programming procedure that finds optimal policies \( \pi^h_{n} \) for \( n = N, \ldots, 1 \) and \( h = 0, \ldots, H \), using a backward recursion.

**DP Procedure**

**Step 0.** Set the stage and the inner loop counters as \( n = N + 1 \) and \( k = K - 1 \), respectively.

**Step 1.** If \( n = 1 \), terminate the procedure with the optimal expected reward \( v^h_1 \). Otherwise, decrement the stage counter \( n \leftarrow n - 1 \), put \( H = \min\{n, H - 1\} \), initialize \( h = 0 \), and proceed to Step 2.

**Step 2.** Solve (4) to obtain \( \pi^h_{n,k} \) and \( w^h_{n,k}(x_1, \ldots, x_{k-1}; \pi^h_{n,k}) \). If \( k > 1 \), decrement \( k \leftarrow k - 1 \) and repeat Step 2. Else, put \( k = K - 1 \) and proceed to Step 3.

**Step 3.** Put \( v^h_n = \max \left\{ v^h_{n+1}, v^h_{n+1} + E^f_n - c_{n,0}, w^h_{n,1}(\pi^h_{n}) \right\} \). If \( n = H \), return to Step 1. Otherwise, increment \( h \leftarrow h + 1 \), and return to Step 2.
3. Special Case: Linear Additive Reward Function

In this section, we demonstrate how to implement the foregoing DP procedure for a reward function that is the sum of attribute values, i.e., \( f(x_{n,1}, \ldots, x_{n,K}) = \sum_{k=1}^{K} x_{n,k}. \) (Note that this reward function is expectation-monotonic.) We assume that we can select at most one option, i.e., \( H = 1. \) Furthermore, assume that \( X_{n,k} \) for all \( n \) and \( k \), are independently distributed random variables whose density function is continuous on a range \([a_{n,k}, b_{n,k}]\) such that \(-\infty \leq a_{n,k} < b_{n,k} \leq \infty. \) Since \( H = 1, \) we have \( v_{n,i} = 0, \forall n. \) The remaining task is to find optimal policies, \( \pi_{n}^{*} \) for \( n = 1, \ldots, N. \) To this end, we will omit the superscript \( h = 0, i.e., v_{n} = v_{n0}, \pi_{n} = \pi_{n}^{0}, \) etc. After observing the \( k \)th attribute at stage \( n, \) a decision is made based on \( E_{n}^{f}(x_{n,1}, \ldots, x_{n,K}) = \sum_{i=1}^{K} x_{n,i} + \sum_{i=k+1}^{K} E[X_{n,i}]. \) That is, (1) if \( \sum_{i=1}^{K} x_{n,i} \geq \alpha_{n,k} = a_{n,k} - \sum_{j=k+1}^{K} E[X_{n,j}], \) select the current option, (2) else if \( \sum_{i=1}^{K} x_{n,i} < \beta_{n,k} = b_{n,k} - \sum_{j=k+1}^{K} E[X_{n,j}], \) drop the option and continue to the next stage, (3) else, inspect the next attribute of the current option. To simplify this discussion, we find optimal values of \((\alpha_{n,k}^{h}, \beta_{n,k}^{h})\) instead of \((a_{n,k}, b_{n,k})\). Accordingly, we use \( \pi_{n,k}^{*} \) in lieu of \( \pi_{n,k}. \) In addition, let us denote \( \sum_{i=1}^{k} x_{n,i} = s_{n,k}, \sum_{i=k+1}^{K} E[X_{n,i}] = E_{n,k+1}, \) and \( w_{n,k+1}(x_{n,1}, \ldots, x_{n,K}; \pi_{n,k+1}) = w_{n,k+1}(s_{n,k}; \pi_{n,k+1}). \) Then, we have the following expected reward when observing \( X_{n,k+1}. \)

i) For \( k = 1, \ldots, K - 2, \)

\[
\begin{align*}
w_{n,k+1}(s_{n,k}; \pi_{n,k+1}) &= -c_{n,k+1} + Pr(X_{n,k+1} + s_{n,k} < \beta_{n,k+1}^{*}) \times v_{n+1} \\
&+ Pr(X_{n,k+1} + s_{n,k} \geq \alpha_{n,k+1}^{*}) \\
&\times E[s_{n,k} + X_{n,k+1} + E_{n,k+2} \mid X_{n,k+1} + s_{n,k} \geq \alpha_{n,k+1}^{*}] \\
&+ Pr(\beta_{n,k+1}^{*} \leq X_{n,k+1} + s_{n,k} < \pi_{n,k+1}^{*}) \\
&\times E[w_{n,k+1}(s_{n,k} + X_{n,k+1}; \pi_{n,k+1}^{*}) \mid \beta_{n,k+1}^{*} \leq X_{n,k+1} + s_{n,k} < \pi_{n,k+1}^{*}].
\end{align*}
\]

ii) For \( k = K - 1, \)

\[
\begin{align*}
w_{n,K}(s_{n,K-1}; \pi_{n,K}) &= -c_{n,K} + Pr(X_{n,K} + s_{n,K-1} < v_{n+1}) \times v_{n+1} \\
&+ Pr(X_{n,K} + s_{n,K-1} \geq v_{n+1}) \\
&\times E[X_{n,K} + s_{n,K-1} \mid X_{n,K} + s_{n,K-1} \geq v_{n+1}].
\end{align*}
\]

In Lemmas 1 and 2, and in Proposition 2, we consider stage \( n, \) where we have observed \( x_{n,1}, \ldots, x_{n,K} \) and have an optimal partial TP \( \pi_{n,i}^{*} \) for \( i = k + 1, \ldots, K - 1. \) Furthermore, note that if \( c_{n,k+1} = 0, \) we have a trivial optimal policy in which we always observe \( X_{n,k+1}. \) Hence, we only consider \( k \) such that \( c_{n,k+1} > 0. \)

**Lemma 1.** The following claims hold true.

i) \( w_{n,k+1}(s_{n,k}; \pi_{n,k+1}^{*}) \) is continuous and nondecreasing in \( s_{n,k}. \)

ii) For any \( s_{n,k} \in \mathbb{R}, \) we have

\[
\lim_{\delta \to 0^+} \frac{w_{n,k+1}(s_{n,k} + \delta; \pi_{n,k+1}^{*}) - w_{n,k+1}(s_{n,k}; \pi_{n,k+1}^{*})}{\delta} \leq 1.
\]

iii) We have that

\[
\begin{align*}
w_{n,k+1}(s_{n,k}; \pi_{n,k+1}^{*}) &\to -c_{n,k+1} + v_{n+1} \quad \text{as} \quad s_{n,k} \to -\infty \\
w_{n,k+1}(s_{n,k}; \pi_{n,k+1}^{*}) &\to -c_{n,k+1} + s_{n,k} + E_{n,k+1} \quad \text{as} \quad s_{n,k} \to \infty.
\end{align*}
\]
Proof. Since the probability density function of \( X_{n,k} \) is continuous, \( w_{n,k+1}(s_{n,k}; \pi_{n,k+1}) \) is continuous. Noting that \( E_t^n(x_1, \ldots, x_{n,k}) = s_{n,k} + E_{n+1} \), we have from Lemma 3 that \( w_{n,k+1}(s_{n,k}; \pi_{n,k+1}) \) is nondecreasing in \( s_{n,k} \). Lemma 4 together with the continuity of our distribution functions establishes that (6a) holds true. (See the Appendix for the statement and proofs of Lemmas 3 and 4.) Finally, from (5), as \( s_{n,k} \to -\infty \), we have \( Pr(X_{n,k+1} + s_{n,k} \geq \alpha_{n,k+1}) \to 0 \), \( Pr(\beta_{n,k+1} \leq X_{n,k+1} + s_{n,k} < \alpha_{n,k+1}) \to 0 \), and \( Pr(X_{n,k+1} + s_{n,k} < \beta_{n,k+1}) \to 0 \). Similarly, as \( s_{n,k} \to \infty \), we have \( Pr(X_{n,k+1} + s_{n,k} < \beta_{n,k+1}) \to 0 \), \( Pr(\beta_{n,k+1} \leq X_{n,k+1} + s_{n,k} < \alpha_{n,k+1}) \to 0 \), and \( Pr(X_{n,k+1} + s_{n,k} \geq \alpha_{n,k+1}) \to 0 \). Therefore, (6b) holds true. This completes the proof. □

**Lemma 2.** Consider \( \pi_{n,k} \) such that \( w_{n,k+1}(\pi_{n,k}; \pi_{n,k+1}) > \max\{v_{n+1}, \pi_{n,k} + E_{n+1}\} \). Then, there exist nonempty convex sets \( \pi_{n,k}^L = \{s_{n,k} : w_{n,k+1}(s_{n,k}; \pi_{n,k+1}^*) = v_{n+1}, s_{n,k} < \pi_{n,k}\} \) and \( \pi_{n,k}^U = \{s_{n,k} : w_{n,k+1}(s_{n,k}; \pi_{n,k+1}^*) = s_{n,k} + E_{n+1}, s_{n,k} \geq \pi_{n,k}\} \).

Proof. Since \( c_{n,k+1} > 0 \), \( w_{n,k+1}(s_{n,k}; \pi_{n,k+1}) \) is continuously nondecreasing, and \( w_{n,k+1}(s_{n,k}; \pi_{n,k+1}) \to -c_{n,k+1} + v_{n+1} \) as \( s_{n,k} \to -\infty \), we have a nonempty convex set \( \pi_{n,k}^L \). Likewise, since \( c_{n,k+1} > 0 \), \( w_{n,k+1}(s_{n,k}; \pi_{n,k+1}) \) is continuous, its rate of change does not exceed 1, and \( w_{n,k+1}(s_{n,k}; \pi_{n,k+1}) \to -c_{n,k+1} + s_{n,k} + E_{n+1} \) as \( s_{n,k} \to \infty \), we have a nonempty convex set \( \pi_{n,k}^U \).

Given \( s_{n,k} \), if we reject the current option, the reward function value is \( v_{n+1} \). Also, if we select the option, the expected reward will be \( s_{n,k} + E_{n+1} \). Hence, in order to find an optimal policy, we need to identify \( \pi_{n,k}^L \) and \( \pi_{n,k}^U \) such that

\[
\begin{align*}
\pi_{n,k}^L & = \max\{w_{n,k+1}(s_{n,k}; \pi_{n,k+1}) = v_{n+1}, s_{n,k} < \pi_{n,k}\} & \text{if } \pi_{n,k}^L \leq \pi_{n,k}^U, \\
\pi_{n,k}^U & = \max\{w_{n,k+1}(s_{n,k}; \pi_{n,k+1}) = s_{n,k} + E_{n+1}, s_{n,k} \geq \pi_{n,k}\} & \text{otherwise}.
\end{align*}
\]

**Proposition 2.** Let \( \pi_{n,k} = v_{n+1} - E_{n+1} \). Furthermore, let \( s_{n,k}^L \in \pi_{n,k}^L \) and \( s_{n,k}^U \in \pi_{n,k}^U \) as defined in Lemma 2. Then, the following policy is optimal:

\[
\pi_{n,k}^L = \pi_{n,k}^U = \alpha_{n,k} \quad \text{if } w_{n,k+1}(\pi_{n,k}; \pi_{n,k+1}) \leq v_{n+1},
\]

\[
\pi_{n,k}^L = \pi_{n,k}^U = \pi_{n,k} \quad \text{otherwise}.
\]

Proof. Note that if \( s_{n,k} \leq \pi_{n,k} \), we have \( v_{n+1} \geq s_{n,k} + E_{n+1} \). Otherwise, we have \( v_{n+1} \leq s_{n,k} + E_{n+1} \). Also, note that \( v_{n+1} = \pi_{n,k} + E_{n+1} \). First, suppose that \( \pi_{n,k+1}(\pi_{n,k}; \pi_{n,k+1}) \leq v_{n+1} \). Then, since \( w_{n,k+1}(s_{n,k}; \pi_{n,k+1}) \) is continuously nondecreasing, we have \( v_{n+1} \geq w_{n,k+1}(s_{n,k}; \pi_{n,k+1}) \) for \( s_{n,k} \leq \pi_{n,k} \). Similarly, from (6a), we have \( s_{n,k} + E_{n+1} \geq v_{n+1} \) for \( s_{n,k} \geq \pi_{n,k} \). Hence, this proves the first case (see Figure 2(a) for an illustration of this case). Now, suppose that \( \pi_{n,k+1}(s_{n,k}; \pi_{n,k+1}) > v_{n+1} \). From Lemma 2, we have the existence of values \( s_{n,k}^L \) and \( s_{n,k}^U \). For the same reasons as in the first case, we have \( v_{n+1} \geq w_{n,k+1}(s_{n,k}; \pi_{n,k+1}) \) for \( s_{n,k} \leq s_{n,k}^L \) and \( s_{n,k} + E_{n+1} \geq w_{n,k+1}(s_{n,k}; \pi_{n,k+1}) \) for \( s_{n,k} \geq s_{n,k}^U \). Also, by the continuity of the function \( w_{n,k+1}(s_{n,k}; \pi_{n,k+1}) \), we have \( s_{n,k} \geq s_{n,k}^L \). Therefore, the policy prescribed in (7) is optimal. □

From Proposition 2, we can summarize the dynamic programming procedure to find an optimal TP policy for the linear additive reward function as follows.

**DP Procedure for a Linear Additive Reward Function**

**Step 0.** Set the stage and the inner loop counters as \( n = N + 1 \) and \( k = K - 1 \), respectively.

**Step 1.** If \( n = 1 \), terminate the procedure with the optimal expected reward \( v_1 \). Otherwise, decrement counters \( n \leftarrow n - 1 \) and \( k \leftarrow k - 1 \).
Fig. 2 Illustration of Proposition 2.

\( s_{n,k} = \beta_{n,k}^* = \alpha_{n,k}^* \)

(a) \( w_{n+1} (s_{n,k}; \pi_{n+1}^*) \leq v_{n+1} \).

(b) \( w_{n+1} (s_{n,k}; \pi_{n+1}^*) > v_{n+1} \).

**Step 2.** If \( c_{n,k+1} = 0 \), put \( \beta_{n,k}^* = -\infty \), \( \alpha_{n,k}^* = \infty \), and proceed to Step 3. Otherwise, put \( s_{n,k} = v_{n+1} - E_{n,k+1} \) and compute \( w_{n+1} (s_{n,k}; \pi_{n,k+1}^*) \). If \( w_{n+1} (s_{n,k}; \pi_{n,k+1}^*) \leq v_{n+1} \), then put \( \beta_{n,k}^* = \alpha_{n,k}^* = s_{n,k}^* \). Else, put \( \beta_{n,k}^* = s_{n,k}^L \) and \( \alpha_{n,k}^* = s_{n,k}^U \), where \( s_{n,k}^L \in S_{n,k}^L \) and \( s_{n,k}^U \in S_{n,k}^U \) are defined as in Lemma 2. Proceed to Step 3.

**Step 3.** If \( k > 1 \), decrement \( k \leftarrow k - 1 \) and return to Step 2. Else, put \( k = K - 1 \) and \( v_n = \max \{ v_{N+1}, v_{n+1}, E_n^f - c_{n,0}, w_{n,1} (\pi_{n,1}^*) \} \), and return to Step 1.

**Remark 1.** Although \( s_{n,k} \) can have values only in a range \([\sum_{i=1}^k a_{n,k}, \sum_{i=1}^k b_{n,k}]\), we can have any real values for \( \beta_{n,k}^* \) and \( \alpha_{n,k}^* \). For example, if \( \beta_{n,k}^* < \sum_{i=1}^k a_{n,k} \), we never reject the option before observing \( X_{n,k+1} \) or selecting without the observation. On the other hand, if \( \beta_{n,k}^* > \sum_{i=1}^k b_{n,k} \), then we always reject the current option. \( \square \)

In what follows, we illustrate an implementation of the foregoing DP procedure for an example in which attributes are uniform random variables.

**Example 2.** This example illustrates the behavior of the SMOSP under uniformly distributed attributes. We examine a current option \( n \), and suppose that \( X_{n,k} \) for \( k = 1, 2, 3 \), are i.i.d. uniform random variables on the interval \([0, 100]\). Other values are given as \( v_{n+1} = 150, c_{n,1} = c_{n,2} = c_{n,3} = 5 \), and \( c_{n,0} = 0 \).
From (5b), we have that

\[ w_{n,3}(s_{n,2}; \pi_{n,3}^*) = -5 + \begin{cases} 
  150 & \text{for } s_{n,2} < 50 \\
  g_3(s_{n,2}) & \text{for } 50 \leq s_{n,2} < 150 \\
  s_{n,2} + 50 & \text{for } s_{n,2} \geq 150,
\end{cases} \]

where

\[ g_3(s_{n,2}) = \left( \frac{150 - s_{n,2}}{100} \right) 150 + \left( \frac{s_{n,2} - 50}{100} \right) \left( \frac{s_{n,2} + 250}{2} \right). \]

From Step 2, we have \( \tau_{n,2} = 100 \), and hence, \( w_{n,3}(\tau_{n,2}; \pi_{n,3}^*) = 157.5 > 150 = v_{n+1} \). Equating \( w_{n,3}(s_{n,2}; \pi_{n,3}^*) = 150 \) and \( w_{n,3}(s_{n,2}; \pi_{n,3}^*) = s_{n,2} + 50 \) yields \( \tau_{n,2}^L = s_{n,2}^L = 81.63 \) and \( \tau_{n,2}^U = s_{n,2}^U = 118.37 \), respectively.

Next, consider \( k = 1 \). Then, from (5a), we have that

\[ w_{n,2}(s_{n,1}; \pi_{n,2}^*) = -5 + \begin{cases} 
  150 & \text{for } s_{n,1} < -18.37 \\
  g_2^1(s_{n,1}) & \text{for } -18.37 \leq s_{n,1} < 18.37 \\
  g_2^2(s_{n,1}) & \text{for } 18.37 \leq s_{n,1} < 81.63 \\
  g_2^3(s_{n,1}) & \text{for } 81.63 \leq s_{n,1} < 118.37 \\
  s_{n,1} + 100 & \text{for } s_{n,1} \geq 118.37,
\end{cases} \]

where

\[ g_2^1(s_{n,1}) = \left( \frac{81.63 - s_{n,1}}{100} \right) 150 \]
\[ + \left( \frac{s_{n,1} + 18.37}{100} \right) E \left[ w_{n,3}(s_{n,1} + X_{n,2}; \pi_{n,3}^*) \mid 81.63 - s_{n,1} \leq X_{n,2} < 100 \right], \]
\[ g_2^2(s_{n,1}) = \left( \frac{81.63 - s_{n,1}}{100} \right) 150 \]
\[ + \left( \frac{s_{n,1} - 18.37}{100} \right) \left( \frac{s_{n,1} + 18.37}{2} \right) \]
\[ + \left( \frac{36.74}{100} \right) E \left[ w_{n,3}(s_{n,1} + X_{n,2}; \pi_{n,3}^*) \mid 81.63 \leq X_{n,2} + s_{n,1} < 118.37 \right], \]
\[ g_2^3(s_{n,1}) = \left( \frac{s_{n,1} - 18.37}{100} \right) \left( \frac{s_{n,1} + 18.37}{2} \right) \]
\[ + \left( \frac{118.37 - s_{n,1}}{100} \right) E \left[ w_{n,3}(s_{n,1} + X_{n,2}; \pi_{n,3}^*) \mid 0 \leq X_{n,2} < 118.37 - s_{n,1} \right]. \]

The conditional expectation of reward when observing \( X_{n,3} \) can be computed as follows.

i) \(-18.37 \leq s_{n,1} < 18.37\)

\[ E \left[ w_{n,3}(s_{n,1} + X_{n,2}; \pi_{n,3}^*) \mid 81.63 - s_{n,1} \leq X_{n,2} < 100 \right] \]
\[ = -5 + \Pr[X_{n,2} + X_{n,3} < 150 - s_{n,1} \mid 81.63 - s_{n,1} \leq X_{n,2} < 100] 150 \]
\[ + \Pr[X_{n,2} + X_{n,3} \geq 150 - s_{n,1} \mid 81.63 - s_{n,1} \leq X_{n,2} < 100] \]
\[ \times E [s_{n,1} + X_{n,2} + X_{n,3} \mid X_{n,2} + X_{n,3} \geq 150 - s_{n,1}, \ 81.63 - s_{n,1} \leq X_{n,2} < 100] \]
\[ = -5 + \left( \frac{81.37 - s_{n,1}}{200} \right) 150 + \left( \frac{s_{n,1} + 87.5}{s_{n,1} + 18.37} \right) \frac{1250}{81.37} . \]
Table 1 Summary of Examples for Uniformly and Exponentially Distributed Attributes.

<table>
<thead>
<tr>
<th>Attributes</th>
<th>$\beta_{n,1}^\star$</th>
<th>$\alpha_{n,1}^\star$</th>
<th>$\beta_{n,2}^\star$</th>
<th>$\alpha_{n,2}^\star$</th>
<th>$v_{n+1}$</th>
<th>$v_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform over $[0, 100]$</td>
<td>27.30</td>
<td>72.70</td>
<td>81.63</td>
<td>118.37</td>
<td>150</td>
<td>159.30</td>
</tr>
<tr>
<td>Exponential with mean 50</td>
<td>-26.89</td>
<td>90.92</td>
<td>34.87</td>
<td>125.84</td>
<td>150</td>
<td>168.77</td>
</tr>
</tbody>
</table>

ii) $18.37 \leq s_{n,1} < 81.63$

$$E\left[w_{n,3}(s_{n,1} + X_{n,2}; \pi_{n,3}^\star) \mid 81.63 \leq X_{n,2} + s_{n,1} < 118.37\right] = 158.06.$$  

iii) $81.63 \leq s_{n,1} < 118.37$

$$E\left[w_{n,3}(s_{n,1} + X_{n,2}; \pi_{n,3}^\star) \mid 0 \leq X_{n,2} < 118.37 - s_{n,1}\right] = -5 + \left(\frac{181.63 - s_{n,1}}{200}\right) + \frac{s_{n,1}^2}{3} + 139.46s_{n,1} + 4007.49.$$  

Solving $w_{n,2}(s_{n,1}; \pi_{n,2}^\star) = 150$ and $w_{n,2}(s_{n,1}; \pi_{n,2}^\star) = s_{n,1} + 100$, we have $\beta_{n,1}^\star = s_{n,1}^L = 27.30$ and $\alpha_{n,1}^\star = s_{n,1}^U = 72.70$, respectively. □

Similarly, we can derive an optimal policy for i.i.d. exponential random attributes with a mean of 50. We report the results of this exercise in Table 1 along with the results from Example 2. Note that $\beta_{n,1}^\star$ is negative (-26.89) for exponential distribution, and hence, we never drop the current option after observing the first attribute.

4. Case Study: Searching for an Employee

In this section, we present a brief case study that illustrates the impact of certain parameters on the optimal search strategy. It is possible, of course, to vary any number of parameters of the search problem. Here, we have restricted consideration to search costs and the quality of status quo options because we consider these to be the most interesting.

Consider the problem of searching for an employee to fill an open position. The position may require good technical skills, interpersonal skills, knowledge of legal matters, and so on. It is more costly to learn about some of these attributes than others. Basic interpersonal skills, for example, can be observed in a relatively brief interview. Getting a good sense of an applicant’s technical skills may require costly testing. Let us now consider how assessment costs and the value of the company’s status quo option affect the interviewer’s search policy.

Suppose each applicant is represented by $K = 2$ attributes, that the attribute values are sampled from a uniform distribution on the interval $[0, 100]$ (assume the traits have all been scaled to fall within this range), and that the quality of an applicant is an additive measure of her attributes. According to the personnel selection literature, additive combinations of applicant traits are most often used in making hiring decisions (see, e.g. van den Berg and Feij 2003). The same sorts of decision rules also often play a role in college admissions.

Imagine that the company is looking to fill a single administrative assistant position. Hence, $H = 1$ and $v^i_i = 0 \forall i = 1, \ldots, N + 1$; and, as in the above examples, for economy, we will omit the $h$ superscript when expressing the optimal policies (i.e., $v_n = v^0_n$, etc.). The position requires someone who is friendly (first attribute) and who has good basic computer skills (second attribute). For the company, a short interview to get a sense of an applicant’s interpersonal skills has a relatively

\footnote{Details of this derivation are provided at Online Supplements (http://da.pubs.informs.org/online-supp.html).}
Table 2: Optimal Policies for
\( \upsilon_{N+1} = 50, \quad c_{n, 1} = 2, \quad \text{and} \quad c_{n, 2} = 10 \forall n \)

<table>
<thead>
<tr>
<th>n</th>
<th>( \beta_{n, 1}^* )</th>
<th>( \alpha_{n, 1}^* )</th>
<th>( \upsilon_n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>64.6025</td>
<td>75.1598</td>
<td>122.5464</td>
</tr>
<tr>
<td>2</td>
<td>60.6783</td>
<td>71.2356</td>
<td>119.8811</td>
</tr>
<tr>
<td>3</td>
<td>54.4331</td>
<td>64.9904</td>
<td>115.9570</td>
</tr>
<tr>
<td>4</td>
<td>42.8509</td>
<td>53.4082</td>
<td>109.7118</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>5.2786</td>
<td>98.1295</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>-</td>
<td>50.0000</td>
</tr>
</tbody>
</table>

Table 3: Optimal Policies for
\( \upsilon_{N+1} = 50, \quad c_{n, 1} = 2, \quad \text{and} \quad c_{n, 2} = 5 \forall n \)

<table>
<thead>
<tr>
<th>n</th>
<th>( \beta_{n, 1}^* )</th>
<th>( \alpha_{n, 1}^* )</th>
<th>( \upsilon_n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>54.1087</td>
<td>90.8632</td>
<td>125.5459</td>
</tr>
<tr>
<td>2</td>
<td>49.7566</td>
<td>86.5110</td>
<td>122.4859</td>
</tr>
<tr>
<td>3</td>
<td>43.0377</td>
<td>79.7921</td>
<td>118.1338</td>
</tr>
<tr>
<td>4</td>
<td>30.8976</td>
<td>67.6521</td>
<td>111.4149</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>18.3772</td>
<td>99.2749</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>-</td>
<td>50.0000</td>
</tr>
</tbody>
</table>

Inexpensive cost of \( c_{n, 1} = 2, \forall n \). However, getting a sufficient measure of an applicant’s computer skills requires sending the applicant to a testing center where the evaluation cost \( c_{n, 1} = 10, \forall n \), is greater. Let us look at a case in which the company currently has a poor-quality person filling the administrative assistant position; i.e., their status quo is relatively undesirable (\( \upsilon_{N+1} = 50 \)). The optimal policy for a search with five (\( N = 5 \)) applicants is shown in Table 2.

Let us contrast the optimal search policy shown in Table 2 with the one for the case in which the cost of the computer skills assessment is cut in half (i.e., \( c_{n, 2} = 5 \)). These results are shown in Table 3. First, note that the DM should put applicants through the computer assessment more often when the assessment is less expensive. This is revealed by the greater spread between \( \alpha_{n, 1}^* \) and \( \beta_{n, 1}^* \) when the cost of assessment is decreased. Further, the DM in the higher-cost case is more likely to accept applicants right away on the basis of their interpersonal skills; this can be seen in the lower values of \( \alpha_{n, 1}^* \) in the high-cost case.

Also observe that the DM’s expected payoff (\( \upsilon_1 = 125.55 \)) for the search is greater when the assessment costs are lower than when they are higher (\( \upsilon_1 = 122.55 \)). In addition to earning more (on average), the DM in the lower-cost scenario is expected to search deeper into the set of applicants (to applicant 2.25) than the DM in the higher-cost scenario (to applicant 1.79). Less obvious perhaps is that the DM in the higher-cost case is expected to incur lower costs (average incurred cost 5.48) than the DM in the lower-cost case (average incurred cost 8.56). To summarize, the DM in the higher-cost case more quickly dismisses applicants on the basis of their first trait, and is less likely to put applicants through the computer skills assessment. On the other hand, the DM in the lower-cost scenario searches through more applicants and ultimately receives a greater reward than the corresponding DM in the higher-cost case.

Now, imagine that our company already has a good administrative assistant (\( \upsilon_{N+1} = 120 \)); nonetheless, they are looking to replace him with a better one. We will keep the skills assessment costs low (\( c_{n, 2} = 5 \)) and examine the effects of this status quo on the optimal search policy. Comparing Tables 3 and 4, we observe that the availability of a good status quo option makes the DM a bit more choosy. She is less apt to hire an applicant right away after the interview (\( \alpha_{n, 1}^* \) is greater
Table 4 Optimal Policies for $v_{N+1} = 120$, $c_{n,1} = 2$, and $c_{n,2} = 5 \forall n.$

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\beta^*_n,1$</th>
<th>$\alpha^*_n,1$</th>
<th>$v_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>61.7303</td>
<td>98.4848</td>
<td>131.3609</td>
</tr>
<tr>
<td>2</td>
<td>60.1501</td>
<td>96.9045</td>
<td>130.1075</td>
</tr>
<tr>
<td>3</td>
<td>58.1113</td>
<td>94.8657</td>
<td>128.5273</td>
</tr>
<tr>
<td>4</td>
<td>55.3976</td>
<td>92.1521</td>
<td>126.4885</td>
</tr>
<tr>
<td>5</td>
<td>51.6228</td>
<td>88.3772</td>
<td>123.7749</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>-</td>
<td>120.0000</td>
</tr>
</tbody>
</table>

when $v_{N+1}$ is higher). Further, she is more likely to dismiss intermediate-quality folks immediately ($\beta^*_n,1$ is greater when $v_{N+1}$ is higher). And with the good status quo available, the DM tends to search deeper into the pool of applicants (to applicant 2.69) and also incurs greater costs (average incurred costs 10.34). In summary, having a good-quality incumbent in hand already, the DM can be choosier, search longer, and obtain a better-quality optimal solution (with $v_1 = 131.36$).

A more complete picture of the ways in which our DM’s search policies are affected by assessment costs and the quality of the status quo option can be obtained by examining Figures 3 and 4. To keep matters simple, we restrict focus to the optimal stage 1 policies and to problems in which costs are the same for both attributes. Consider left-hand plot in Figure 3 for which $v_{N+1} = 110$. Note that as the costs increase, $\alpha^*_n,1$ decreases, and, consequently, the DM is more apt to immediately hire applicants with poorer values on the first attribute, without learning about the applicant’s second attribute. At the same time, however, $\beta^*_n,1$ increases as the assessment costs increase, thereby also causing the DM to more quickly reject applicants after viewing only their first attribute. The net effect, then, is that as the assessment costs increase, the DM becomes progressively less likely to fully evaluate the applicants and more likely to base her hiring decisions only on the basis of interpersonal skills assessment (attribute 1). Further, from the behavior of $v_2$, we can see that as the costs increase, she is less demanding about the applicants’ overall worths. (Recall that the DM selects the $n$th applicant whenever selecting that applicant leads to a payoff that exceeds the expected payoff for moving to the next applicant and choosing optimally thereafter, $v_{n+1}$.) This same pattern persists for $v_{N+1} = 130$ (right-hand plot in Figure 3). However, with a better status quo option available, the DM is less likely both to hire applicants knowing only their first attribute and to pay to learn applicants’ second attribute; and, once the second attribute is known, she is more demanding, as evidenced by the increase in $v_2$. In short, a DM with a valuable status quo option can, therefore, be more choosy, especially when it is cheap to assess applicants. Figure 4 summarizes the behavior expected payoffs ($v_1$) as a function of a range of status quo and assessment cost values. The result is clear and to be expected: Expected earnings increase in the value of the status quo option and decrease in assessment costs.

5. Concluding Remarks

The Sequential Multi-Attribute Option Search Problem captures important properties of a number of realistic search problems. We are often faced with problems in which we must decide how much time, energy, and money to invest in evaluating particular decision options, and also how deeply into the set of available options we should search. By insufficiently evaluating a particular option, we may either pass it up when it is of high worth, or we may accept it prematurely when it is of low worth. Likewise, we can spend too much time evaluating clearly poor options or even clearly exceptional ones. Most previous work on optional stopping has bypassed the within-option search problem. The single-attribute problems do capture problems in which all available options are identical (e.g., when multiple sellers are selling identical goods) or when only a single feature
is important to a DM (e.g., when a seller is only concerned with the selling price). The single-
attribute problems, however, do not have a perfect correspondence to problems in which the decision 
alternatives are multi-attribute.

There exist several theoretical avenues that can be explored in future research. Although we have 
examined some elementary comparative statics in Section 4, we have not proven anything about 
the characteristics of the $\alpha$-, $\beta$-, or $\nu$-values as different problem characteristics change, such as 
varying search costs, attribute values, and the amount of risk in the attribute distributions (in the 
sense of mean-preserving changes of risk; see Lippman and McCall (1976)). A different problem 
would examine the order in which different options and attributes should be investigated if the DM 
has this luxury, under static or dynamic ordering scenarios. While this introduces a combinatorial
element to the decision maker, it is likely that simple, effective rules of thumb exist for prescribing optimal orderings under certain restrictions of the problem investigated herein.

In addition to encouraging theoretical work on multi-attribute sequential search problems, we hope that this paper will encourage empirical work to determine how actual DMs solve these kinds of problems. Already, Bearden and Connolly (2004) have experimentally studied the SMOSP with linear additive rewards. They found that subjects in their experiments, who earned real money for their decisions, had a strong tendency to search too deeply into particular options and insufficiently deeply into the set of options. These findings suggest that decision makers faced with real problems captured by the SMOSP could benefit from decision aids that can be developed based on theoretical work on the SMOSP. We believe this is a rich avenue for work in decision analysis that will open up a number of new, interesting problems.

Acknowledgments
We gratefully acknowledge financial support by a contract F49620-03-1-0377 from the AFOSR/MURI to the Department of Systems and Industrial Engineering and the Department of Management and Policy at the University of Arizona. We are also indebted to two referees and the editors for their insightful comments and suggestions that improved the presentation of this paper.

Appendix. Lemmas and Propositions Regarding TP Optimality
We state here formal proofs for two lemmas leading to Proposition 1, which establishes the optimality of the TP rules for expectation-monotonic functions, followed by a formal proof of Proposition 1 itself. First, define $\omega^h_n(x)$ to be the optimal expected future objective function value for choosing the remaining $H-h$ options, where $h \in \{0, \ldots, H-1\}$, given that we (a) are currently examining option $n$, (b) have observed attribute values $x$ belonging to some set $\Omega_{n,1} \times \ldots \times \Omega_{n,k}$ for some $k = 1, \ldots, K-1$, and (c) will examine the value of the $k+1$ attribute. (The value $\omega^h_n(x)$ includes the attribute investment cost of $c_{n,k+1}$.) Note that the dependence of this definition on $k$ is captured by the dimension of $x$.

**Lemma 3.** For any expectation-monotonic function $f$, if $E'_n(x') \leq E'_n(x'')$ for some $x'$ and $x''$ that belong to $\Omega_{n,1} \times \ldots \times \Omega_{n,k}$, where $k = 1, \ldots, K-1$, then $\omega^h_n(x') \leq \omega^h_n(x'')$.

**Proof.** Given $x'$ and $x''$ such that $E'_n(x') \leq E'_n(x'')$, where $f$ is expectation-monotonic, consider the policy that leads to $\omega^h_n(x')$. Suppose that we use this policy (perhaps suboptimally) for the collection $x''$ of outcomes too, and suppose that we next decide to examine $x_{n,k+1} \in \Omega_{n,k+1}$. First, examine the case of $k = K-1$. Rejecting option $n$ after viewing the outcome $x_{n,k}$ leads to the same objective function value for either $x'$ or $x''$, while accepting the option after viewing $x_{n,K}$ leads to an objective function value at least as large for $x''$ as for $x'$ (since $f$ is expectation-monotonic). For values of $k \leq K-2$, we proceed by induction. Assume that the conclusion of the lemma holds for all $x'$ and $x''$ belonging to $\Omega_{n,1} \times \ldots \times \Omega_{n,k}$, where $k \in \{1, \ldots, K-2\}$, and examine the case for $k$. Again, if the decision to reject option $n$ is made after learning $x_{n,k+1}$, then there is no change in the objective function in using $x''$ over $x'$, while if option $n$ is accepted after learning $x_{n,k+1}$, then there cannot be an improvement by discovering $x'$ instead of $x''$. If the decision is made to examine the value of the next attribute $(k+2)$, then by induction and the fact that $E'_n([x', x_{n,k+1}]) \leq E'_n([x'', x_{n,k+1}])$, the objective improves or remains the same by using $x''$ over $x'$.

Since $\omega^h_n(x'')$ is at least as large as the objective we would obtain by mirroring the optimal policy used for $\omega^h_n(x')$, and since this mirror policy leads to at least as good of an expected objective function as $\omega^h_n(x')$, we have the desired result that $\omega^h_n(x'') \geq \omega^h_n(x')$. This completes the proof. □

**Lemma 4.** For any expectation-monotonic function $f$, if $E'_n(x') \leq E'_n(x'')$ for some $x'$ and $x''$ that belong to $\Omega_{n,1} \times \ldots \times \Omega_{n,k}$, $k = 1, \ldots, K-1$, then $E'_n(x') - E'_n(x'') \leq \omega^h_n(x') - \omega^h_n(x'')$.

**Proof.** We can compare $E'_n(x)$ to $\omega^h_n(x)$, for some $x \in \Omega_{n,1} \times \ldots \times \Omega_{n,k}$, $k \in \{1, \ldots, K-1\}$, by rewriting these expectations as

$$E'_n(x) = \int_{x_{n,k+1} \in \Omega_{n,k+1}} f_{x_{n,k+1}}(x_{n,k+1}) E'_n([x', x_{n,k+1}]) \, dx_{n,k+1}$$

$$\omega^h_n(x) = \int_{x_{n,k+1} \in \Omega_{n,k+1}} f_{x_{n,k+1}}(x_{n,k+1}) \gamma^h_n([x', x_{n,k+1}]) \, dx_{n,k+1},$$

where $\gamma^h_n([x', x_{n,k+1}])$ is expectation-monotonic. For values of $h$, we have observed attribute values $x$ belonging to some set $\Omega_{n,1} \times \ldots \times \Omega_{n,k}$ for some $k = 1, \ldots, K-1$, and (c) will examine the value of the $k+1$ attribute. (The value $\omega^h_n(x)$ includes the attribute investment cost of $c_{n,k+1}$.) Note that the dependence of this definition on $k$ is captured by the dimension of $x$. For values of $k \leq K-2$, we proceed by induction. Assume that the conclusion of the lemma holds for all $x'$ and $x''$ belonging to $\Omega_{n,1} \times \ldots \times \Omega_{n,k}$, where $k \in \{1, \ldots, K-2\}$, and examine the case for $k$. Again, if the decision to reject option $n$ is made after learning $x_{n,k+1}$, then there is no change in the objective function in using $x''$ over $x'$, while if option $n$ is accepted after learning $x_{n,k+1}$, then there cannot be an improvement by discovering $x'$ instead of $x''$. If the decision is made to examine the value of the next attribute $(k+2)$, then by induction and the fact that $E'_n([x', x_{n,k+1}]) \leq E'_n([x'', x_{n,k+1}])$, the objective improves or remains the same by using $x''$ over $x'$. Since $\omega^h_n(x'')$ is at least as large as the objective we would obtain by mirroring the optimal policy used for $\omega^h_n(x')$, and since this mirror policy leads to at least as good of an expected objective function as $\omega^h_n(x')$, we have the desired result that $\omega^h_n(x'') \geq \omega^h_n(x')$. This completes the proof. □
where $f_{X_{n,k+1}}$ is the probability density function for $X_{n,k+1}$, and $\gamma^*_n([x, x_{n,k+1}])$ is the optimal expected objective function value given observations $[x, x_{n,k+1}]$ of option $n$ with $h$ selected options thus far. Formally, for a set of outcomes $\mathbf{y} \in \Omega_{n,1} \times \ldots \times \Omega_{n,k}$, $k \in \{1, \ldots, K\}$, an option $n \in \{1, \ldots, N\}$, and $h \in \{0, \ldots, H-1\}$, define

$$\gamma^*_n(y) = \begin{cases} v^h_{n+1} & \text{if option } n \text{ is immediately rejected} \\ \omega^h_{n}(y) & \text{if we purchase information on attribute } k+1 \\ E'_n(y) + v^h_{n+1} & \text{if option } n \text{ is immediately accepted}, \end{cases}$$

after observing $y$. (Note that the second possibility of (8) is not applicable if $k = K$.) Hence, we can compute $(\omega^h_{n}(\mathbf{x}') - E'_n(\mathbf{x}')) - (\omega^h_{n}(\mathbf{x}'') - E'_n(\mathbf{x}''))$ as

$$\int_{x_{n,k+1} \in \Omega_{n,k+1}} f_{X_{n,k+1}}(x_{n,k+1}) \left[ (\gamma^*_n([x, x_{n,k+1}]) - E'_n([x, x_{n,k+1}])) - (\gamma^*_n([x'', x_{n,k+1}]) - E'_n([x'', x_{n,k+1}])) \right] dx_{n,k+1}. \quad (9)$$

Following similar logic as in the proof of Lemma 3, suppose that $\mathbf{x}'$ and $\mathbf{x}''$ are realizations of the first $\tilde{k} = K-1$ attributes, and that we have determined the outcome $x_{n,K}$ of the $K$th attribute. For $k = 1, \ldots, K-1$, consider the objective value $\gamma^*_n([x', x_{n,k+1}])$ obtained by mirroring the policy for $\gamma^*_n([x', x_{n,k+1}])$ after learning $x_{n,k+1}$. If the decision is made to reject option $n$, then $\gamma^*_n([x', x_{n,k+1}]) = \gamma^*_n([x'', x_{n,K}]) - v^h_{n+1}$. Since $E'_n(\mathbf{x}') \leq E'_n(\mathbf{x}'')$, we have that $\gamma^*_n([x', x_{n,k+1}]]) \leq E'_n([x', x_{n,K}])$ for all $n \in \{1, \ldots, N\}$, and therefore that $\gamma^*_n([x', x_{n,K}]) - E'_n([x', x_{n,K}]) \geq \gamma^*_n([x', x_{n,K}]) - E'_n([x'', x_{n,K}])$ in this case. Otherwise, if we accept option $n$, we have that $\gamma^*_n([x', x_{n,K}]) - E'_n([x', x_{n,K}]) = \gamma^*_n([x'', x_{n,K}]) - E'_n([x'', x_{n,K}]) = v^h_{n+1}$. Since $\gamma^*_n(\mathbf{x}') \geq \gamma^*_n(\mathbf{x}'')$, the lemma holds for $\tilde{k} = K-1$.

Suppose this lemma holds true for all $\tilde{k} + 1$, where $\tilde{k} \in \{1, \ldots, K-2\}$, and examine the case for $\tilde{k}$. We construct the same argument as before for the case in which we accept or reject the option after learning $x_{n,k+1}$. Also, if we choose to examine attribute number $\tilde{k} + 2$ for option $n$, we have that $\gamma^*_n([x', x_{n,k+1}]) - E'_n([x', x_{n,K}]) \geq \gamma^*_n([x'', x_{n,k+1}]) - E'_n([x'', x_{n,K}])$ by induction. Hence, noting again that $\gamma^*_n(\mathbf{x}') \geq \gamma^*_n(\mathbf{x}'')$, each element inside the integral given by (9) is nonnegative, and thus $E'_n(\mathbf{x}') - E'_n(\mathbf{x}'') \leq \omega^h_{n}(\mathbf{x}') - \omega^h_{n}(\mathbf{x}'')$. This completes the proof.

**Proof of Proposition 1.** The optimality of the first and the third rules in TP is obvious. Now, consider any optimal policy that does not obey Rule 2 of TP. Then for some option $n \in \{1, \ldots, N\}$ and $h \in \{0, \ldots, H-1\}$, there exist outcomes $\mathbf{x}' = (x'_{n,1}, \ldots, x'_{n,k})$ and $\mathbf{x}'' = (x''_{n,1}, \ldots, x''_{n,k})$, $1 \leq k \leq K-1$, such that $E'_n(\mathbf{x}') \leq E'_n(\mathbf{x}'')$ such that this optimal policy must either (i) select option $n$ after observing $\mathbf{x}'$ but reject it after observing $\mathbf{x}''$, (ii) purchase information regarding attribute $k+1$ for option $n$ after observing $\mathbf{x}'$ but reject option $n$ after observing $\mathbf{x}''$, or (iii) select option $n$ after observing $\mathbf{x}'$ but purchase more information regarding attribute $k+1$ for option $n$ after observing $\mathbf{x}''$. We show that there exists an alternative optimal policy that complies with Rule 2 of TP as follows.

In case (i), there must exist two nonempty sets of joint outcomes $Y'$ and $Y''$ for $X_{n,1}, \ldots, X_{n,k}$, such that $E'_n(\mathbf{x}') \leq E'_n(\mathbf{x}'')$ for all $\mathbf{x}' \in Y'$ and $\mathbf{x}'' \in Y''$, and such that the optimal policy selects option $n$ after observing $\mathbf{x}'$ but rejects it after observing $\mathbf{x}''$. Define $p_1$ to be the probability that some $\mathbf{x} = (x_{n,1}, \ldots, x_{n,k})$ is in $Y'$ and $p_2$ to be the probability that $\mathbf{x} \in Y''$. Consider the alternative policy in which option $n$ is rejected if $\mathbf{x} \in Y'$. Then the change in the objective function by using this alternative policy is given by

$$\Delta_1 = p_1[v^h_{n+1} - E'_n(\mathbf{x}|\mathbf{x} \in Y') - v^h_{n+1}], \quad (10a)$$

where $E'_n(\mathbf{x}|\mathbf{x} \in Y')$ denotes the expected value for option $n$ given observations $\mathbf{x}$ that belong to some set $Y \subseteq \Omega_{n,1} \times \ldots \times \Omega_{n,k}$ for some $n \in \{1, \ldots, N\}$ and $k \in \{1, \ldots, K\}$. Additionally, consider a second alternative to the optimal policy in which option $n$ is accepted if $\mathbf{x} \in Y''$. The change in the objective function incurred by using this second alternative policy is given by

$$\Delta_2 = p_2[E'_n(\mathbf{x}|\mathbf{x} \in Y'') + v^h_{n+1} - v^h_{n+1}], \quad (10b)$$

However, at least one of these alternative policies must not worsen the objective function. If $\Delta_1 < 0$ and $\Delta_2 < 0$ as assumed in our argument by contradiction, then (10a) and (10b) imply that

$$E'_n(\mathbf{x}|\mathbf{x} \in Y') < v^h_{n+1} - v^h_{n+1} < E'_n(\mathbf{x}|\mathbf{x} \in Y'). \quad (10c)$$

However, since $E'_n(\mathbf{x}|\mathbf{x} \in Y'') \geq E'_n(\mathbf{x}|\mathbf{x} \in Y')$ by definition, (10c) is impossible, and an optimal policy must exist in which case (i) does not occur.
Next, in case (ii), we set up $Y'$ and $Y''$ as before. In this case, if an outcome $x$ of $X_{n,1}, \ldots, X_{n,k}$ belongs to $Y'$, we decide to purchase information about attribute $k+1$, while if $x \in Y''$, we reject option $n$. In one alternative policy, we could reject option $n$ if $x \in Y''$. The change in the objective function would be given by

$$\Delta_1 = p[n^h_{n+1} - E[w_n^h(x)|x \in Y']]. \quad (11a)$$

A second alternative to the optimal policy would purchase information about attribute $k+1$ if $x \in Y''$. The objective change in this case is

$$\Delta_2 = p[n^h_{n+1} - v_h^h_{n+1}]. \quad (11b)$$

If both of these alternative policies worsen the objective function, (11a) and (11b) would imply that $E[w_n^h(x)|x \in Y'] > E[w_n^h(x)|x \in Y'']$. However, Lemma 3 (along with the fact that $E'_n(x') \leq E'_n(x'')$ for all $x' \in Y'$ and $x'' \in Y''$), proves that this is not possible, and an optimal policy exists in which case (ii) does not arise.

Finally, in case (iii), we employ the same steps as before. In this case, if an outcome $x \in Y'$, we accept option $n$, while if $x \in Y''$, we purchase more information about the option. If we decide to purchase more information about option $n$ when $x \in Y'$, the change in the objective function is given by

$$\Delta_1 = p[n E[w_n^h(x)|x \in Y'] - E_n^h(x|x \in Y') - v_h^h_{n+1}]. \quad (12a)$$

A second alternative would be to accept option $n$ if $x \in Y''$, which would yield the objective change

$$\Delta_2 = p[n E'_n(x|x \in Y'') + v_h^h_{n+1} - E[w_n^h(x)|x \in Y'']]. \quad (12b)$$

By contradiction, suppose $\Delta_1 < 0$ and $\Delta_2 < 0$ (and thus that we could not optimally impose our dual threshold rule for this problem). These conditions would imply that

$$E[w_n^h(x)|x \in Y'] - E_n^h(x|x \in Y') < E[w_n^h(x)|x \in Y''] - E'_n(x|x \in Y''). \quad (12c)$$

which is impossible, due to Lemma 4. Hence, an optimal policy exists in which case (iii) does not appear. This completes the proof. \qed

References


