*J. Appl. Prob.* **53,** 1052–1063 (2016)
*doi:*10.1017/jpr.2016.64
© *Applied Probability Trust* 2016

# SEQUENTIAL STOCHASTIC ASSIGNMENT PROBLEM WITH TIME-DEPENDENT RANDOM SUCCESS RATES

GOLSHID BAHARIAN,* *University of Montreal*
ARASH KHATIBI ** *** AND
SHELDON H. JACOBSON,** **** *University of Illinois at Urbana-Champaign*

**Abstract**

The sequential stochastic assignment problem (SSAP) allocates distinct workers with deterministic values to sequentially arriving tasks with stochastic parameters to maximize the expected total reward. In this paper we study an extension of the SSAP, in which the worker values are considered to be random variables, taking on new values upon each task arrival. Several SSAP models with different assumptions on the distribution of the worker values and closed-form expressions for optimal assignment policies are presented.

*Keywords:* Sequential stochastic assignment; optimal policy

2010 Mathematics Subject Classification: Primary 60G20
Secondary 68R01

## 1. Introduction

Consider the sequential stochastic assignment problem (SSAP) introduced in [6], where $n$ workers are available to perform $n$ independent and identically distributed (i.i.d.) sequentially arriving tasks so as to maximize the expected total reward. A random variable $X_j$ denotes the value of the $j$th task that arrives during time period $j$, with a fixed value (or success rate) $p_i$ associated with worker $i$. If the $i$th worker is assigned to the $j$th task with observed value $x_j$, the worker becomes unavailable for future assignments, and the expected reward due to this assignment is given by $p_i x_j$. The objective is to assign these $n$ workers to $n$ tasks so as to maximize the expected total reward. It was shown in [6] that there exists numbers

$$-\infty = a_{0,n} \le a_{1,n} \le a_{2,n} \le \cdots \le a_{n,n} = +\infty$$

such that the optimal choice in the initial stage is to assign the $i$th best available worker (i.e. that with the $i$th highest success rate) if the random variable $X_1$ falls within the $i$th highest interval. Similarly, $n - j + 1$ intervals are defined for the $j$th stage and the optimal choice in the $j$th stage is to assign the $i$th best available worker if the random variable $X_j$ falls within the $i$th highest interval. The optimal expected total reward obtained from assigning all the $n$ tasks to workers is $\sum_{i=1}^{n} p_i a_{i,n+1}$, and, hence, $a_{i,n+1}$ is the expected value of the quantity assigned to the $i$th best worker. Moreover, the $a_{i,n}$ are independent of the worker values and depend on

Received 22 September 2015; revision received 14 January 2016.
* Postal address: CHU Sainte Justine, University of Montreal, 3175 Chemin de la Cote Sainte-Catherine, Montreal, Québec, H3T 1C5, Canada. Email address: golshidb@gmail.com
** Postal address: Department of Computer Science, University of Illinois at Urbana-Champaign, 201 North Goodwin Avenue, Urbana, IL 61801, USA.
*** Email address: khatibi2@illinois.edu
**** Email address: shj@illinois.edu



https://doi.org/10.1017/jpr.2016.64 Published online by Cambridge University Press

the task values' distribution function. Specifically, these breakpoints are computed recursively from

$$a_{i,n+1} = \int_{a_{i-1,n}}^{a_{i,n}} u \, \mathrm{d}G(u) + a_{i-1,n}G(a_{i-1,n}) + a_{i,n}(1 - G(a_{i,n})) \tag{1}$$

for $i = 1, 2, \ldots, n$, with $a_{0,n} = -\infty$ and $a_{n,n} = +\infty$, where $G$ is the cumulative distribution function of task values [6].

The SSAP has applications in several areas, and various extensions to the problem have been discussed in the literature. For example, Nikolaev *et al.* [10] studied a variation of the SSAP in aviation security screening systems where sequentially arriving passengers are assigned to the available screening devices, while Su and Zenios [12] addressed the problem of allocating sequentially arriving kidneys to patients on a waiting list. Another application of the SSAP is the asset selling problem [3], where one needs to choose the best offers out of a sequence of bids from potential buyers. Moreover, Albright [1] analyzed a problem where i.i.d. tasks arrive sequentially at random times according to a renewal process and must be assigned to a fixed set of workers. The same problem is studied in [2], where various arrival time distributions are considered. An extension of the SSAP with tasks arriving according to a Poisson process while having random deadlines is presented in [11]. A variation of the SSAP in which the number of tasks is unknown until after the final arrival and follows a given probability distribution is considered in [9]. Several extensions of the problem with various assumptions on the number of tasks and workers, and dependency between task values is studied in [7]. Asymptotic behavior of a special case of the SSAP, where worker values are only allowed to be 0 or 1, is analyzed in [4], and an optimal policy has been established as the number of arriving tasks approaches $\infty$. A generalization of this problem is studied in [5], where workers are allowed to take on any value. Also, Baharian and Jacobson [5] considered this model under the assumption of incomplete information, where task distributions are unobservable to the decision maker, and presented a stationary optimal policy to achieve the long-run expected reward per task. The assumption of fixed known success rates is relaxed in [8] and it is assumed that the workers success rates are random, but do not change over different time intervals.

The existing SSAP literature assumes the worker values to be either deterministic or fixed over different time intervals. In this paper we study an extension of the SSAP in which the success rates are random variables, taking on new values upon each task arrival. For example, consider the asset selling problem, where a sequence of bids (tasks) arrive from potential buyers, and a set of items (workers) are available to be sold to these buyers so as to maximize the expected total profit. A time-dependent random price is associated with each item. Each item price is dependent on the economic conditions upon the arrival of the offer and, hence, takes on a new value upon each task arrival, since economic conditions vary from time to time. This can be modeled as a set of distribution functions, governed by a Markov chain, which generate item prices. Markov chain transitions represent the variations in economic conditions (equivalently, switching from one price distribution to another). Specifically, let $Z_j$ denote the state of the Markov chain with state space $\mathcal{S} = \{1, 2, \ldots, r\}$ at time period $j$. Then $Z_j = k$ implies that item prices upon the arrival of offer $j$ are derived from the distribution function $F_k$, which corresponds to the economic conditions at time $j$.

The remainder of this paper is organized as follows. In Section 2 we present mathematical models of three classes of SSAPs with random success rates, providing closed-form expressions for optimal assignment policies so as to maximize the expected total reward. In Section 3 we present concluding comments and future research directions.

https://doi.org/10.1017/jpr.2016.64 Published online by Cambridge University Press

## 2. Model description

In this paper we study various versions of the SSAP, where worker success rates are no longer assumed to be fixed or deterministic. Specifically, three models are introduced which analyze this problem from different aspects.

I. The distribution function of worker values are identical or distinct.

II. The worker values in a given time period are dependent on those in the preceding time periods (or within the same time period).

III. The worker values are deterministic, but assumed to be functions of time (possibly deteriorating with time). Worker values are observable to the decision maker (i.e. the person who makes the assignments) and take on new values upon each task arrival.

### 2.1. Model I

Incorporating the randomness of worker values into the problem, in this subsection we discuss three distinct models. In model I.1 we assume that the sequence of arriving tasks is i.i.d., while the success rate of worker $i$ through time, $\{p_k(i)\}_{k=1}^n$, forms a sequence of dependent random variables for any arbitrarily fixed $i = 1, 2, \ldots, n$ (note that we assume that the workers are re-numbered at each stage such that at stage $j$, the set of remaining workers success rates is denoted by $\{p_j(i)\}$, $i = 1, 2, \ldots, n - j + 1$). Moreover, the set of worker success rates is assumed to be exchangeable (i.e. identically distributed, but not necessarily i.i.d.), given the past values. In the following theorem we present the optimal assignment policy to maximize the expected total reward for this model.

**Theorem 1.** *The optimal decision at each time period in model I.1 is to assign the best available worker to the current task. The resulting optimal expected total reward is then given by* $\mathbb{E}_{\phi^*}[R_n] = \mathbb{E}[X]\sum_{k=1}^n \mathbb{E}[M_p(k)]$, *where* $M_p(k) := \max_{i=1,\ldots,n-k+1} p_k(i)$, $R_n$ *is the total reward over the last n time periods, and* $\phi^*$ *is the optimal assignment policy.*

*Proof.* The proof is by induction on $n$. Starting from $n = 2$, the optimal conditional expected total reward after observing the first task and worker values is given by

$$\begin{aligned}
&\mathbb{E}_{\phi^*}[R_2 \mid X_1, P_1] \\
&\quad = \max\{\mathbb{E}[p_1(1)X_1 + p_2(2)X_2 \mid X_1, P_1], \mathbb{E}[p_1(2)X_1 + p_2(1)X_2 \mid X_1, P_1]\} \\
&\quad = \max\{p_1(1)X_1 + \mathbb{E}[p_2(2)X_2 \mid P_1], p_1(2)X_1 + \mathbb{E}[p_2(1)X_2 \mid P_1]\} \\
&\quad = \max\{p_1(1)X_1 + \mathbb{E}[p_2(2) \mid P_1]\mathbb{E}[X], p_1(2)X_1 + \mathbb{E}[p_2(1) \mid P_1]\mathbb{E}[X]\} \\
&\quad = \max_{i=1,2}\{p_1(i)X_1 + \mathbb{E}[M_p(2) \mid P_1]\mathbb{E}[X]\} \\
&\quad = M_p(1)X_1 + \mathbb{E}[M_p(2) \mid P_1]\mathbb{E}[X], \qquad (2)
\end{aligned}$$

where the second equality follows since $\sigma(P_2X_2, P_1)$ is independent of $\sigma(X_1)$, and the third equality follows since $\sigma(X_2)$ is independent of $\sigma(P_1, P_2)$. It can be deduced from (2) that the optimal decision upon the arrival of the first task is to assign it to the best available worker, since $\mathbb{E}[M_p(2) \mid P_1]\mathbb{E}[X]$ is independent of $i$. Now suppose that the statement of Theorem 1 holds for some $n \geq 2$, where the optimal conditional expected total reward is given by $\mathbb{E}_{\phi^*}[R_n \mid X_1, P_1] = M_p(1)X_1 + \mathbb{E}[X]\sum_{k=2}^n \mathbb{E}[M_p(k) \mid P_1]$, with $\phi^*$ denoting the optimal assignment policy over the last $n$ stages. To prove the claim, note that the optimal conditional

https://doi.org/10.1017/jpr.2016.64 Published online by Cambridge University Press

expected total reward for the case with $n+1$ workers, after observing the first task and worker values, is given by

$$\begin{aligned}
&\mathbb{E}_{\phi^*}[R_{n+1} \mid X_1, P_1] \\
&\quad = \max_{i=1,\dots,n+1} \{p_1(i)X_1 + \mathbb{E}_{\phi^*}[R_{n+1\setminus\{i\}} \mid X_1, P_1]\} \\
&\quad = \max_{i=1,\dots,n+1} \{p_1(i)X_1 + \mathbb{E}_{\phi^*}[\mathbb{E}_{\phi^*}[R_{n+1\setminus\{i\}} \mid X_2, P_2] \mid X_1, P_1]\} \\
&\quad = \max_{i=1,\dots,n+1} \left\{ p_1(i)X_1 + \mathbb{E}_{\phi^*}\left[ M_p(2)X_2 + \mathbb{E}[X] \sum_{k=3}^{n+1} \mathbb{E}[M_p(k) \mid P_2] \,\middle|\, X_1, P_1 \right] \right\} \\
&\quad = \max_{i=1,\dots,n+1} \left\{ p_1(i)X_1 + \mathbb{E}[X]\mathbb{E}[M_p(2) \mid P_1] + \mathbb{E}[X] \sum_{k=3}^{n+1} \mathbb{E}[M_p(k) \mid P_1] \right\} \\
&\quad = \max_{i=1,\dots,n+1} \left\{ p_1(i)X_1 + \mathbb{E}[X] \sum_{k=2}^{n+1} \mathbb{E}[M_p(k) \mid P_1] \right\} \\
&\quad = M_p(1)X_1 + \mathbb{E}[X] \sum_{k=2}^{n+1} \mathbb{E}[M_p(k) \mid P_1]\}, \qquad (3)
\end{aligned}$$

where the third equality follows from the induction assumption. It follows from (3) that it is optimal to assign the best available worker to the first arriving task and that $\mathbb{E}_{\phi^*}[R_{n+1}] = \mathbb{E}[X]\sum_{k=1}^{n+1}\mathbb{E}[M_p(k)]$. □

In model I.2 we analyze the case where i.i.d. sequentially arriving tasks are assigned to a set of workers, with $\{p_k(i)\}_{k=1}^n$ forming a martingale for arbitrarily fixed $i = 1, 2, \dots, n$. Worker success rates at a given time period are no longer assumed to be i.i.d. In fact, any kind of dependency is allowed. However, it is assumed that the worker values (although being random variables) have a fixed order throughout the problem (note that this assumption only holds for models I.2 and I.3 defined below). Specifically, suppose that the order of success rates observed in $P_1$ is assumed to be fixed throughout the problem. For simplicity, once this order is revealed at the beginning of assignments, the workers are re-labeled such that $p_1(i)$ would correspond to the worker with the $i$th smallest value. In the following theorem we detail how to tackle this problem in order to obtain the optimal expected total reward.

**Theorem 2.** *The optimal decision in model I.2 is to assign the $i$th best worker to the first task with value $X_1$ if it falls within the $i$th highest interval formed by the $\{a_{i,n}\}_{i=1}^{n-1}$, as calculated in (1). The resulting optimal expected total reward is then given by $\mathbb{E}_{\phi^*}[R_n] = \sum_{i=1}^n a_{i,n+1}\mathbb{E}[p_1(i)]$.*

*Proof.* The proof is by induction on $n$. Starting from $n = 2$, the optimal conditional expected total reward after observing the first task and worker values is given by

$$\begin{aligned}
&\mathbb{E}_{\phi^*}[R_2 \mid X_1, P_1] \\
&\quad = \max\{\mathbb{E}[p_1(1)X_1 + p_2(2)X_2 \mid X_1, P_1], \mathbb{E}[p_1(2)X_1 + p_2(1)X_2 \mid X_1, P_1]\} \\
&\quad = \max\{p_1(1)X_1 + \mathbb{E}[p_2(2)X_2 \mid P_1], p_1(2)X_1 + \mathbb{E}[p_2(1)X_2 \mid P_1]\} \\
&\quad = \max\{p_1(1)X_1 + \mathbb{E}[p_2(2) \mid P_1]\mathbb{E}[X], p_1(2)X_1 + \mathbb{E}[p_2(1) \mid P_1]\mathbb{E}[X]\} \\
&\quad = \max\{p_1(1)X_1 + p_1(2)\mathbb{E}[X], p_1(2)X_1 + p_1(1)\mathbb{E}[X]\} \\
&\quad = p_1(1)(X_1 \wedge \mathbb{E}[X]) + p_1(2)(X_1 \vee \mathbb{E}[X]), \qquad (4)
\end{aligned}$$

https://doi.org/10.1017/jpr.2016.64 Published online by Cambridge University Press

where the second equality follows since $\sigma(P_2X_2, P_1)$ is independent of $\sigma(X_1)$. The third equality follows since $\sigma(X_2)$ is independent of $\sigma(P_1, P_2)$, while the fourth equality is a direct result of the martingale property. Let '$\vee$' and '$\wedge$' denote the maximum and the minimum, respectively. Note that, according to (4), the optimal action to take upon the arrival of the first task is to compare its value with $a_{1,2} := \mathbb{E}[X]$ and assign it to the worker with value $p_1(1)$ if $X_1 = X_1 \wedge \mathbb{E}[X]$ and to $p_1(2)$ otherwise. As a consequence of (4), it follows that

$$\mathbb{E}_{\phi^*}[R_2 \mid P_1] = p_1(1)\mathbb{E}[X_1 \wedge \mathbb{E}[X]] + p_1(2)\mathbb{E}[X_1 \vee \mathbb{E}[X]] = p_1(1)a_{1,3} + p_1(2)a_{2,3}.$$

Now suppose that the statement of Theorem 2 holds for some $n \geq 2$, where the optimal assignment policy over the last $n$ stages is denoted by $\phi^*$ and $\mathbb{E}_{\phi^*}[R_n \mid P_1] = \sum_{j=1}^{n} a_{i,n+1} p_1(j)$. To prove the claim, note that the optimal conditional expected total reward, upon the arrival of first task, for the case with $n+1$ workers is given by

$$\begin{aligned}
&\mathbb{E}_{\phi^*}[R_{n+1} \mid X_1, P_1] \\
&\quad = \max_{i=1,\ldots,n+1} \{p_1(i)X_1 + \mathbb{E}_{\phi^*}[R_{n+1\setminus\{i\}} \mid X_1, P_1]\} \\
&\quad = \max_{i=1,\ldots,n+1} \{p_1(i)X_1 + \mathbb{E}_{\phi^*}[R_{n+1\setminus\{i\}} \mid P_1]\} \\
&\quad = \max_{i=1,\ldots,n+1} \{p_1(i)X_1 + \mathbb{E}_{\phi^*}[\mathbb{E}_{\phi^*}[R_{n+1\setminus\{i\}} \mid P_2] \mid P_1]\} \\
&\quad = \max_{i=1,\ldots,n+1} \left\{ p_1(i)X_1 + \mathbb{E}_{\phi^*}\left[ \sum_{j=1}^{i-1} a_{j,n+1} p_2(j) + \sum_{k=i}^{n} a_{k,n+1} p_2(k+1) \;\middle|\; P_1 \right] \right\} \\
&\quad = \max_{i=1,\ldots,n+1} \left\{ p_1(i)X_1 + \sum_{j=1}^{i-1} a_{j,n+1} p_1(j) + \sum_{k=i}^{n} a_{k,n+1} p_1(k+1) \right\},
\end{aligned}$$

where the fourth equality is a result of the induction assumption and the assumption that the order of worker success rates is fixed (i.e. $p_2(l)$ is an increasing function of $l$). The last equality follows from the martingale property. Recall that $\{p_1(l)\}_{l=1}^{n+1}$ is an increasing sequence of size $n+1$, and that $\{a_{m,n+1}\}_{m=1}^{n}$ is an increasing sequence of size $n$. According to Hardy's theorem, the optimal decision upon the arrival of the first task with value $X_1$ is to match $X_1$ with the $i$th best worker if it falls within the $i$th highest interval formed by $\{a_{m,n+1}\}_{m=1}^{n}$. Therefore,

$$\mathbb{E}_{\phi^*}[R_{n+1} \mid P_1] = \sum_{i=1}^{n+1} a_{i,n+2} p_1(i),$$

since $\{a_{i,n+2}\}_{i=1}^{n+1}$ is the set of ordered values of $X_1$ and $\{a_{m,n+1}\}_{m=1}^{n}$ (see (1)). □

**Remark 1.** In a similar fashion, model I.2 and the result of Theorem 2 can be extended to the case of incomplete information, where the actual worker values are unobservable to the decision maker; however, the decision maker knows the order of these values, which is assumed to be fixed throughout the problem. Specifically, the optimal assignment policy and the optimal expected total reward in this case is the same as that presented in Theorem 2.

In model I.3 we generalize model I.2 to the case where the sequence of arriving tasks is no longer i.i.d., while keeping all the other assumptions of model I.2. To obtain the optimal policy for this model, a set of breakpoints must be calculated upon the arrival of each task, and the value of that task is compared with these breakpoints to determine the index of the worker to

https://doi.org/10.1017/jpr.2016.64 Published online by Cambridge University Press

assign to the task. These breakpoints are computed in a recursive manner, as the breakpoints $\{a_{i,n}\}$ computed in (1) that characterize the optimal policy of Theorem 2. However, the set of breakpoints upon the arrival of task $j$ (for any arbitrarily fixed $j$) are random variables, the values of which are realized upon observing the value of $X_j$. These breakpoints, which make up the optimal assignment policy for model I.3, are computed as follows. Recall that $n$ indicates the total number of tasks to arrive. For any $r \in \{1, 2, \ldots, n+1\}$, define the random variable $A_{m,r}^{(n)}$ as

$$A_{m,r}^{(n)} = \begin{cases} -\infty, & m = 0, \\ \mathbb{E}[\tilde{A}_{m,r}^{(n)} \mid \mathcal{F}_{r-1}], & 1 \le m \le n-r+1, \\ +\infty, & m > n-r+1, \end{cases} \tag{5}$$

where $\tilde{A}_{m,r}^{(n)} = (X_r \vee A_{m-1,r+1}^{(n)}) \wedge A_{m,r+1}^{(n)}$. In the following theorem we present the optimal assignment policy with a threshold structure for model I.3 with random success rates, where the threshold values are computed iteratively via (5).

**Theorem 3.** *The optimal decision in model I.3 at time period $r$ for $r = 1, 2, \ldots, n$ is to assign the $i$th best worker to the $r$th task with value $X_r$ if it falls within the $i$th highest interval formed by the $\{A_{m,r+1}^{(n)}\}_{m=1}^{n-r}$, as computed in (5). The resulting optimal expected total reward is then given by $\mathbb{E}_{\phi^*}[R_n \mid X_1, P_1] = \sum_{i=1}^{n} p_1(i)\tilde{A}_{i,1}^{(n)}$.*

*Proof.* The proof is by induction on $n$. Starting from $n = 2$, the optimal conditional expected total reward after observing the first task and worker values is given by

$$\begin{aligned} &\mathbb{E}_{\phi^*}[R_2 \mid X_1, P_1] \\ &\quad = \max\{\mathbb{E}[p_1(1)X_1 + p_2(2)X_2 \mid X_1, P_1], \mathbb{E}[p_1(2)X_1 + p_2(1)X_2 \mid X_1, P_1]\} \\ &\quad = \max\{p_1(1)X_1 + \mathbb{E}[p_2(2)X_2 \mid X_1, P_1], p_1(2)X_1 + \mathbb{E}[p_2(1)X_2 \mid X_1, P_1]\} \\ &\quad = \max\{p_1(1)X_1 + \mathbb{E}[p_2(2) \mid P_1]\mathbb{E}[X_2 \mid X_1], p_1(2)X_1 + \mathbb{E}[p_2(1) \mid P_1]\mathbb{E}[X_2 \mid X_1]\} \\ &\quad = \max\{p_1(1)X_1 + p_1(2)\mathbb{E}[X_2 \mid X_1], p_1(2)X_1 + p_1(1)\mathbb{E}[X_2 \mid X_1]\} \\ &\quad = p_1(1)(X_1 \wedge A_{1,2}^{(2)}) + p_1(2)(X_1 \vee A_{1,2}^{(2)}) \\ &\quad = p_1(1)\tilde{A}_{1,1}^{(2)} + p_1(2)\tilde{A}_{2,1}^{(2)}, \end{aligned} \tag{6}$$

where the fourth equality is a direct result of the martingale property. Note that the optimal action to take upon the arrival of the first task with value $X_1$ is to compare its value with $A_{1,2}^{(2)}$ and assign it to the first worker if $X_1 = X_1 \wedge A_{1,2}^{(2)}$ and to the second worker otherwise. The last equality in (6) implies that $\tilde{A}_{m,1}^{(2)}$ is the expected value of the task assigned to the $m$th smallest worker, upon the arrival (and observing the value) of $X_1$, when there are a total of $n = 2$ tasks to arrive. Now suppose that the statement of Theorem 2 holds for some $n \ge 2$ and that the optimal assignment policy over the last $n$ stages is denoted by $\phi^*$. To prove the claim, note that the optimal conditional expected total reward for the case with $n+1$ workers, after observing the first task and worker values, is given by

$$\begin{aligned} &\mathbb{E}_{\phi^*}[R_{n+1} \mid X_1, P_1] \\ &\quad = \max_{i=1,\ldots,n+1} \{p_1(i)X_1 + \mathbb{E}_{\phi^*}[R_{n+1\setminus\{i\}} \mid X_1, P_1]\} \\ &\quad = \max_{i=1,\ldots,n+1} \{p_1(i)X_1 + \mathbb{E}_{\phi^*}[\mathbb{E}_{\phi^*}[R_{n+1\setminus\{i\}} \mid X_2, P_2] \mid X_1, P_1]\} \\ &\quad = \max_{i=1,\ldots,n+1} \left\{ p_1(i)X_1 + \mathbb{E}_{\phi^*}\left[ \sum_{j=1}^{i-1} p_2(j)\tilde{A}_{j,2}^{(n+1)} + \sum_{k=i}^{n} p_2(k+1)\tilde{A}_{k,2}^{(n+1)} \,\middle|\, X_1, P_1 \right] \right\} \end{aligned}$$

https://doi.org/10.1017/jpr.2016.64 Published online by Cambridge University Press

$$= \max_{i=1,\ldots,n+1} \left\{ p_1(i)X_1 + \sum_{j=1}^{i-1} p_1(j) A_{j,2}^{(n+1)} + \sum_{k=i}^{n} p_1(k+1) A_{k,2}^{(n+1)} \right\}$$

$$= \sum_{i=1}^{n+1} p_1(i) \tilde{A}_{i,1}^{(n+1)}, \tag{7}$$

where the last equality follows since $\{\tilde{A}_{m,1}^{(n+1)}\}_{m=1}^{n+1}$ is the set of ordered values of

$$\{X_1, A_{k,2}^{(n+1)} \text{ for all } k = 1, 2, \ldots, n\},$$

by definition. Note that $\{p_1(l)\}_{l=1}^{n+1}$ is an increasing sequence of size $n+1$, and that $\{A_{m,2}^{(n+1)}\}_{m=1}^{n}$ is an increasing sequence of size $n$. Therefore, according to Hardy's theorem, the optimal decision upon the arrival of the first task with value $X_1$ is to match it with the $i$th best worker if it falls within the $i$th highest interval formed by $\{A_{m,2}^{(n+1)}\}_{m=1}^{n}$. The last equality in (7) implies that $\tilde{A}_{m,1}^{(n+1)}$ is the expected value of the task assigned to the $m$th smallest worker, upon the arrival (and observing the value) of $X_1$, when there are a total of $n+1$ tasks to arrive. □

### 2.2. Model II

Models II and III relax the assumptions that the workers success rates are either identically distributed or have a fixed order. In model II we assume that the workers form two classes: one class includes the workers with a zero success rate and the other class includes a set of identical workers with nonzero success rates. Assume that there are $n$ i.i.d. sequentially arriving tasks (and, hence, $n$ remaining time periods), while $m$ ($\leq n$) identical workers (i.e. with the same success rate) with nonzero values are available to be assigned to these tasks. In model II.1 we consider the $m$ worker values to be deterministic and time dependent. Specifically, if the $j$th task with value $X_j$ arrives at time $t_j$ then the nonzero worker values at that time are given by $p(t_j)$. As a result, once a task arrives, the decision maker must decide whether to accept the task (i.e. assign it to one of the available workers with a nonzero value) or to reject it (i.e. assign it to a worker with value 0). Let $V_j^{(l)}(x)$ be the maximum conditional expected total reward obtained over the last $n-j+1$ time periods (specifically, time periods $j, j+1, \ldots, n$), when there are $l$ workers with nonzero values available upon the arrival of the $j$th task with value $X_j = x$. Moreover, define $V_j^{(l)} := \mathbb{E}[V_j^{(l)}(X)]$ for any $j = 1, 2, \ldots, n$ and $l = 1, 2, \ldots, n-j+1$. In the following theorem we present the optimal assignment policy for model II.1, which has a threshold structure.

**Theorem 4.** *There exist thresholds $y^{(m)}(t), y^{(m-1)}(t), \ldots, y^{(1)}(t)$ such that if the $j$th task arrives at time $t_j$ and $l$ workers are available for assignment at that time, then the optimal decision is to accept the $j$th task with value $X_j$ if and only if*

$$X_j \geq y^{(l)}(t_j), \tag{8}$$

*where*

$$y^{(l)}(t_j) := \frac{V_{j+1}^{(l)} - V_{j+1}^{(l-1)}}{p(t_j)}.$$

*Moreover, for any fixed $l = 1, 2, \ldots, m$, $\{V_j^{(l)}\}_{j=1}^{n-l+1}$ is computed recursively from $\{V_{\tilde{j}}^{(l-1)}\}_{\tilde{j}=1}^{n-l+2}$ by*

$$V_{n-l+1}^{(l)} = p(t_{n-l+1})\mathbb{E}[X] + V_{n-l+2}^{(l-1)}, \tag{9}$$

https://doi.org/10.1017/jpr.2016.64 Published online by Cambridge University Press

*and*

$$V_j^{(l)} = p(t_j)\mathbb{E}[X \vee y^{(l)}(t_j)] + V_{j+1}^{(l-1)} \quad \text{for } j = 1, 2, \ldots, n-l, \tag{10}$$

*with* $V_j^{(0)} := 0$ *for all* $j$.

*Proof.* To prove (9), note that at time period $n-l+1$, $l$ tasks have yet to arrive (specifically, with values $X_{n-l+1}, X_{n-l+2}, \ldots, X_n$), and considering that $l$ workers are available to be assigned to these arriving tasks, the best decision upon the arrival of each of these $l$ tasks is to accept the task and assign it to one of the available workers upon its arrival. Therefore, once the task with value $X_{n-l+1} = x$ arrives, it is assigned to a worker of value $p(t_{n-l+1})$, and $l-1$ tasks (specifically, with values $X_{n-l+2}, X_{n-l+3}, \ldots, X_n$) remain to be assigned to $l-1$ identical remaining workers. In other words,

$$V_{n-l+1}^{(l)}(x) = p(t_{n-l+1})x + V_{n-l+2}^{(l-1)},$$

and (9) follows. Proving (10) follows along the same lines; however, the number of remaining tasks in this case is not equal to the number of available workers. Therefore, upon the arrival of the $j$th task with value $X_j$ at time $t_j$, the decision maker must decide whether to accept the task or discard it. In mathematical form, this is shown by

$$V_j^{(l)}(x) = \max\{p(t_j)x + V_{j+1}^{(l-1)}, V_{j+1}^{(l)}\} = p(t_j)(x \vee y^{(l)}(t_j)) + V_{j+1}^{(l-1)},$$

which results in (8) and (10). Specifically, the $j$th task with value $X_j = x$ is accepted if $p(t_j)x + V_{j+1}^{(l-1)} \geq V_{j+1}^{(l)}$, which is equivalent to (8). Also, (10) follows since

$$V_j^{(l)} = \mathbb{E}[V_j^{(l)}(X)] = p(t_j)\mathbb{E}[X \vee y^{(l)}(t_j)] + V_{j+1}^{(l-1)}. \qquad \square$$

Observe that in order to obtain the optimal threshold values for a problem with $m$ available workers in model II.1, the threshold values for problems with $l = 1, 2, \ldots, m-1$ available workers must be computed in advance. Moreover, the threshold values computed in (8) are functions of arrival times, task distributions, the number of available workers, and worker values, as opposed to the models studied in the previous sections where the thresholds depend only on arrival times and task distributions.

Under the same setting and assumptions, model II.2 extends the results of model II.1 to the case where the nonzero identical worker values are random variables with time dependent probability density functions (PDFs) (i.e. $p(t_j) \sim f_j$). Specifically, the worker values at time period $j$ only become observable upon the arrival of task $j$ with value $X_j$. Moreover, task and worker values are no longer assumed to be independent. The following corollary is a generalization of the results of Theorem 4 to the case with random worker values.

**Corollary 1.** *There exist thresholds* $\tilde{y}^{(m)}(t), \tilde{y}^{(m-1)}(t), \ldots, \tilde{y}^{(1)}(t)$ *such that if the* $j$*th task arrives at time* $t_j$ *and* $l$ *workers with random values are available for assignment at that time, then the optimal decision is to accept the* $j$*th task with value* $X_j$ *if and only if*

$$W_j \geq \tilde{y}^{(l)}(t_j),$$

*where* $W_j := p(t_j)X_j$ *and* $\tilde{y}^{(l)}(t_j) := V_{j+1}^{(l)} - V_{j+1}^{(l-1)}$. *Moreover, for any fixed* $l = 1, 2, \ldots, m$, $\{V_j^{(l)}\}_{j=1}^{n-l+1}$ *is computed recursively from* $\{V_{\tilde{j}}^{(l-1)}\}_{\tilde{j}=1}^{n-l+2}$ *by*

$$V_{n-l+1}^{(l)} = \mathbb{E}[W_{n-l+1}] + V_{n-l+2}^{(l-1)},$$

https://doi.org/10.1017/jpr.2016.64 Published online by Cambridge University Press

*and*

$$V_j^{(l)} = \mathbb{E}[W_j \vee \tilde{y}^{(l)}(t_j)] + V_{j+1}^{(l-1)} \quad \textit{for } j = 1, 2, \ldots, n-l,$$

*with* $V_j^{(0)} := 0$ *for all* $j$.

*Proof.* The proof is an extension of that of Theorem 4 and is therefore omitted. □

### 2.3. Model III

In Subsection 2.2 we studied the SSAP with random worker values, where the workers form two distinct classes based on their success rates: one class includes workers with a success rate of 0 and the other class consists of identical workers with the same success rate. In this subsection we extend this problem to the case where workers are allowed to form two or more classes with nonzero success rates and worker values at a given time are derived from distinct distribution functions. Specifically, in model III.1 we study an SSAP with $n = 2$, where two tasks arrive sequentially to be assigned to two workers, with random values $\alpha$ and $\beta$. Worker values at time period $j$ are denoted by $\alpha_j \sim f_\alpha$ and $\beta_j \sim f_\beta$, where $f_\alpha$ and $f_\beta$ are distinct PDFs. As assumed in the previous sections, $\alpha_j$ and $\beta_j$ are unobservable until the arrival of the $j$th task, and take on new values with each task arrival. The goal, as in previous sections, is to assign the arriving tasks to the workers with values $\alpha$ and $\beta$ so as to maximize the expected total reward. A difficulty in performing assignments in model III.1 arises since both the task and worker values are random. Moreover, worker values are not assumed to be identically distributed anymore, and, hence, no fixed ranking of worker values exists in advance. Recall from (1) that $a_{1,2} := \mathbb{E}[X]$, $a_{1,3} := \mathbb{E}[X \wedge a_{1,2}]$, and $a_{2,3} := \mathbb{E}[X \vee a_{1,2}]$. For an arbitrary random variable $\gamma$, let $a_{i,n}^{(\gamma)}$ denote the breakpoint $a_{i,n}$ defined by Equation (1) of [6], under the assumption that the sequentially arriving task values are given by $\gamma X$ (instead of $X$). In the following theorem we present the optimal assignment policy for model III.1.

**Theorem 5.** *The optimal decision for model III.1 upon the arrival of the first task is to assign it to the worker with value* $\beta$ *if and only if*

$$W_1^{(\beta-\alpha)} \geq a_{1,2}^{(\beta-\alpha)}, \tag{11}$$

*where* $W^{(\gamma)} := \gamma X$ *for any random variable* $\gamma$. *The optimal expected total reward obtained under this policy is given by*

$$V^{(\alpha,\beta)} := V_1^{(2)} = 2\,a_{1,2}^{(\alpha)} + a_{2,3}^{(\beta-\alpha)}. \tag{12}$$

*Proof.* Observe that

$$V_1^{(2)}(x) = \max\{\alpha_1 x + \mathbb{E}[W^{(\beta)}],\, \beta_1 x + \mathbb{E}[W^{(\alpha)}]\},$$

and, hence, it is optimal to choose the worker with value $\beta$ upon the arrival of the first task if and only if $\beta_1 x + \mathbb{E}[W^{(\alpha)}] \geq \alpha_1 x + \mathbb{E}[W^{(\beta)}]$, which is equivalent to (11). To compute the expected total reward, let $F$ denote the cumulative distribution function of $W^{(\beta-\alpha)}$, with $\bar{F} := 1 - F$, and note that

$$\begin{aligned}
V_1^{(2)} &= F(a_{1,2}^{(\beta-\alpha)})\mathbb{E}[W^{(\alpha)} + \mathbb{E}[W^{(\beta)}] \mid W^{(\beta-\alpha)} \leq a_{1,2}^{(\beta-\alpha)}]] \\
&\quad + \bar{F}(a_{1,2}^{(\beta-\alpha)})\mathbb{E}[W^{(\beta)} + \mathbb{E}[W^{(\alpha)}] \mid W^{(\beta-\alpha)} > a_{1,2}^{(\beta-\alpha)}] \\
&= F(a_{1,2}^{(\beta-\alpha)})\mathbb{E}[W^{(\beta)}] + \bar{F}(a_{1,2}^{(\beta-\alpha)})\mathbb{E}[W^{(\alpha)}] + F(a_{1,2}^{(\beta-\alpha)})\mathbb{E}[W^{(\alpha)} \mid W^{(\beta-\alpha)} \leq a_{1,2}^{(\beta-\alpha)}] \\
&\quad + \bar{F}(a_{1,2}^{(\beta-\alpha)})\mathbb{E}[W^{(\beta-\alpha)} \mid W^{(\beta-\alpha)} > a_{1,2}^{(\beta-\alpha)}] \\
&\quad + \bar{F}(a_{1,2}^{(\beta-\alpha)})\mathbb{E}[W^{(\alpha)} \mid W^{(\beta-\alpha)} > a_{1,2}^{(\beta-\alpha)}]
\end{aligned}$$

https://doi.org/10.1017/jpr.2016.64 Published online by Cambridge University Press

$$
\begin{aligned}
&= F(a_{1,2}^{(\beta-\alpha)})a_{1,2}^{(\beta-\alpha)} + \mathbb{E}[W^{(\alpha)}] + F(a_{1,2}^{(\beta-\alpha)})\mathbb{E}[W^{(\alpha)} \mid W^{(\beta-\alpha)} \le a_{1,2}^{(\beta-\alpha)}] \\
&\quad + \bar{F}(a_{1,2}^{(\beta-\alpha)})\mathbb{E}[W^{(\alpha)} \mid W^{(\beta-\alpha)} > a_{1,2}^{(\beta-\alpha)}] \\
&\quad + \bar{F}(a_{1,2}^{(\beta-\alpha)})\mathbb{E}[W^{(\beta-\alpha)} \mid W^{(\beta-\alpha)} > a_{1,2}^{(\beta-\alpha)}] \\
&= 2\mathbb{E}[W^{(\alpha)}] + F(a_{1,2}^{(\beta-\alpha)})a_{1,2}^{(\beta-\alpha)} + \bar{F}(a_{1,2}^{(\beta-\alpha)})\mathbb{E}[W^{(\beta-\alpha)} \mid W^{(\beta-\alpha)} > a_{1,2}^{(\beta-\alpha)}] \\
&= 2\mathbb{E}[W^{(\alpha)}] + \mathbb{E}[W^{(\beta-\alpha)} \vee a_{1,2}^{(\beta-\alpha)}] \\
&= 2a_{1,2}^{(\alpha)} + a_{2,3}^{(\beta-\alpha)}. \qquad \square
\end{aligned}
$$

Now consider the following SSAP (labeled S1). Two tasks arrive sequentially to be assigned to the following two workers: a worker with a random success rate of $\beta - \alpha$ and a worker with value 0. To obtain the maximum expected total reward, $X_1$ is matched with $\beta - \alpha$ if and only if $W_1^{(\beta-\alpha)} \ge \mathbb{E}[W^{(\beta-\alpha)}]$, which is similar to (11). This motivates us to break model III.1 into two separate SSAPs. The first SSAP is S1, and the second SSAP (labeled S2) is a problem with two tasks and two identical workers with random success rate $\alpha$. Suppose that the stream of arriving tasks is the same for S1 and S2. Specifically, when the first task with value $X_1 = x_1$ arrives, it is automatically assigned to the worker with success rate $\alpha$ in S2 (since workers are identical). For S1, $X_1 = x_1$ is matched with $\beta - \alpha$ if (11) is satisfied; otherwise, it is assigned to the worker with success rate 0 (i.e. rejected). When the second task with value $X_2 = x_2$ is observed, it is again assigned to the worker with value $\alpha$ in S2. For S1, $X_2 = x_2$ is matched with the last remaining worker (with value $\beta - \alpha$ or 0, depending on the decision made at time period 1). The optimal expected total reward is equal to $\mathbb{E}[W^{(\beta-\alpha)} \vee \mathbb{E}[W^{(\beta-\alpha)}]]$ and $2\mathbb{E}[W^{(\alpha)}]$ for S1 and S2, respectively. Therefore, solving S1 and S2 simultaneously yields an overall optimal expected total reward equal to that of model III.1, as denoted in (12). In other words, one can break model III.1 into two simpler SSAPs, S1 and S2 (recall that the optimal assignment policy for S1 is the same as that of model III.1).

Model III.2 extends this setting to a problem of size $n = 3$, where three tasks arrive sequentially to be assigned to two workers with random success rates $\alpha$ and $\beta$, and one worker with fixed value 0 (assigning a task to this worker is equivalent to rejecting that task). We define a new set of breakpoints as follows, which characterize the optimal assignment policy for model III.2 in Theorem 6 below:

$$
\tilde{a}_{1,3}^{(\alpha)} := a_{1,3}^{(\beta)} - a_{1,3}^{(\beta-\alpha)}, \qquad \tilde{a}_{1,3}^{(\beta)} := a_{1,3}^{(\alpha)} - a_{1,3}^{(\alpha-\beta)}, \qquad \tilde{a}_{2,3}^{(\alpha-\beta)} := a_{2,3}^{(\alpha)} - a_{2,3}^{(\beta)}.
$$

**Theorem 6.** *Upon the arrival of $X_1$ in model III.2, it is optimal to assign the first task with value $X_1$ to the worker with success rate $l \in \{\alpha, \beta, 0\}$, if event $\mathcal{C}_l$ happens, where*

$$
\begin{aligned}
\mathcal{C}_\alpha &:= \{W^{(\alpha)} > \tilde{a}_{1,3}^{(\alpha)},\ W^{(\alpha-\beta)} > \tilde{a}_{2,3}^{(\alpha-\beta)}\}, \\
\mathcal{C}_\beta &:= \{W^{(\beta)} > \tilde{a}_{1,3}^{(\beta)},\ W^{(\alpha-\beta)} \le \tilde{a}_{2,3}^{(\alpha-\beta)}\}, \\
\mathcal{C}_0 &:= \{W^{(\alpha)} \le \tilde{a}_{1,3}^{(\alpha)},\ W^{(\beta)} \le \tilde{a}_{1,3}^{(\beta)}\},
\end{aligned}
$$

*and the resulting optimal expected total reward is given by*

$$
V_1^{(3)} = (\mathbb{E}[W^{(\alpha)} \mid \mathcal{C}_\alpha] + V^{(0,\beta)})P\{\mathcal{C}_\alpha\} + (\mathbb{E}[W^{(\beta)} \mid \mathcal{C}_\beta] + V^{(0,\alpha)})P\{\mathcal{C}_\beta\} + V^{(\alpha,\beta)}P\{\mathcal{C}_0\}.
$$

*Proof.* Observe that if we opt to assign the first task with value $X_1 = x$ to the worker with success rate $\alpha$, then we are left with two workers (with success rates $\beta$ and 0) and two

https://doi.org/10.1017/jpr.2016.64 Published online by Cambridge University Press

arriving tasks. The optimal expected total reward for this problem over the last two remaining time periods is $a_{2,3}^{(\beta)} = \mathbb{E}[W^{(\beta)} \vee \mathbb{E}[W^{(\beta)}]]$. Comparing this value with the expression in (12) reveals that it is equal to $V^{(0,\beta)}$. A similar argument can be made for the other two cases where $X_1$ is assigned to worker $\beta$ or to the worker with value 0. It follows that

$$V_1^{(3)}(x) = \max\{\alpha x + V^{(0,\beta)}, \beta x + V^{(0,\alpha)}, V^{(\alpha,\beta)}\}.$$

It is optimal to reject the first task with value $X_1$ (i.e. to assign it to the worker with value 0) if $V^{(\alpha,\beta)} \geq \alpha x + V^{(0,\beta)}$ and $V^{(\alpha,\beta)} \geq \beta x + V^{(0,\alpha)}$, which are equivalent to the occurrence of the event $\mathcal{C}_0$. To see why this is the case, note that an argument analogous to that presented in Theorem 5 leads to

$$V^{(\alpha,\beta)} = 2a_{1,2}^{(\beta)} + a_{2,3}^{(\alpha-\beta)},$$

where $2a_{1,2}^{(\beta)} = a_{1,3}^{(\beta)} + a_{2,3}^{(\beta)}$ and $a_{2,3}^{(\alpha-\beta)} = -a_{1,3}^{(\beta-\alpha)}$. Therefore, $V^{(\alpha,\beta)} \geq \alpha x + V^{(0,\beta)}$ is equivalent to $W^{(\alpha)} \leq \tilde{a}_{1,3}^{(\alpha)}$. In a similar fashion, it is proven that $V^{(\alpha,\beta)} \geq \beta x + V^{(0,\alpha)}$ is the same as $W^{(\beta)} \leq \tilde{a}_{1,3}^{(\beta)}$.

It is optimal to assign the first task with value $X_1$ to the worker with value $\alpha$ if

$$\alpha x + V^{(0,\beta)} \geq V^{(\alpha,\beta)}, \qquad \alpha x + V^{(0,\beta)} \geq \beta x + V^{(0,\alpha)},$$

or, equivalently,

$$\alpha x \geq V^{(\alpha,\beta)} - V^{(0,\beta)}, \qquad (\alpha - \beta)x \geq V^{(0,\alpha)} - V^{(0,\beta)}. \tag{13}$$

From the argument made for the first case (i.e. rejecting $X_1$), it follows that $V^{(\alpha,\beta)} = a_{1,3}^{(\beta)} + a_{2,3}^{(\beta)} - a_{1,3}^{(\beta-\alpha)}$. Therefore, the first inequality in (13) simplifies to

$$\alpha x \geq V^{(\alpha,\beta)} - V^{(0,\beta)} = (a_{1,3}^{(\beta)} + a_{2,3}^{(\beta)} - a_{1,3}^{(\beta-\alpha)}) - a_{2,3}^{(\beta)} = a_{1,3}^{(\beta)} - a_{1,3}^{(\beta-\alpha)} = \tilde{a}_{1,3}^{(\alpha)}, \tag{14}$$

and the second inequality in (13) is re-written as

$$(\alpha - \beta)x \geq V^{(0,\alpha)} - V^{(0,\beta)} = a_{2,3}^{(\alpha)} - a_{2,3}^{(\beta)} = \tilde{a}_{2,3}^{(\alpha-\beta)}. \tag{15}$$

Note that (14) and (15) are equivalent to the event $\mathcal{C}_\alpha$, which implies that it is optimal to pick the worker with value $\alpha$ upon the arrival of the first task if the event $\mathcal{C}_\alpha$ occurs. The rest of the proof follows along the same lines. □

Unlike model III.1 and judging from the optimal policy presented in Theorem 6, model III.2 cannot be solved by breaking the problem down into a set of smaller SSAPs. Specifically, the points characterizing the events $\mathcal{C}_\alpha$, $\mathcal{C}_\beta$, and $\mathcal{C}_0$ are functions of the breakpoints defined by Equation (1) of [6], and computing these points involves solving stochastic sequential assignment problems for task values given by $\alpha X$, $\beta X$, and $(\alpha - \beta)X$. However, $\tilde{a}_{1,3}^{(\alpha)}$, $\tilde{a}_{1,3}^{(\beta)}$, and $\tilde{a}_{2,3}^{(\alpha-\beta)}$ are not equal to the $\{a_{i,n}\}$ calculated for SSAPs of a smaller size. Another issue that differentiates the optimal policy of this problem from that of a SSAP with deterministic success rates is that the ordering of the points $\tilde{a}_{1,3}^{(\alpha)}$, $\tilde{a}_{1,3}^{(\beta)}$, and $\tilde{a}_{2,3}^{(\alpha-\beta)}$ is dependent on the specific task and worker value distributions, and, hence, varies from one problem to another, as opposed to the sequence $\{a_{i,n}\}$ that depends only on the task distributions and is proven to be a nondecreasing function of $i$. We can solve problems of size $n \geq 3$ using a similar approach to that demonstrated by Theorem 6; however, no specific pattern can be found which relates the optimal breakpoints of a problem of size $n$ to that of a problem of size $n + 1$. Therefore, the problem can also be modeled as a Markov decision process and solved numerically using backward optimality equations.

https://doi.org/10.1017/jpr.2016.64 Published online by Cambridge University Press

## 3. Conclusion

The existing SSAP literature studies variations of the problem under the general assumption that worker values are deterministic and fixed numbers. In this paper we extended the SSAP model to the case where worker success rates are observable random variables, taking on new values upon each task arrival. Various models of the SSAP with random success rates are studied with different assumptions regarding the success rates' distribution and their dependency on one another. Optimal assignment policies that maximize the expected total reward are presented for each model. These models are categorized into three classes as follows: (i) an exchangeable set of success rates at a given time period, with worker success rates at a given time being dependent on that of preceeding time periods; (ii) time-dependent random success rates which form two distinct worker classes (based on their values), with one class containing workers with value 0; and (iii) time-dependent random success rates with distinct generic distributions, forming two or more than two worker classes.

Further research is required to address a more general version of the SSAP with random success rates. One challenge is to study this problem for task arrival times that are no longer deterministic; e.g. tasks arriving according to a Poisson process. Another research direction is the case where worker values are updated at each task arrival according to the realized task value (i.e. worker and task values are dependent on one another). Studying this problem under risk sensitive objective functions, such as the threshold criteria, is another possible extension.

## Acknowledgement

This research has been supported in part by the Air Force Office of Scientific Research under grant $FA9550-15-1-0100$. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views the Air Force Office of Scientific Research.

## References

[1] Albright, S. C. (1974). A Markov-decision-chain approach to a stochastic assignment problem. *Operat. Res.* **22,** 61–64.
[2] Albright, S. C. (1974). Optimal sequential assignments with random arrival time. *Manag. Sci.* **21,** 60–67.
[3] Albright, S. C. (1977). A Bayesian approach to a generalized house selling problem. *Manag. Sci.* **24,** 432–440.
[4] Albright, S. C. and Derman, C. (1972). Asymptotic optimal policies for the stochastic sequential assignment problem. *Manag. Sci.* **19,** 46–51.
[5] Baharian, G. and Jacobson, S. H. (2013). Limiting behavior of the stochastic sequential assignment problem. *Naval Res. Logistics* **60,** 321–330.
[6] Derman, C., Lieberman, G. J. and Ross, S. M. (1972). A sequential stochastic assignment problem. *Manag. Sci.* **18,** 349–355.
[7] Khatibi, A., Baharian, G., Behzad, B. and Jacobson, S. H. (2015). Extensions of the sequential stochastic assignment problem. *Math. Meth. Operat. Res.* **82,** 317–340.
[8] Khatibi, A., Baharian, G., Kone, E. R. and Jacobson, S. H. (2014). The sequential stochastic assignment problem with random success rates. *IIE Trans.* **46,** 1169–1180.
[9] Nikolaev, A. G. and Jacobson, S. H. (2010). Stochastic sequential decision-making with a random number of jobs. *Operat. Res.* **58,** 1023–1027.
[10] Nikolaev, A. G., Jacobson, S. H. and McLay, L. A. (2007). A sequential stochastic security system design problem for aviation security. *Transportation Sci.* **41,** 182–194.
[11] Righter, R. (1987). The stochastic sequential assignment problem with random deadlines. *Prob. Eng. Inf. Sci.* **1,** 189–202.
[12] Su, X. and Zenios, S. A. (2005). Patient choice in kidney allocation: a sequential stochastic assignment model. *Operat. Res.* **53,** 443–455.

https://doi.org/10.1017/jpr.2016.64 Published online by Cambridge University Press