

STOCHASTIC SEQUENTIAL ASSIGNMENT PROBLEM WITH ARRIVALS

RHONDA RIGHTER

*Department of Industrial Engineering and Operations Research
University of California
Berkeley, CA 94720
E-mail: RRighter@IEOR.Berkeley.edu*

We extend the classic sequential stochastic assignment problem to include arrivals of workers. When workers are all of the same type, we show that the socially optimal policy is the same as the individually optimal policy for which workers are given priority according to last come–first served. This result also holds under several variants in the model assumptions. When workers have different types, we show that the socially optimal policy is determined by thresholds such that more valuable jobs are given to more valuable workers, but now the individually optimal policy is no longer socially optimal. We also show that the overall value increases when worker or job values become more variable.

1. INTRODUCTION AND SUMMARY

We consider the sequential stochastic assignment problem, as originally introduced by Derman, Lieberman, and Ross (DLR) [5], in which a set of workers of different values are assigned sequentially arriving jobs with random values, where the value of an assignment of a job to a worker is the product of the job and worker values and where the objective is to maximize the sum of the assignment values. We extend their analysis to permit arrivals of workers as well as of jobs. Jobs must be assigned or rejected immediately upon arrival, without recall, but arriving workers stay until they are assigned jobs.

In the classical DLR model, there are n workers, where worker i has value p_i , $p_1 \geq p_2 \geq \dots \geq p_n$, and where $m \geq n$ jobs will arrive sequentially and must be immediately either assigned to a worker or rejected. Job values X are independent and identically distributed (i.i.d.) random variables revealed upon arrival, and if a job of

value x is assigned to a worker with value p , the reward is px . They show that there are thresholds $\infty = t_{0,m} \geq t_{1,m} \geq \dots \geq t_{n,m} \geq t_{n+1}$, $m = 0$, such that it is optimal to assign an arriving job of value x to worker i if $t_{i-1} < x \leq t_i$ and where worker $n + 1$ corresponds to rejection. Moreover, the thresholds are independent of the values of the p_i s, and threshold i can be interpreted as the expected job value for worker i , assuming each worker follows an individually optimal policy of accepting or rejecting jobs, where jobs are offered to workers in descending order of their values (so worker i has the i th highest priority). We will consider extensions to a slight modification of the DLR model in which there are an infinite number of jobs arriving according to a Poisson process with rate λ and rewards are exponentially discounted with rate α . Again, the same threshold result applies, but now the thresholds can be simply written as t_i [2]. We extend this model to permit worker arrivals.

Since the sequential stochastic assignment problem was introduced, it has been extended in numerous directions, such as having random arrival times of jobs and arbitrary discounting [2], having a random number of jobs with nonidentically distributed values [8], developing asymptotic results [10], and allowing workers with different skill sets [1]. The results have been applied in many contexts, including kidney transplantation [3,4,11], aviation security [9], buying decisions in supply chains [12], and real estate [10]. The only work we know of in which there are arrivals of workers (patients) as well as jobs (kidneys) is that of David and Yechiali [3], in which pairs of patients and kidneys arrive and there are two types of patients and kidneys. If a patient is assigned a matching kidney, a reward R is earned; the reward for assigning a kidney that does not match is $r < R$. They show that to maximize the long-term average award, only matches should be assigned and the corresponding long-term average reward is R . When rewards are discounted or there is a finite horizon, the optimal policy is a threshold policy such that a mismatch is assigned only if there are no matches and the number of (nonmatching) patients exceeds some threshold. Our work differs by permitting arrivals that are not in pairs and an arbitrary number of worker and job types, and our reward structure for assignments is different.

We first consider the problem in which workers are identical (also known as the house selling problem) and they arrive according to a Poisson process with rate γ and leave only after being assigned a job. We show that, as in the classical case with no worker arrivals, the optimal policy can be implemented by allowing each worker to accept or reject his or her own job, subject to a last come–first served (LCFS) priority among the workers; that is, a worker will only be offered a job that has been rejected by all workers who have arrived after that worker. This means that the socially optimal policy is a threshold policy, such that a job is assigned to a worker only if its value is above a threshold that is decreasing in the number of workers presently available. This result also holds when there is a finite buffer for workers, workers are impatient and leave after an exponential time (which would apply, e.g., in the kidney allocation application), there is a random environment that determines worker and job arrival rates and discount rate or abandonment rate, there are holding costs associated with waiting workers, there are batch arrivals of workers and jobs, and the job arrival process is an arbitrary renewal process.

We then consider the heterogeneous worker case, for which a worker of type i has value p_i , $p_1 \geq p_2 \geq \dots \geq p_n$, and we show that the optimal policy is again a threshold policy, where in state $n = (n_1, n_2, \dots, n_K)$, representing the current number of available workers of each type, a job of value x is assigned to a worker of type i if $t_i(n) \leq x < t_{i-1}(n)$, with $t_0(n) = \infty$ for all n , and where the thresholds are increasing in each of the numbers of the different types of workers. However, unlike the case with no arriving workers, these thresholds are not the same as would be obtained for an individually optimal policy where within each class, workers that arrived later have priority and where workers with higher values have higher priority over other workers. Moreover, the thresholds depend on the p_i s and on the numbers of each type of worker present, not just the total number present. We show that the overall value in the heterogeneous worker case increases if the worker values and/or job values become more variable (to be defined precisely latter).

2. IDENTICAL WORKERS

We first suppose we have identical workers, that workers and jobs arrive according to Poisson processes with respective rates γ and λ , and that rewards are exponentially discounted with rate α . Job values are i.i.d. All of these assumptions will be relaxed later.

We first consider an individually optimal policy, called the IO policy, in which workers are given priority in inverse order of their arrival times (or seniority); so the worker that arrived most recently has first right of refusal for jobs, and workers try to maximize their own expected job values. Note that this priority has the effect that the decisions of individual workers have no impact on future workers. Label the workers so that worker i is the i th most recently arriving worker and therefore will be the i th to be offered an arriving job if a job arrives before another worker. The IO policy is the Nash equilibrium of a noncooperative game; that is, it is such that if all workers follow the IO policy and each worker is trying to maximize its own expected job value, then no worker will have incentive to deviate from the IO policy.

LEMMA 2.1: *The IO policy is unique (except for the indifference of accepting or rejecting jobs with values equal to a threshold) and is determined by thresholds t_i such that worker i should accept a job of value x if $x \geq t_i$, where $t_1 > t_2 > \dots$ and where, for each i , t_i is the expected discounted job value that will be assigned to worker i under the IO policy, so the t_i s satisfy the following equations, for $i = 1, 2, \dots$, where $t_0 = \infty$:*

$$t_i = \gamma t_{i+1} + \lambda [t_i P(X < t_i) + E[X | t_{i-1} > X \geq t_i] P(t_{i-1} > X \geq t_i) + t_{i-1} P(X \geq t_{i-1})]. \quad (1)$$

PROOF: We consider the equivalent discrete-time problem by uniformizing with uniformization rate $\lambda + \gamma + \alpha = 1$, and we use value iteration (i.e., induction on a finite time horizon, n). Let t_i^n be the discounted expected job value (henceforth we will

abbreviate this as the EJV) for the i th worker under the IO policy when the time horizon is n , and there is not a job currently available so $t_i^0 \equiv 0$, and let $t_i^n(x)$ be the EJV for worker i under the IO policy given that a job has just arrived with value x and it is offered to worker i and there are n more steps to go. Let $t_0^n = \infty$ for all n . We will show the following by induction on n :

- (i) $t_i^n = 0$ for $i > n$.
- (ii) $t_i^n = \gamma t_{i+1}^{n-1} + \lambda [t_i^{n-1} P(X < t_i^{n-1}) + E[X | t_{i-1}^{n-1} > X \geq t_i^{n-1}] P(t_{i-1}^{n-1} > X \geq t_i^{n-1}) + t_{i-1}^{n-1} P(X \geq t_{i-1}^{n-1})]$ for $i \leq n$.
- (iii) $t_{i-1}^n > t_i^n$ for all $2 \leq i \leq n + 1$.
- (iv) It is optimal for worker i to accept a job of value x that is offered to it when there are n more steps to go if $x \geq t_i^n$, for any i .

Suppose (i)–(iv) are true for $n = 0, \dots, k$ and consider $k + 1$. Then (ii) follows from the induction hypothesis for (iv), (i) follows from the induction hypotheses for (i) and (iv) (no jobs will ever be offered to worker i for $i > k + 1$), (iii) follows from the induction hypotheses for (ii) and (iii), and (iv) follows from the dynamic programming equation $t_i^n(x) = \max\{x, t_i^n\}$. We can rewrite (ii) as

$$t_i^n = \gamma t_{i+1}^{n-1} + \lambda E \max\{t_i^{n-1}, \min[X, t_{i-1}^{n-1}]\}$$

from which it is easy to see by induction that for each $i, t_i^n, n = 0, 1, \dots$, is an increasing sequence. Since it is also bounded, the result follows in the limit. Note also that the IO policy is a Nash equilibrium, because, under our priority scheme, a worker’s actions has no impact on the job values it is offered, so it has no incentive to deviate from the IO policy. ■

THEOREM 2.2: *The IO policy is socially optimal (it maximizes the total discounted return) when all workers are identical. Therefore, the socially optimal policy is a threshold policy.*

PROOF: Again we consider the equivalent discrete-time problem through uniformization. Our proof is by policy iteration; we show that following the IO policy starting with the first decision, at time 0, is better (for all workers) than following an arbitrary policy for the first decision and then switching to the IO policy thereafter (call the latter policy π). Suppose there are n workers and a job of value x arrives at time 0, and first suppose that $t_{i-1} > x \geq t_i$ for some $i \leq n$, but that π rejects the job. Thus, the IO policy assigns the job to worker i and follows the IO policy thereafter. From time 1 on, the EJVs for workers $1, \dots, i - 1$ and all future workers will be the same under both IO and π , but the EJVs for workers $k, k = i + 1, \dots, n$ will be $t_{k-1} \geq t_k$, which will be the EJV under π . Finally, the EJV for worker i under IO will be $x \geq t_i$, its EJV under π . Hence, all workers will be better off under IO than under the alternative policy. Now, consider the case for which $x < t_n$, and π assigns a job to some worker. Let us label that worker worker n and label the other workers $1, \dots, n - 1$. Then,

under π , worker n will have the lowest priority, and since π follows IO from time 1 on, worker n will have no effect on any other workers. Thus, all workers besides worker n will have the same EJVs under π and under IO, but worker i 's EJV under IO will be $t_n > x$, the EJV under π . Again, all workers are better off under IO than under π , so the IO policy cannot be improved upon in a social sense and therefore it is also socially optimal. ■

It is not hard to show that the thresholds are decreasing in γ and increasing in λ and the overall value is increasing in both. The following extensions are also not hard to show. We could have a finite waiting space for the workers. Instead of an exponential discount rate, we could have each worker leave the system after an exponential time with rate α , although in this case we would need an arbitrary upper bound on the number of workers permitted in the system in order to do the uniformization. We could have arbitrary Markov modulated arrival processes for workers and jobs and a randomly varying discount rate or abandonment rate, although now the thresholds would depend on the environmental state modulating the arrivals, discount rate, and abandonment rate. We could have costs associated with waiting, so that now the EJV for a worker would be net of its waiting cost. We could have batch arrivals of workers and jobs with random independent batch sizes (i.e., compound Poisson processes) as long as workers are given an arbitrary priority order within a batch of workers (but later arrivals still have priority over later arrivals). In this case, the optimal policy is to start with the job with the largest value in an arriving job batch and offer it sequentially to the workers according to their priorities until it is accepted by one or rejected by all of them, and then repeat with the next largest job value and so on (once a job is rejected, all jobs of lesser value will be rejected). We could also have a general (batch) renewal process for job arrivals, which would, in effect, create batches of workers arriving between job arrivals. Under all of these extensions, Theorem 2.2 would hold; that is, the socially optimal policy would correspond to the IO policy for which worker priorities are assigned on a LCFS basis, and it would be a threshold policy. The details are left to the reader.

3. HETEROGENEOUS WORKERS

Now, consider the problem with K worker types in which workers of type i have value p_i , $p_1 \geq p_2 \geq \dots \geq p_K$, and the total reward for assigning a job of value x to a worker with value p is px . For a fixed policy π , let $W^\pi(n, \gamma, p)$ be the total expected discounted return under π , where $n = (n_1, \dots, n_K)$, n_k is the number of type k workers present at time 0. Let $\gamma = (\gamma_1, \dots, \gamma_K)$, where γ_k is the arrival rate of type k workers, and let p be the vector of worker values. When $\gamma_i \equiv 0$, the IO policy is still optimal, as long as the workers with the higher p_i are given higher priority (otherwise it does not depend on the values of the p_i s) [5]. We include a slightly different proof for completeness. Let $t_i^\pi(n, \gamma)$ be the EJV for worker i under π when we start with n workers and the workers are identical (i.e., $p_i \equiv 1$) and γ is the total worker arrival rate, and let $V_k^\pi(n, \gamma) = \sum_{i=1}^k t_i(n, \gamma)$ be the total expected discounted return for the first k workers.

COROLLARY 3.1: *When $\gamma_i \equiv 0$, the IO policy is socially optimal for heterogeneous workers.*

PROOF: Since there are no worker arrivals, suppose, without loss of generality, that we start with one worker of each type, $n_k = 1$ for $k = 1, \dots, K$, and $K = n$. We know from Theorem 2.2 that for any i and K , the IO policy maximizes $V_i^\pi(K, 0)$. It therefore maximizes $W^\pi(K, 0, p)$ because we can write $W^\pi(K, 0, p)$ as follows, where $p_{K+1} = 0$:

$$\begin{aligned} W^\pi(K, 0, p) &= \sum_{i=1}^K p_i t_i^\pi(K, 0) = \sum_{i=1}^K (p_i - p_{i+1}) \sum_{j=1}^i t_j^\pi(K, 0) \\ &= \sum_{i=1}^K (p_i - p_{i+1}) V_i^\pi(K, 0). \end{aligned} \tag{2}$$

■

When there are arrivals, the IO policy is no longer socially optimal. With worker arrivals, we define the IO policy so that type i workers always have priority over type $j > i$ workers, and within type, workers have priority according to LCFS. Let $N_k = \sum_{j=1}^k n_j$ and $\Gamma_k = \sum_{j=1}^k \gamma_j$, and let $T_i^{IO}(\Gamma_{i-1}, \gamma_i)$ be the total expected discounted job value (EJV) for all future arrivals of type i under the IO policy. Let $t_l^{IO}(\Gamma_i)$ be the EJV under the IO policy for the l th customer initially present, assuming it is type i , $l = N_{i-1} + 1, N_{i-1} + 2, \dots, N_i$. Then

$$W^{IO}(n, \gamma, p) = \sum_{i=1}^K p_i \left(\sum_{l=N_{i-1}+1}^{N_i} t_l^{IO}(\Gamma_i) + T_i^{IO}(\Gamma_{i-1}, \gamma_i) \right). \tag{3}$$

Note that the IO policy cannot be socially optimal. For example, for $K = 2$, $p_1 = p_2$, and $n_2 = 0$, the socially optimal policy, from Theorem 2.2, is to assign an arriving job of value x to a type 1 worker if and only if $x \geq t_{n_1}(\gamma_1 + \gamma_2)$, whereas the IO policy is to assign an arriving job of value x to a type 1 worker if and only if $x \geq t_{n_1}(\gamma_1) \neq t_{n_1}(\gamma_1 + \gamma_2)$ (unless $\gamma_2 = 0$). (The latter would be the socially optimal policy when $p_1 > p_2 = 0$.) In general, the socially optimal policy will depend on the values of the p_i s.

PROPOSITION 3.2: *For heterogeneous workers, the IO policy is not, in general, socially optimal.*

Let $W(n, \gamma, p)$ ($W(n, \gamma, p|x)$) be the maximal expected discounted return under the optimal policy for the heterogeneous worker model starting with the vector n of numbers of each worker type at a random time (when a job of value x has just arrived). Let $n^{-i} = (n_1, \dots, n_{i-1}, n_i - 1, n_{i+1}, \dots, n_K)$ and $n^{+i} = (n_1, \dots, n_{i-1}, n_i + 1, n_{i+1}, \dots, n_K)$. Then

$$W(n, \gamma, p|x) = \max\{W(n, \gamma, p); \max_{i=1, \dots, K} I\{n_i > 0\}(p_i x + W(n^{-i}, \gamma, p))\} \tag{4}$$

and

$$W(n, \gamma, p) = \sum_{i=1}^K \gamma_i W(n^{+i}, \gamma, p) + \lambda EW(n, \gamma, p|X). \tag{5}$$

Additionally, from our previous result,

$$W(n, \gamma, \mathbf{1}) = W_{N_K}^{IO}(N_K, \Gamma_K, 1). \tag{6}$$

We can show that the optimal policy with heterogeneous workers is still a threshold policy, but now the thresholds depend on the worker values.

THEOREM 3.3: *The optimal policy for the heterogeneous worker problem is determined by a set of state-dependent thresholds, $\infty = t_0(n, \gamma, p) > t_1(n, \gamma, p) \geq t_2(n, \gamma, p) \geq \dots \geq t_K(n, \gamma, p)$, such that if an arriving job has value x , it is optimal to assign it to a worker of type i if $t_{i-1}(n, \gamma, p) > x \geq t_i(n, \gamma, p)$ and to reject it if $x < t_K(n, \gamma, p)$.*

PROOF: If we accept the job and assign it to a type i worker, the value is $p_i x + W(n^i, \gamma, p)$, which is increasing in x , whereas if we reject it, the value is $W(n, \gamma, p)$, so there must be a rejection threshold, $t_K(n, \gamma, p)$, such that a job will be rejected if its value is lower than the threshold and accepted if it is greater. Similarly, we will prefer to assign the job to a worker of type $i - 1$ rather than type i if $x \geq (W(n^i, \gamma, p) - W(n^{i-1}, \gamma, p)) / (p_{i-1} - p_i)$. ■

We now consider the impact of variability in the worker values and job values on the optimal return. For two vectors p and p' , we say that p' weakly majorizes p , $p' \succ_w p$, if

$$\sum_1^k p'_{[i]} \geq \sum_1^k p_{[i]}, \quad k = 1, \dots, K, \tag{7}$$

where $p_{[i]}$ denotes the components of p in decreasing order [7]. For two random variables X' and X , we say that X' is larger than X in the increasing convex sense, $X' \geq_{icx} X$, if $Ef(X') \geq Ef(X)$ for all increasing and convex functions f . Intuitively, p' is both larger and more variable than p , and X' is larger and more variable than X . A function f is increasing and Schur-convex if $f(p') \geq f(p)$ for any $p' \succ_w p$.

We have the following result: More variability yields greater returns. Part (ii) is a special case of the result of Lippman and Ross [6] for marked point processes. We give a simple proof for completeness.

THEOREM 3.4: *Let $W'(n, \gamma, p)$ be the maximal expected discounted return under the optimal policy when the job values have the same distribution as X' instead of X .*

- (i) *If $p' \succ_w p$, then $W(n, \gamma, p') \geq W(n, \gamma, p)$.*
- (ii) *If $X' \geq_{icx} X$, then $W'(n, \gamma, p) \geq W(n, \gamma, p)$.*

PROOF:

(i) We can write $W(n, \gamma, p)$ as follows:

$$W(n, \gamma, p) = \sum_{i=1}^K p_i U_i(n, \gamma, p), \tag{8}$$

where $U_i(n, \gamma, p)$ is the total EJV for type i workers under the optimal policy given (n, γ, p) . Now, suppose that the worker values are p' but we follow the policy that is optimal when worker values are p . Then the total discounted return will be $\sum_{i=1}^K p'_i U_i(n, \gamma, p)$, which is less than the optimal total return; that is,

$$W(n, \gamma, p') \geq \sum_{i=1}^K p'_i U_i(n, \gamma, p) \geq \sum_{i=1}^K p_i U_i(n, \gamma, p) = W(n, \gamma, p), \tag{9}$$

where the second inequality follows because $\sum p_i u_i$ is an increasing Schur-convex function of p for any $u = (u_1, \dots, u_K)$ [7].

(ii) We can write $W(n, \gamma, p|x)$ as follows, where $t_{K+1}(n, \gamma, p) = p_{K+1} = 0$ and $n^{K+1} = n$:

$$W(n, \gamma, p|x) = \sum_{i=1}^{K+1} (p_i x I\{t_{i-1}(n, \gamma, p) > x \geq t_i(n, \gamma, p)\} I\{n_i > 0\} + W(n^i, \gamma, p)). \tag{10}$$

Since this is an increasing convex function of x , the result follows. ■

A consequence of part (i) is that we can bound the total optimal return with the easier to compute IO policy. Let $\bar{p} = \sum p_i / K$.

COROLLARY 3.5:

$$W(n, \gamma, p) \geq W(n, \gamma, \bar{p}) = \bar{p} W^{IO}(N_K, \Gamma_K, 1). \tag{11}$$

References

1. Akçay, Y., Balakrishnan, A., & Xu, S.H. (2010) Dynamic assignment of flexible service resources. *Production and Operations Management* 19: 279–304.
2. Albright, S.C. (1974). Optimal sequential assignments with random arrival times. *Management Science* 21: 60–67.
3. David, I. & Yechiali, U. (1990). Sequential assignment match processes with arrivals of candidates and offers. *Probability in the Engineering and Informational Sciences* 4: 413–430.
4. David, I., & Yechiali, U. (1995). One-attribute sequential assignment match processes in discrete time. *Operations Research* 43: 879–884.

5. Derman, C., Lieberman, G., & Ross, S.M. (1972). A sequential stochastic assignment problem. *Management Science* 18: 349–355.
6. Lippman, S.A. & Ross, S.M. (2008). Variability is beneficial in marked stopping problems. *Economic Theory* 35: 333–342.
7. Marshall, A. & Olkin, I. (1979). *Inequalities: Theory of majorization and its applications*. Orlando, FL: Academic Press.
8. Nikolaev, A. & Jacobson, S.H. (2010). Stochastic sequential decision-making with a random number of jobs. *Operations Research* 58: 1023–1027.
9. Nikolaev, A., Jacobson, S.H., & McLay, L.A. (2007). A sequential stochastic security system design problem for aviation security. *Transportation Science* 41: 182–194.
10. Saario, V. (1985). Limiting properties of the discounted house-selling problem. *European Journal Operational Research* 20: 206–210.
11. Su, X.M. & Zenios, S.A. (2005). Patient choice in kidney allocation: A sequential stochastic assignment model. *Operations Research* 53: 443–455.
12. You, P.S. (2000). Sequential buying policies. *European Journal of Operational Research* 120: 535–544.