The Dynamic and Stochastic Knapsack Problem with Deadlines
Author(s): Jason D. Papastavrou, Srikanth Rajagopalan and Anton J. Kleywegt
Reviewed work(s):
Source: *Management Science*, Vol. 42, No. 12 (Dec., 1996), pp. 1706-1718
Published by: INFORMS
Stable URL: http://www.jstor.org/stable/2634548
Accessed: 10/01/2013 07:01

The Dynamic and Stochastic Knapsack Problem with Deadlines

Jason D. Papastavrou • Srikanth Rajagopalan • Anton J. Kleywegt
School of Industrial Engineering, Purdue University, West Lafayette, Indiana 47907-1287

In this paper a dynamic and stochastic model of the well-known knapsack problem is developed and analyzed. The problem is motivated by a wide variety of real-world applications. Objects of random weight and reward arrive according to a stochastic process in time. The weights and rewards associated with the objects are distributed according to a known probability distribution. Each object can either be accepted to be loaded into the knapsack, of known weight capacity, or be rejected. The objective is to determine the optimal policy for loading the knapsack within a fixed time horizon so as to maximize the expected accumulated reward. The optimal decision rules are derived and are shown to exhibit surprising behavior in some cases. It is also shown that if the distribution of the weights is concave, then the decision rules behave according to intuition.
(*Dynamic Programming; Sequential Stochastic Resource Allocation*)

1. Introduction

The *knapsack* problem is one of the most studied problems in operations research (Martello and Toth 1990). There are two primary reasons this problem is of interest: (1) even though it is a simple problem to describe, it is hard to solve, and (2) it has many practical applications, for example in resource allocation problems. The knapsack problem is a static and deterministic approximation of a problem that often is both dynamic and stochastic. Requests for the resource arrive one-by-one stochastically in time, and must either be accepted or be rejected on the spot (on-line) without the benefit of complete information, which includes the arrival times, the amounts requested, and the associated rewards of all future requests. This lack of complete information is a critical factor in the determination of optimal policies.

In this paper the *Dynamic and Stochastic Knapsack Problem with Deadlines* is defined and analyzed. This problem has the following characteristics:

a. The resource is limited (i.e., the knapsack has a fixed capacity).

b. Requests for the resource (i.e., objects to be included in the knapsack) arrive in time according to a stochastic process.

c. The demands for the resource (i.e., weights) and their associated rewards are random, and become known upon arrival.

d. A time deadline exists after which requests cannot be accepted.

e. If a demand is rejected, it cannot be recalled.

f. Decisions on the acceptance or rejection of demands have to be made in real-time.

g. The objective is to maximize the expected reward accumulated by the deadline.

Applications that motivated this research, because they can be modeled as dynamic and stochastic resource allocation problems, include:

1. *Loading containers and vehicles.* Requests for the transportation of loads arrive randomly. If the load is accepted it is placed in the container, otherwise it is serviced by another carrier. Usually there exists a time deadline after which loading stops because the container has to be shipped. The objective is to determine the optimal policy for accepting requests as they arrive, or for setting the transportation fees, in order to maximize the expected profit accumulated by the deadline. The optimal policy is a function of the state of the system. For example, the policy usually becomes more lenient as the deadline approaches. Moreover, the larger

0025-1909/96/4212/1706$01.25

the size (weight or volume) of a load, the larger the reward should be for the load to be accepted.

2. *Selling real-estate and selling cars.* Consider the case of a land developer trying to sell the lots in a new subdivision, or a real-estate agent trying to sell condominiums in a new apartment complex. The developer or the agent would like to sell the units within a given time period. The arrival process of potential buyers is random, as is the amount that each customer offers for a unit. The objective is to determine and set the optimal prices for each unit, and also to determine the optimal policy for accepting offers from potential buyers. Again the optimal policy is a function of the state of the system. A similar problem is faced by car dealers who need to decide on the offers they are willing to accept. The policies depend on the popularity of the model, the number of cars in stock at the dealership, and the time until the next year's models come out.

3. *Accepting loan requests.* Consider a bank that has a given amount of funds to invest in credit card loans, mortgages, and personal loans. Requests for credit and loans arrive according to some stochastic process, and the bank officers have to decide which should be accepted and for what amount. The bank strives to maximize its expected return taking into account the yield that each loan bears. The bank has to adjust its policies to reflect changes in the economy, the interest rates, and the amount of funds available.

4. *Taking reservations at a restaurant.* The advantages of taking reservations are that (a) they increase the probability that parties interested in dining out show up, and (b) they reduce the uncertainty in the number of customers; thus the management can better estimate the number of waiters and cooks that will be needed on a particular night. The disadvantage is that the utilization of tables is lower, and that can lead to lost profits. If reservations are not accepted, the potential customers will likely go to a different restaurant. The management would like to determine the optimal policy for accepting reservations for different sized parties. The policy should depend not only on the size of the party (larger parties mean larger profits), but also on the number of reservations already accepted (if too many reservations are accepted, then walk-in customers will be dissatisfied because they will have to wait too long).

5. *Selling tickets for air travel and for sports events.* Airlines offer different fares depending on the number of available seats and the time until departure. A limited number of super-saver fares are available if reservations are made at least two to three weeks prior to departure because early reservations help the airlines to schedule their fleets more efficiently, as well as cover some of the fixed costs of the flights. Moreover, if an airline does not offer special fares ahead of time, potential travelers will make reservations with a competitor. On the other hand, as the number of available seats decreases, the airlines are less willing to offer special prices because they expect to sell a sufficient number of tickets at regular prices. Some airlines also offer special prices with stand-by tickets for seats that would otherwise have been empty. The objective for an airline is to determine a policy to maximize the expected profit. A similar problem is encountered by (popular) sports teams that have to decide (a) how many season tickets to sell, (b) how many promotional discounts should be offered and for which games, and (c) how the different tickets should be priced.

6. *Scheduling a batch processor.* Consider the problem faced by the scheduler of a batch processor with fixed capacity. Jobs arrive over time, and the capacity requirements and rewards of jobs are unknown before arrival. Fixed schedules and commitments lead to deadlines. The objective is to determine a policy for accepting jobs in order to maximize the expected profit. The manager would also like to know how the optimal policy and the expected profit will be affected by changes in the operating conditions; for example, the manager may want to study the effects of acquiring a larger processor before deciding to proceed with this investment.

Static versions of the stochastic knapsack problem have been studied. In these, the set of objects is known and the rewards and/or weights are random (Carraway et al. 1993; Henig 1990; Sniedovich 1980, 1981; Steinberg and Parks 1979). The objective typically is to maximize the probability of attaining some prespecified level of utility or reward.

Dynamic versions of the stochastic knapsack problem have also been considered. Objects arrive over time and the rewards and/or weights are unknown prior to arrival. Some stopping time problems and optimal selection problems are similar to our problem (Bruss 1984;

Freeman 1983; Nakai 1986; Presman and Sonin 1972; Sakaguchi 1984a; Stewart 1981; Tamaki 1986a, b; Yasuda 1984). A well-known example is the *secretary problem* where candidates arrive one at a time. The objective is to maximize the probability of selecting the best candidate or group of candidates, or to maximize the expected value of the selected candidates from a given or random set of candidates.

Another related problem is the *Sequential Stochastic Assignment Problem* (SSAP). Derman et al. (1972) defined the problem as follows: a given number n of persons, with known values p_i, $i = 1, \ldots, n$, are to be assigned sequentially to n jobs, which arrive one at a time. The jobs have values x_j, $j = 1, \ldots, n$, which are unknown before arrival, but become known upon arrival, and which are independent and identically distributed with a known probability distribution. If a person with value p_i is assigned to a job with value x_j, the reward is $p_i x_j$. The objective is to maximize the expected total reward. Different extensions of the SSAP were studied by Albright (1974), Sakaguchi (1984b, c), Nakai (1986b, c), Kennedy (1986) and Righter (1989).

Many investment problems are variations of the model studied in this paper. For example, Prastacos (1983) studied the problem of allocating a given amount of resource before a deadline to irreversible investment opportunities that arrive according to a geometric process. Prastacos assumed that each investment opportunity is large enough to absorb all the available capital; but in our problem the sizes of investment opportunities are given, and cannot be chosen.

A class of problems similar to the dynamic and stochastic knapsack problem is known as *Perishable Asset Revenue Management* (PARM) problems (Weatherford and Bodily 1992), or as *yield management* problems. Another type of PARM problem, which occurs when a perishable inventory has to be sold before a deadline, has been studied by Gallego and Van Ryzin (1994). In their problem demands arrive according to a Poisson process with price dependent rate. The major difference with our model is that in our model offers arrive, and the offers can be accepted or rejected, as is typical with large contracts such as selling of real estate, whereas in the model of Gallego and Van Ryzin prices are set and all demands are accepted as long as supplies last, which is typical in retail.

Hassin and Henig (1986) studied a dynamic control problem where demand and supply offers arrive according to a Poisson process, and can be accepted or rejected. The objective is to maximize the difference between the discounted demand offers and the discounted supply offers. They show that the optimal policy is to accept or reject an offer depending on whether its value is above or below a critical value, which depends on the state of the system. Similar versions of the dynamic and stochastic knapsack problem have been studied for communication applications (Kaufman 1981, Ross and Tsang 1989, Ross and Yao 1990).

Kleywegt and Papastavrou (1995) studied a problem similar to the one in this paper with demands arriving in continuous time, and including a waiting cost.

The *Dynamic and Stochastic Knapsack Problem with Deadlines* is analyzed as follows: It is shown that the optimal policy is a *threshold* type policy. These thresholds are analyzed to characterize their behavior under a variety of operating conditions (i.e., as the remaining capacity, the time, the weights and rewards change). Several cases of the problem are considered.

In the first case, objects have equal weights, but random rewards. This situation describes, for example, the selling of tickets for air travel where each passenger requires one seat, but potential passengers are willing to pay different fares. It is shown that the optimal decision rules are "well-behaved," that is, according to intuition, the optimal policy becomes more lenient as the capacity of the knapsack increases and as the deadline approaches. The case where all objects have the same reward, but different weights is considered next. This describes the situation encountered in truck leasing, where trucks are leased at a fixed rate, irrespective of the load to be transported. Then the general model is considered where both the rewards and the weights are random. This models the problem faced by the manager of an LTL trucking operation. The sizes and rewards of the loads that will be received are random. Finally the special case is considered where the reward is proportional to the weight. This describes the situation faced by the loan department of a bank; the interest collected by the bank is proportional to the amount of the loan. In the last three cases, the weights of the objects are not equal, and this causes the optimal decision rules to exhibit surprising counterintuitive behavior in some cases. It is also shown that the optimal decision rules become "well-behaved" if the distribution of the weights satisfies some special conditions (i.e., the conditional probability distribution of the weights given the rewards is concave in weight for all positive weights and rewards). In colloquial terms, the implication of this condition is that "a bird in the hand is worth two in the bush."

Some comments are in order before proceeding. Since both the case of equal weights and the case of equal rewards are special cases of the general model, their analyses could have followed the analysis of the general model. But it was decided to present them before the general model because (1) the optimal policy for the case with equal weights is intuitive and establishes what we call "well-behaved" decision rules, and (2) it facilitates the interested reader who desires to study the proofs because the proofs for the case with equal rewards are simpler versions of the proofs for the general model. All the proofs can be found in Papastavrou et al. (1994).

The rest of the paper is organized as follows: The *Dynamic and Stochastic Knapsack Problem with Deadlines* is defined in §2. In §3, it is established that the optimal policy is a threshold type policy. The case with equal weights is analyzed in §4, and the case with equal rewards is considered in §5. The model with random weights and random rewards is studied in §6, and the case where the reward is proportional to the weight is investigated in §7. Our concluding remarks follow in §8.

2. Model Description

Consider a knapsack of given weight capacity. Objects arrive over a time horizon of T discrete periods. The periods are numbered from 1 to T, with T being the last period in the horizon. In every period there is a constant probability p of one object arriving, and probability $1 - p$ of no arrivals. The weights and rewards of different arrivals are independent. As soon as an object arrives, its weight W and reward R become known. The weights and rewards are positive random variables, and are distributed according to a known joint probability distribution $F_{WR}(w, r)$. Whenever an object is rejected, that object is "lost" (i.e., it cannot be recalled at a later time). Of course, an arriving object can be accepted only if its weight is less than or equal to the remaining capacity of the knapsack. The objective is to determine decision rules for accepting or rejecting objects in order to maximize the expected reward accumulated at the end of the time horizon.

Let Π denote the class of policies that take only past events into account, and that prescribe the acceptance or rejection of arrivals as they occur. Therefore no "look-ahead," postponement of decisions or recall are allowed. We restrict attention to this class of policies. Because the weights and rewards of different arrivals are independent, the expected reward accumulated from (and including) period t until the deadline, if the remaining capacity is c and policy $\pi \in \Pi$ is implemented, depends only on t, c and π, and not on the full history of the process. Let $E[V_t^c | \pi]$ denote this expected accumulated reward. Let EV_t^c denote the optimal expected accumulated reward, i.e.,

$$EV_t^c = \sup_{\pi \in \Pi} \{E[V_t^c | \pi]\}$$

For the same reason, we can further restrict attention to the class of policies that map the appropriate set of 4-tuples (t, c, w, r) into the action space $\{accept, reject\}$.

3. General Results

In this section, some general preliminary results are presented.

LEMMA 1. (i) *EV_t^c is a nondecreasing function of c.*
(ii) *EV_t^c is a nonincreasing function of t.*

THEOREM 1. *Suppose that, at time t, the remaining capacity is c, and a reward r arrives. The optimal decision rule is a threshold rule defined by:*

$$\pi^*(t, c, w, r) = \begin{cases} \textit{accept if} \quad r + EV_{t+1}^{c-w} \geq EV_{t+1}^c \quad \textit{and} \quad w \leq c, \\ \textit{reject if} \quad r + EV_{t+1}^{c-w} < EV_{t+1}^c \quad \textit{or} \quad w > c. \end{cases}$$

Note that the quantity $r + EV_{t+1}^{c-w}$, which depends on both the reward and the weight of the incoming object, is compared to the constant threshold EV_{t+1}^c. Moreover, if $r + EV_{t+1}^{c-w} = EV_{t+1}^c$, it is equally beneficial to accept and to reject the object. It was decided to arbitrarily break the ties by accepting the object.

COROLLARY 1. *The optimal accumulated reward can be evaluated recursively from*:

$$\begin{aligned} EV_t^c &= P[W \le c, R + EV_{t+1}^{c-W} \ge EV_{t+1}^c] \\ &\quad \times E[R + EV_{t+1}^{c-W} \mid W \le c, R + EV_{t+1}^{c-W} \ge EV_{t+1}^c] \\ &\quad + \{P[R + EV_{t+1}^{c-W} < EV_{t+1}^c, W \le c] \\ &\quad + P[W > c]\} EV_{t+1}^c \end{aligned}$$

with boundary condition:

$$EV_t^c = 0 \quad \forall c \quad \textit{and} \quad t > T.$$

The optimal decision rule of Theorem 1 can also be expressed as follows:

COROLLARY 2. *Suppose that, at time period t, the remaining capacity is c and that an object of weight w arrives. The optimal decision rule is the threshold rule*:

$$\pi^*(t, c, w, r) = \begin{cases} \textit{accept if} & r \ge R_t^c(w), \\ \textit{reject if} & r < R_t^c(w), \end{cases}$$

where the **critical reward** $R_t^c(w)$ *is given by*:

$$R_t^c(w) = \begin{cases} EV_{t+1}^c - EV_{t+1}^{c-w}, & w \le c, \\ \infty, & w > c. \end{cases}$$

COROLLARY 3. *Suppose that, at time period t, the remaining capacity is c and that an object of reward r arrives. The optimal decision rule is the threshold rule*:

$$\pi^*(t, c, w, r) = \begin{cases} \textit{accept if} & w \le W_t^c(r), \\ \textit{reject if} & w > W_t^c(r), \end{cases}$$

where the **critical weight** $W_t^c(r)$ *is given by*:

$$W_t^c(r) = \begin{cases} \sup\{w \le c \mid EV_{t+1}^c - EV_{t+1}^{c-w} \le r\}, & c > 0, \\ 0, & c = 0. \end{cases}$$

COROLLARY 4. $R_t^c(w)$ *is a nondecreasing function of w for all t and c.*

COROLLARY 5. $W_t^c(r)$ *is a nondecreasing function of r for all t and c.*

4. Equal Weights

In this section, a special case of the stochastic knapsack problem is considered where objects have equal weights w. Without loss of generality, let $w = 1$. Assume that in every period exactly one object arrives (i.e., $p = 1$). As will be explained later the results also hold for the case where arrivals occur with probability $p < 1$ in each period.

The problem with $p = 1$ is a special case of the Sequential Stochastic Assignment Problem studied by Derman et al. (1972). This is obtained by letting the remaining number of men n in their problem be equal to the remaining number of time periods until the deadline $T - t + 1$; letting $p_i = 1$ for $\min\{c, n\}$ of these men, and $p_i = 0$ for $\max\{0, n - c\}$ men; and letting the job value x be equal to the value of the reward r that arrives in period t. We study the behavior of the optimal expected accumulated reward and critical reward as functions of time (or n) and remaining capacity. These issues were not addressed in Derman et al. (1972).

Since the weights are equal ($w = 1$), the optimal decision rules of Corollary 2 reduce to:

$$\pi^*(t, c, r) = \begin{cases} \text{accept if} & r \ge R_t^c, \\ \text{reject if} & r < R_t^c, \end{cases}$$

where the critical reward is defined by:

$$R_t^c = \begin{cases} EV_{t+1}^c - EV_{t+1}^{c-1}, & c \ge 1, \\ \infty, & c < 1. \end{cases} \tag{1}$$

Note that R_t^c is nonnegative from Lemma 1. Moreover, using Corollaries 1 and 2, the expected reward can be obtained from:

$$\begin{aligned} EV_t^c &= \int_0^{R_t^c} EV_{t+1}^c dF_R(r) + \int_{R_t^c}^{\infty} (r + EV_{t+1}^{c-1}) dF_R(r) \\ &= EV_{t+1}^c F_R(R_t^c) + EV_{t+1}^{c-1}[1 - F_R(R_t^c)] + \int_{R_t^c}^{\infty} r dF_R(r) \\ &= EV_{t+1}^{c-1} + R_t^c F_R(R_t^c) + \int_{R_t^c}^{\infty} r dF_R(r) \end{aligned} \tag{2}$$

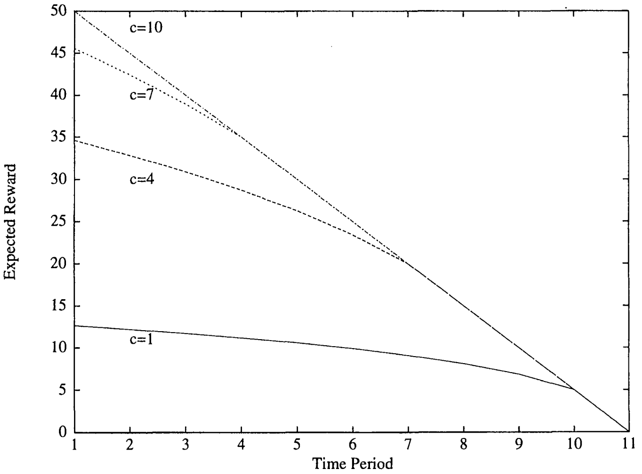
with boundary condition:

$$EV_t^c = 0 \quad \forall c \quad \text{and} \quad t > T. \tag{3}$$

The following results characterize the behavior of the critical reward and the expected accumulated reward.

THEOREM 2. *If objects have equal weights, then*
(i) EV_t^c *is a concave nondecreasing function of c for all t.*
(ii) EV_t^c *is a concave nonincreasing function of t for all c.*
(iii) R_t^c *is nonincreasing with c for all t.*
(iv) R_t^c *is nonincreasing with t for all c.*

Figure 1 **Expected Accumulated Reward Against Time for Different Remaining Capacities c, for Objects of Equal Weight ($w = 1$), Exponentially Distributed Rewards ($E[R] = 5$), and Deadline $T = 10$**



A few remarks about the significance of the above results are in order. The optimal policy becomes more lenient as the deadline approaches. This can be seen by the critical reward that is non-increasing with time. Moreover, extra capacity also makes the optimal policy more lenient (as reflected by the nonincreasing threshold values). The two concavity results are according to intuition. The concavity of the expected accumulated reward with respect to time implies that having extra time is more beneficial closer to the deadline (i.e., the marginal expected accumulated reward of time increases as the deadline approaches). Similarly, the concavity of the expected accumulated reward with respect to capacity implies that extra capacity is of greater benefit when the available capacity of the knapsack is smaller (i.e., the marginal expected accumulated reward of capacity decreases with increasing capacity).

Furthermore, the sequential investment problem with revenue being a convex function of the amount invested (Prastacos 1983) is a special case of our problem; the decision to accept an object is equivalent to investing the *entire* capital available at that point in time (i.e., the weight of each incoming object is always equal to the remaining capacity). Thus revenue being a convex instead of a concave function of the amount invested makes the optimal decision rule completely different.

Suppose that an arrival occurs during a time period with probability $p < 1$. The expected accumulated reward becomes:

$$EV_t^c = p\left[\int_0^{R_t^c} EV_{t+1}^c dF_R(r) + \int_{R_t^c}^{\infty} (r + EV_{t+1}^{c-1}) dF_R(r)\right] + (1 - p)EV_{t+1}^c$$

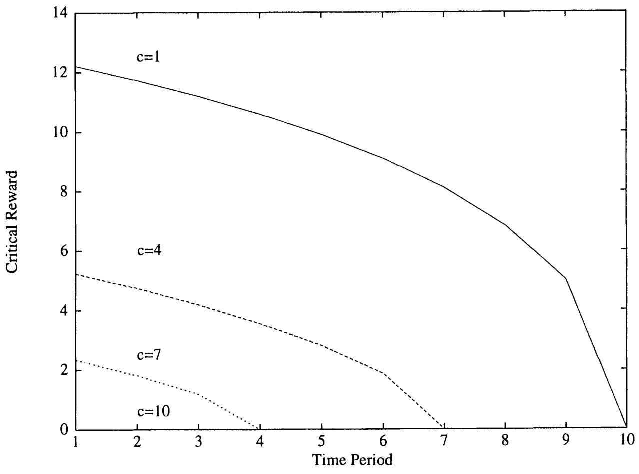
The critical reward and the boundary condition are given by Eqs. (1) and (3).

When an object arrives, the expected accumulated reward is obtained by conditioning on the acceptance or rejection of the object. When no arrival occurs (probability $1 - p$), the expected accumulated reward is EV_{t+1}^c. It can be shown that EV_t^c is a convex combination of two concave monotonic functions of c and t, and hence it is concave and monotonic with c and t as stated in Theorem 2. Thus, the general form of the solution remains the same. This should be intuitive because this case is equivalent to the case where an object arrives during each time period with probability 1, and the reward of the object is 0 with probability $1 - p$.

4.1. Numerical Results

The expected accumulated reward and critical reward can be computed recursively using Eqs. (1) and (2), for all capacities and time periods. Figures 1 and 2 show the expected accumulated reward and critical reward as

Figure 2 **Critical Reward Against Time for Different Remaining Capacities c, for Objects of Equal Weight, ($w = 1$), Exponentially Distributed Rewards ($E[R] = 5$), and Deadline $T = 10$**



a function of time for the case with exponentially distributed reward with mean $E[R] = 5$ and deadline $T = 10$. Recall that the problem has been defined in discrete time; the discrete points of each "curve" have been joined only for clarity. It was found that the shapes of these graphs do not change drastically if the reward distribution is changed to some other commonly used distributions, such as the uniform, normal and triangular distributions.

5. Equal Rewards

In this section, the special case of the stochastic knapsack problem is considered where objects have equal rewards r. Thus objects differ only in weight. If an object with weight greater than the remaining capacity arrives, the object is rejected. Since the rewards are equal, the optimal decision rule of Corollary 3 reduces to:

$$\pi^*(t, c, w) = \begin{cases} \text{accept if} & w \le W_t^c, \\ \text{reject if} & w > W_t^c, \end{cases}$$

where the critical weight is given by:

$$W_t^c = \begin{cases} \sup\{w \le c \mid EV_{t+1}^c - EV_{t+1}^{c-w} \le r\}, & c > 0, \\ 0, & c = 0. \end{cases} \tag{4}$$

Since $r > 0$, W_t^c is nonnegative. Moreover, using Corollaries 1 and 3, the optimal expected reward EV_t^c may be obtained from:

$$\begin{aligned} EV_t^c &= \int_0^{W_t^c} (r + EV_{t+1}^{c-w}) dF_W(w) + \int_{W_t^c}^{\infty} EV_{t+1}^c dF_W(w) \\ &= rF_W(W_t^c) + \int_0^{W_t^c} EV_{t+1}^{c-w} dF_W(w) \\ &\quad + EV_{t+1}^c[1 - F_W(W_t^c)] \end{aligned} \tag{5}$$

with boundary condition:

$$EV_t^c = 0 \quad \forall c \quad \text{and} \quad \forall t > T. \tag{6}$$

Because the weights are not equal, both the expected accumulated reward and critical weight are *not* necessarily "well-behaved" functions of capacity and time, as stated in the following proposition.

PROPOSITION. a. *EV_t^c does not have to be a concave function of c.*
b. *W_t^c does not have to be a nondecreasing function of c.*
c. *EV_t^c does not have to be a concave function of t.*
d. *W_t^c does not have to be a nondecreasing function of t.*

The validity of the above proposition will be demonstrated by the following example. Suppose that the deadline is $T = 8$, the capacity of the knapsack is 14, and the probability mass function of the weight is:

$$f_W(w) = \begin{cases} 0.80 & w = 1, \\ 0.19 & w = 5, \\ 0.01 & w = 7. \end{cases}$$

The problem is recursively solved using Eqs. (4), (5), and (6). The results for periods 1 through 8 are presented in Table 1.

a. EV_8^c is not a concave function of c, since $EV_8^3 = EV_8^4 < EV_8^5 = EV_8^6$

b. W_2^c is not a nondecreasing function of c, since $W_2^{13} = 7 > 5 = W_2^{14}$. That is, if an object of weight $w = 7$ arrives in time period $t = 2$, it is accepted if the remaining capacity is 13 and is rejected if the remaining capacity is 14. Thus, the optimal policy can become stricter as the remaining capacity increases.

c. EV_t^c is not a concave function of t, since $EV_4^6 - EV_5^6 = 0.778 < 0.796 = EV_3^6 - EV_4^6$.

d. W_t^{14} is not a nondecreasing function of t, since $W_2^{14} = 5 < 7 = W_1^{14}$. Therefore, an object of weight $w = 7$ is accepted at $t = 1$ and is rejected at $t = 2$. Thus, the optimal policy can become stricter as the deadline approaches.

The results are interesting because they are so counterintuitive. Intuition suggests that the optimal policy should become more lenient as the deadline approaches or as the capacity increases. This surprising behavior is attributed to the combinatorial nature of the problem; that is, the way that the different weights interact to fill the remaining capacity of the knapsack.

Fortunately, not all the results obtained are "negative." If some structure is imposed on the distribution of the weight, the expected accumulated reward and the critical weight become "well-behaved," and some intuitive results are obtained. Henceforth, the analysis is

Table 1 Expected Accumulated Rewards and Critical Weights

t	c	1	2	3	4	5	6	7	8	9	10	11	12	13	14
8	EV_8^c	0.800	0.800	0.800	0.800	0.990	0.990	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	W_8^c	1	1	1	1	5	5	7	7	7	7	7	7	7	7
7	EV_7^c	0.960	1.600	1.600	1.600	1.639	1.943	1.944	1.960	1.960	1.996	1.996	1.999	1.999	2.000
	W_7^c	1	1	1	1	5	5	5	7	7	7	7	7	7	7
6	EV_6^c	0.992	1.888	2.400	2.400	2.408	2.504	2.868	2.869	2.888	2.895	2.982	2.983	2.992	2.992
	W_6^c	1	1	1	1	1	5	5	5	7	7	7	7	7	7
5	EV_5^c	0.998	1.971	2.794	3.200	3.201	3.227	3.380	3.769	3.770	3.792	3.817	3.955	3.956	3.971
	W_5^c	1	1	1	1	1	1	5	5	5	7	7	7	7	7
4	EV_4^c	0.999	1.993	2.935	3.672	4.000	4.005	4.053	4.262	4.651	4.652	4.677	4.727	4.912	4.915
	W_4^c	1	1	1	1	1	1	1	5	5	5	7	7	7	7
3	EV_3^c	0.999	1.998	2.981	3.882	4.538	4.801	4.815	4.895	5.144	5.517	5.519	5.552	5.631	5.854
	W_3^c	1	1	1	1	1	1	1	1	5	5	5	7	7	7
2	EV_2^c	0.999	1.999	2.994	3.961	4.813	5.390	5.604	5.631	5.745	6.022	6.371	6.376	6.420	6.516
	W_2^c	1	1	1	1	1	1	1	1	1	5	5	5	7	5
1	EV_1^c	0.999	1.999	2.998	3.988	4.932	5.729	6.233	6.409	6.454	6.600	6.896	7.215	7.225	7.283
	W_1^c	1	1	1	1	1	1	1	1	1	1	5	5	5	7

restricted to the case where the following Consistency Condition holds:

CONSISTENCY CONDITION. *The probability distribution of the weight F_W is concave on $(0, \infty)$.*

This condition implies that a nonincreasing density f_W exists. Several commonly used distributions satisfy this condition, such as the exponential, uniform, and some triangular, Weibull and beta distributions. Theorem 3 characterizes the behavior of the critical weight and the expected accumulated reward.

THEOREM 3. *If the probability distribution of W is concave on $(0, \infty)$, then*

(i) *EV_t^c is a concave nondecreasing function of c for all t.*
(ii) *EV_t^c is a concave nonincreasing function of t for all c.*
(iii) *W_t^c is a nondecreasing function of c for all t.*
(iv) *W_t^c is a nondecreasing function of t for all c.*

The Consistency Condition makes the optimal decision rules behave according to intuition. The policy becomes less lenient (critical weight decreases) further away from the deadline, and more lenient as the capacity increases. Concavity of the expected accumulated reward with respect to time implies that the marginal reward of additional time increases as the deadline approaches. Similarly, concavity with respect to capacity implies that the marginal reward of capacity increases as the knapsack is filled. These results parallel similar results obtained for search problems (Prastacos 1983, Saario 1985).

5.1. Numerical Results

Numerical results are shown for two weight distributions. These were computed using Eqs. (4), (5), and (6). The deadline is $T = 10$, and the constant reward of each object is $r = 1$.

In the first case, the weights are Poisson distributed:

$$f_W(w) = \frac{(4)^{w-1}e^{-4}}{(w-1)!}, \quad w = 1, 2, 3, \cdots$$

with mean $E[W] = 5$. The Consistency Condition is not satisfied. The expected accumulated reward as a function of capacity for different time periods are shown in Figure 3. It is clear that the expected accumulated reward is not concave with respect to capacity at time $t = 10$. (At time $t = T = 10$, $EV_T^c = F_W(c)$, which is not concave with respect to c). The nonmonotonic behavior of the critical weight can be seen in Figure 5, where W_t^c decreases from $t = 4$ to $t = 6$ with $c = 50$.

In the second case the weights are distributed according to the following triangular probability density function:

$$f_W(w) = \frac{2}{225}(15 - w), \quad 0 \le w \le 15$$

Figure 3 **Expected Accumulated Reward Against Capacity for Different Time Periods, for Objects with Equal Rewards ($r = 1$), Poisson Distributed Weights (Consistency Condition Not Satisfied) with $E[W] = 5$, and Deadline $T = 10$**

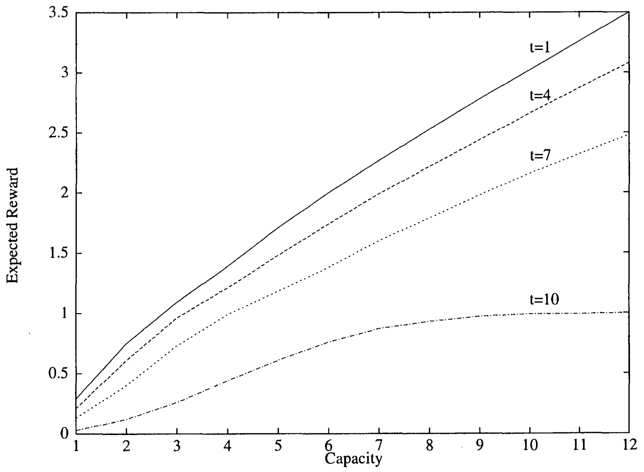
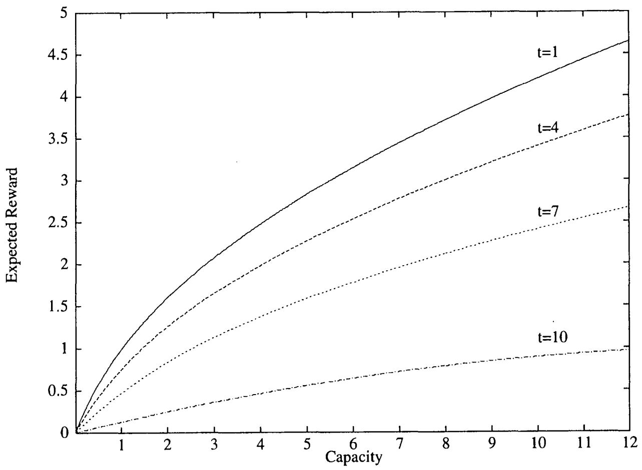


Figure 4 **Expected Accumulated Reward Against Capacity for Different Time Periods, for Objects with Equal Rewards ($r = 1$), Triangular Distributed Weights (Consistency Condition Satisfied) with $E[W] = 5$, and Deadline $T = 10$**



with mean $E[W] = 5$. This probability density function is nonincreasing for $w \geq 0$, hence the Consistency Condition is satisfied. The expected accumulated reward is concave nondecreasing with capacity (Figure 4) and the critical weight is nondecreasing with time (Figure 6).

6. Random Weights and Random Rewards

The general version of the problem, where both the rewards and the weights are random, is analyzed. In the previous section it was shown that if all objects do not have the same weight, the expected accumulated reward and the critical weight are not necessarily "well-behaved," unless the distribution of the weight is concave. This condition is generalized as follows:

CONSISTENCY CONDITION. *The conditional distribution of the weight given the reward $F_{W|R}(w|r)'$ is concave with w on $(0, \infty)$ for all r.*

This implies that a non-increasing conditional density $f_{W|R}(w|r)$ exists. Also, if W and R are independent (as is the case when objects have equal rewards), the above condition reduces to the condition in §5. Theorem 4 states that if the *Consistency Condition* holds, the optimal decision rules behave according to intuition.

THEOREM 4. *If the conditional probability distribution of W given R is concave with w on $(0, \infty)$ for all r, then*

(*i*) *EV_t^c is a concave nondecreasing function of c for all t.*
(*ii*) *EV_t^c is a concave nonincreasing function of t for all c.*
(*iii*) *$W_t^c(r)$ is a nondecreasing function of c for all t and r.*
(*iv*) *$W_t^c(r)$ is a nondecreasing function of t for all c and r.*
(*v*) *$R_t^c(w)$ is a nonincreasing function of c for all t and w.*
(*vi*) *$R_t^c(w)$ is a nonincreasing function of t for all c and w.*

Figure 5 **Critical Weight Against Time for Different Capacities for Objects with Equal Rewards ($r = 1$), Poisson Distributed Weights (Consistency Condition Not Satisfied) with $E[W] = 5$, and Deadline $T = 10$**

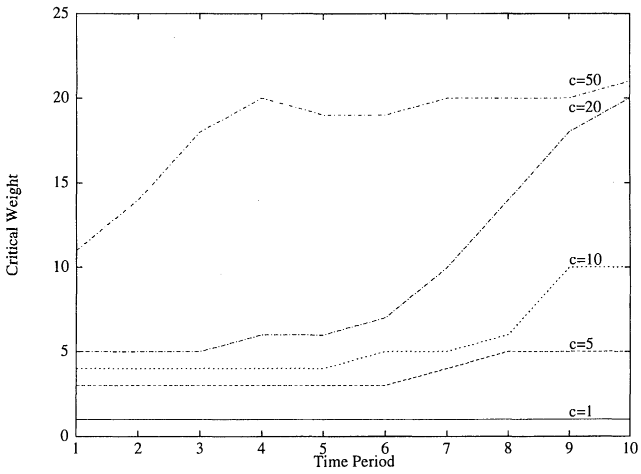
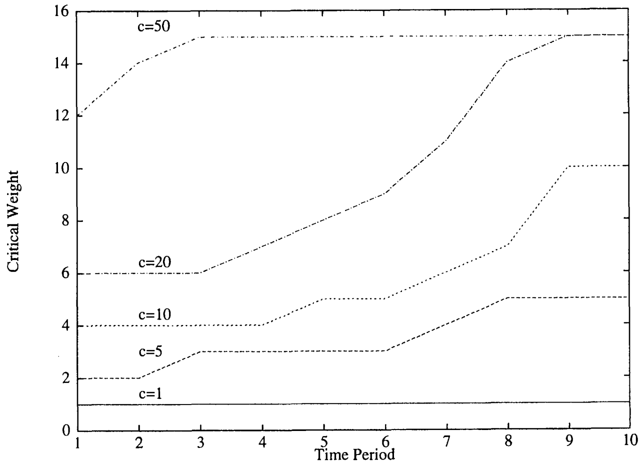


Figure 6 **Critical Weight Against Time for Different Capacities for Objects with Equal Rewards ($r = 1$), Triangular Distributed Weights (Consistency Condition Satisfied) with $E[W] = 5$, and Deadline $T = 10$**



If the Consistency Condition holds, the optimal policy becomes more lenient with increasing capacity and increasing time (i.e., critical weights increase and critical rewards decrease). Also, the concavity results imply that the marginal reward of time increases as the deadline approaches, and the marginal reward of capacity increases as the knapsack is filled.

6.1. Numerical Results

In the first example, the weight is distributed according to the following Poisson distribution:

$$f_W(w) = \frac{(19)^{w-1}e^{-19}}{(w-1)!}, \quad w = 1, 2, 3, \cdots$$

with mean $E[W] = 20$. The conditional probability density function of the reward given the weight is exponential:

$$f_{R|W}(r|w) = \frac{1}{10w} e^{-1/10w^r}, \quad r > 0$$

with $E[R \mid W = w] = 10w$; thus, $E[R] = 200$. The deadline is $T = 20$. The Consistency Condition is not satisfied. In Figure 7, EV_t^c is shown as a function of c for different values of t. It can be seen that EV_t^c is not a concave function of c. In Figure 9, $R_1^c(w)$ is shown as a function of c, for different weights w, for $t = 1$. The critical reward is not monotonic with respect to capacity. The critical reward increases until the capacity is approximately equal to the sum of the weight of the object and the average weight. For capacities less than this, if the object is accepted, the remaining capacity is less than the average weight, and thus the probability that another object will arrive and fit into the knapsack is relatively low. Also, even if such an object arrives, the associated reward will be relatively low (since the expected reward is proportional to the weight). Hence, the policy becomes less lenient. On the other hand, if the capacity is greater than the sum of the weight of the arriving object and the average weight, there is sufficient capacity to accept the object and still have enough capacity to accept another object in the future with a high probability. Thus, the admission policy becomes more lenient. This pattern is also observed for capacities that are close to multiples of the average weight (but not as dramatically) as demonstrated by the peaks of the curves.

In the second example, the rewards are uniformly distributed between 0 and 400:

$$f_R(r) = \frac{1}{400}, \quad 0 \leq r \leq 400$$

with mean $E[R] = 200$. The conditional probability density function of the weight given the reward is decreasing triangular:

Figure 7 **Expected Accumulated Reward Against Capacity for Different Time Periods, with Poisson Distributed Weights and Exponentially Distributed Rewards Given Weights (Consistency Condition Not Satisfied), with Deadline $T = 20$**

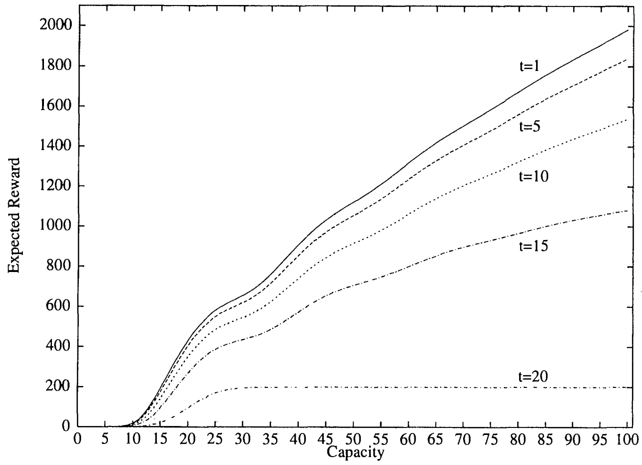


Figure 8 Expected Accumulated Reward Against Capacity for Different Time Periods, with Uniformly Distributed Rewards and Decreasing Triangular Distributed Weights Given Rewards (Consistency Condition Satisfied), with Deadline $T = 20$

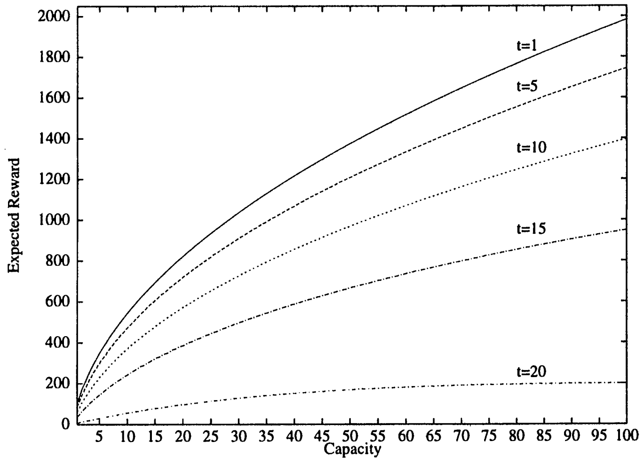
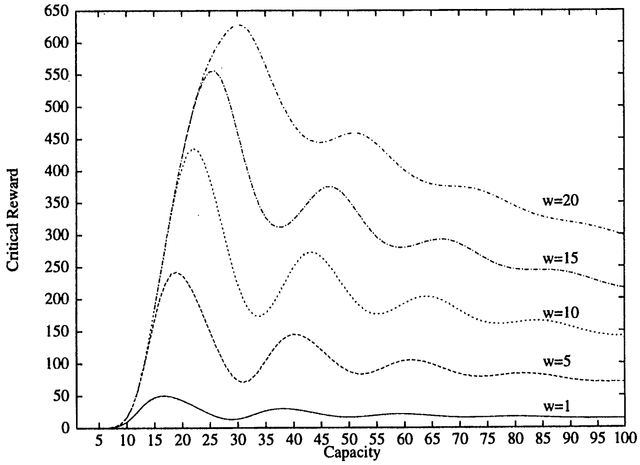


Figure 9 Critical Reward Against Capacity for Objects of Different Weights, with Poisson Distributed Weights and Exponentially Distributed Rewards Given Weights (Consistency Condition Not Satisfied), at Time $t = 1$, with Deadline $T = 20$



$$f_{W|R}(w|r) = \frac{200}{9r^2}\left(\frac{3r}{10} - w\right), \quad 0 \le w \le \frac{3r}{10}$$

with $E[W | R = r] = r/10$; thus, $E[W] = 20$. The deadline is again $T = 20$. This time the Consistency Condition is satisfied. The effects of the condition are clearly demonstrated in Figures 8 and 10. EV_t^c is a concave nondecreasing function of c (Figure 8), and $R_1^c(w)$ is nonincreasing with c (Figure 10).

7. Weight Proportional to Reward

In most cases of practical interest, the reward increases when the weight increases. For this reason, the special case where the reward of an object is proportional to its weight (i.e., $r = kw$, where k is a positive constant) is considered. Since the scale of the capacity of the knapsack is arbitrary, let $k = 1$, without loss of generality. Also, note that the Consistency Condition is not satisfied, since the conditional distribution of the weight given the reward is a step function, with unit step along the line $w = r$, and hence is not concave.

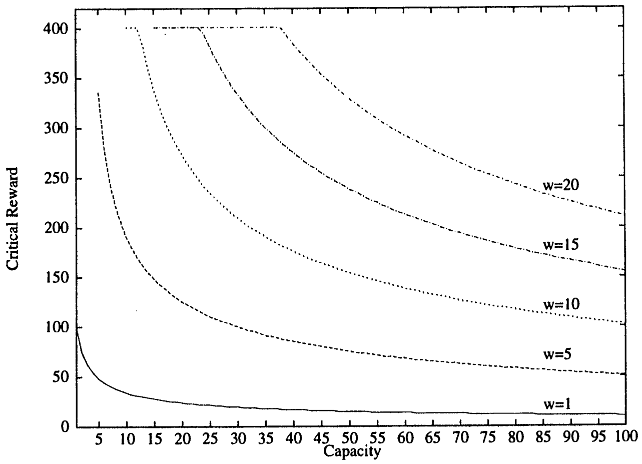
From Theorem 1 and Corollary 1, the optimal decision rule is given by:

$$\pi^*(t, c, r) = \begin{cases} \text{accept} & \text{if} \quad r + EV_{t+1}^{c-r} \ge EV_{t+1}^c \quad \text{and} \quad r \le c, \\ \text{reject} & \text{if} \quad r + EV_{t+1}^{c-r} < EV_{t+1}^c \quad \text{or} \quad r > c. \end{cases}$$

Intuition suggests that if the distribution of R is "reasonable," an optimal policy is to accept any object that arrives, as long as there is sufficient remaining capacity in the knapsack. This is established in the next theorem.

THEOREM 5. *If the weights are proportional to the rewards and the reward distribution F_R is concave on $(0, \infty)$, then*

Figure 10 Critical Reward Against Capacity for Objects of Different Weights, with Uniformly Distributed Rewards and Decreasing Triangular Distributed Weights Given Rewards (Consistency Condition Satisfied) at Time $t = 1$, with Deadline $T = 20$



(*i*) *the optimal policy is to accept any object that fits into the knapsack.*
(*ii*) EV_t^c *is a concave nonincreasing function of t for all c.*

The expected accumulated reward is neither a concave nor a convex function of c, as will be shown with an example in §7.1.

7.1. Uniform Distribution for the Weight and Reward

Suppose that the rewards are uniformly distributed on $[0, r^*]$. This distribution satisfies the condition of Theorem 5; thus, it is optimal to accept all objects that fit into the knapsack. For this case, we obtained a closed-form expression for the expected accumulated reward.

THEOREM 6. *If the reward is uniformly distributed on* $[0, r^*]$ *and the weights are equal to the rewards, then the expected accumulated reward n periods from the deadline T with remaining capacity c* ($c \le r^*$) *is given by*:

$$EV_{T-n}^c = c - \frac{r^*}{n+2}\left[1 - \left(1 - \frac{c}{r^*}\right)^{n+2}\right] \quad \forall c \le r^*.$$

It is straightforward to show that $r_0 + EV_{T-n}^{c-r_0} \ge EV_{T-n}^c$ for all $r_0 \le c \le r^*$ confirming Theorem 5. For capacities greater than r^*, the result may be obtained inductively from:

$$\frac{dEV_{T-n}^c}{dc} = 1 - \left(1 - \frac{c}{r^*}\right)^{n+1} \le 1 \quad \forall c \le r^* \text{ and}$$

$$\frac{dEV_t^c}{dc} = \int_0^{r^*} \frac{dEV_{t+1}^{c-r}}{dc}\, dF_R(r) \le 1 \quad \forall c > r^*.$$

Since the derivative of the expected accumulated reward with respect to capacity is less than 1, it is optimal to accept an object, as long as there is sufficient capacity in the knapsack.

The expected accumulated reward is neither concave nor convex with respect to capacity, as can be seen from:

$$EV_T^c = \begin{cases} c^2/2r^*, & c < r^*, \\ r^*/2, & c \ge r^*. \end{cases}$$

8. Concluding Remarks

A new model for dynamic and stochastic resource allocation problems is presented, that strives to capture the dynamic and stochastic characteristics of the environment in which real-world systems operate. Several diverse areas of applications are presented where the dynamic and stochastic knapsack problem with a deadline is appropriate.

The optimal policy is a threshold type policy. If all objects have the same weight, the decision rules behave according to intuition as the remaining capacity of the knapsack changes, or as the deadline approaches. If the weights are not equal, the decision rules may exhibit counterintuitive behavior. But, if the distribution of the weight satisfies a "consistency condition," the decision rules become "well-behaved." The case where items have equal rewards, the general model with random rewards and random weights, and the case where the rewards are proportional to the weights were also analyzed. Numerical results were used to complement the theoretical results.

Several interesting and challenging opportunities for related research exist: the case with arrivals in continuous time, and a waiting cost for items that have already been accepted, have been studied in Kleywegt and Papastavrou (1995). Also, recall can be introduced, at a cost, for items that have been rejected. The variation of the problem where only the reward or the weight become known upon arrival has important applications; for example, when a reservation is made at a restaurant, the manager knows the number of people in the party, but can only guess the amount of money that the party will spend at the restaurant.[1]

[1] This research was supported by the National Science Foundation under grant DDM-9309579.

References

Albright, S. C., "Optimal Sequential Assignments with Random Arrival Times," *Management Sci.*, 21 (1974), 60–67.
Bruss, F. T., "A Unified Approach to a Class of Best Choice Problems with an Unknown Number of Options," *Ann. Probability*, 12 (1984), 882–889.
Carraway, R. L., R. L. Schmidt, and L. R. Weatherford, "An Algorithm for Maximizing Target Achievement in the Stochastic Knapsack Problem with Normal Returns," *Naval Res. Logistics*, 40, 2 (1993), 161–173.
Derman, C., G. J. Lieberman, and S. M. Ross, "A Sequential Stochastic Assignment Problem," *Management Sci.*, 18 (1972), 349–355.
Freeman, P. R., "The Secretary Problem and its Extensions. A Review," *International Statistical Rev.*, 51 (1983), 189–206.

Gallego, G. and G. van Ryzin, "Optimal Dynamic Pricing of Inventories with Stochastic Demand over Finite Horizons," *Management Sci.*, 40 (1994), 999–1020.

Hassin, R. and M. Henig, "Control of Arrivals and Departures in a State-Dependent Input-Output System," *Oper. Res. Letters*, 5 (1986), 33–36.

Henig, M., "Risk Criteria in a Stochastic Knapsack Problem," *Oper. Res.*, 38 (1990), 820–825.

Kaufman, J. F., "Blocking in Shared Resource Environment," *IEEE Trans. Communications*, 29 (1981), 1474–1481.

Kennedy, D. P., "Optimal Sequential Assignment," *Math. Oper. Res.*, 11 (1986), 619–626.

Kleywegt, A. J. and J. D. Papastavrou, "The Dynamic and Stochastic Knapsack Problem," Technical Report, School of Industrial Engineering, Purdue University, West Lafayette, IN, August 1995.

Martello, S. and P. Toth, *Knapsack Problems. Algorithms and Computer Implementations*, John Wiley & Sons, West Sussex, England, 1990.

Nakai, T., "An Optimal Selection Problem for a Sequence with a Random Number of Applicants per Period," *Oper. Res.*, 34 (1986a), 478–485.

——, "A Sequential Stochastic Assignment Problem in a Partially Observable Markov Chain," *Math. Oper. Res.*, 11 (1986b), 230–240.

——, "A Sequential Stochastic Assignment Problem in a Stationary Markov Chain," *Mathematica Japonica*, 31 (1986c), 741–757.

Papastavrou, J. D., S. Rajagopalan, and A. J. Kleywegt, "The Dynamic and Stochastic Knapsack Problem with Deadlines," Technical Report, School of Industrial Engineering, Purdue University, West Lafayette, IN, April 1995.

Prastacos, G. P., "Optimal Sequential Investment Decisions Under Conditions of Uncertainty," *Management Sci.*, 29 (1983), 118–134.

Presman, E. L. and I. M. Sonin, "The Best Choice Problem for a Random Number of Objects," *Theory of Probability and Its Applications*, 17 (1972), 657–668.

Righter, R., "A Resource Allocation Problem in a Random Environment," *Oper. Res.*, 37 (1989), 329–338.

Ross, K. W. and D. H. K. Tsang, "The Stochastic Knapsack Problem," *IEEE Trans. on Communications*, 37 (1989), 740–747.

—— and D. D. Yao, "Monotonicity Properties of the Stochastic Knapsack," *IEEE Trans. on Information Theory*, 36 (1990), 1173–1179.

Saario, V., "Limiting Properties of the Discounted House-Selling Problem," *European J. Oper. Res.*, 20 (1985), 206–210.

Sakaguchi, M., "Best Choice Problems for Randomly Arriving Offers During a Random Lifetime," *Mathematica Japonica*, 31 (1984a), 107–117.

——, "A Sequential Stochastic Assignment Problem Associated with a Non-homogeneous Markov Process," *Mathematica Japonica*, 31 (1984b), 13–22.

——, "A Sequential Stochastic Assignment Problem with an Unknown Number of Jobs," *Mathematica Japonica*, 31 (1984c), 141–152.

Sniedovich, M., "Preference Order Stochastic Knapsack Problems: Methodological Issues," *J. Oper. Res. Society*, 31 (1980), 1025–1032.

——, "Some Comments on Preference Order Dynamic Programming Models," *J. Math. Analysis & Applications*, 79 (1981), 489–501.

Steinberg, E. and M. S. Parks, "A Preference Order Dynamic Program for a Knapsack Problem with Stochastic Rewards," *J. Oper. Res. Society*, 30 (1979), 141–147.

Stewart, T. J., "The Secretary Problem with an Unknown Number of Options," *Oper. Res.*, 29 (1981), 130–145.

Tamaki, M., "A Full-Information Best-Choice Problem with Finite Memory," *J. Applied Probability*, 23 (1986a), 718–735.

——, "A Generalized Problem of Optimal Selection and Assignment," *Oper. Res.*, 34 (1986b), 486–493.

Weatherford, L. R. and S. E. Bodily, "A Taxonomy and Research Overview of Perishable-Asset Revenue Management: Yield Management, Overbooking, and Pricing," *Oper. Res.*, 40 (1992), 831–844.

Yasuda, M., "Asymptotic Results for the Best-Choice Problem with a Random Number of Objects," *J. Applied Probability*, 21 (1984), 521–536.

Accepted by Linda V. Green; received September 1994. This paper has been with the authors 3 weeks for 2 revisions.