Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

The Dynamic and Stochastic Knapsack Problem

Anton J. Kleywegt, Jason D. Papastavrou,

To cite this article:
Anton J. Kleywegt, Jason D. Papastavrou, (1998) The Dynamic and Stochastic Knapsack Problem. Operations Research 46(1):17-35. https://doi.org/10.1287/opre.46.1.17

ARTICLES

THE DYNAMIC AND STOCHASTIC KNAPSACK PROBLEM

ANTON J. KLEYWEGT

Georgia Institute of Technology, Atlanta, Georgia

and

JASON D. PAPASTAVROU

Purdue University, West Lafayette, Indiana

(Received January 1995; revision received February 1996; accepted October 1996)

The Dynamic and Stochastic Knapsack Problem (DSKP) is defined as follows. Items arrive according to a Poisson process in time. Each item has a demand (size) for a limited resource (the knapsack) and an associated reward. The resource requirements and rewards are jointly distributed according to a known probability distribution and become known at the time of the item's arrival. Items can be either accepted or rejected. If an item is accepted, the item's reward is received; and if an item is rejected, a penalty is paid. The problem can be stopped at any time, at which time a terminal value is received, which may depend on the amount of resource remaining. Given the waiting cost and the time horizon of the problem, the objective is to determine the optimal policy that maximizes the expected value (rewards minus costs) accumulated. Assuming that all items have equal sizes but random rewards, optimal solutions are derived for a variety of cost structures and time horizons, and recursive algorithms for computing them are developed. Optimal closed-form solutions are obtained for special cases. The DSKP has applications in freight transportation, in scheduling of batch processors, in selling of assets, and in selection of investment projects.

The *knapsack problem* has been extensively studied in operations research (Martello and Toth 1990). Items to be loaded into a knapsack with fixed capacity are selected from a given set of items with known sizes and rewards. The objective is to maximize the total reward, subject to capacity constraints. This problem is static and deterministic, because all the items are considered at a point in time, and their sizes and rewards are known a priori. However, in many practical applications, the knapsack problem is encountered in an uncertain and dynamically changing environment. Furthermore, there are often costs associated with delays that are not captured in the static knapsack problem. Applications of the dynamic and stochastic counterpart of the knapsack problem include:

1. In the transportation industry, ships, trains, aircraft, or trucks often carry loads for different clients. Transportation requests arrive stochastically over time, and prices are offered or negotiated for transporting loads. If a load is accepted, costs are incurred for picking up and handling the load, and for the administrative activities involved. These costs are specific to the load, and can be subtracted from the price to give the "reward" of the load. If a load is rejected, some customer goodwill (possible future sales) is lost, which can be taken into account with a penalty for rejecting loads. Loads may have different sizes—such as parcels—or the same size—such as containers. Often there is a fixed schedule for moving vehicles and a deadline after which loads cannot be accepted for a specific shipment. Even when there is not a fixed schedule, an incentive exists to consolidate and dispatch shipments with high frequency, to maintain short delivery times, and to maximize the rate at which revenue is earned with the given investment in capital and labor costs. This incentive can be modeled with a discount rate, and a waiting cost or holding cost per unit time that is incurred until the shipment is dispatched. The waiting cost may be constant or may depend on the number of loads accepted, but not yet dispatched. The dispatcher can decide to dispatch a vehicle at any time before the deadline. There is also a dispatching and transportation cost that is incurred for the shipment as a whole, that may depend on the number of loads in the shipment.
2. A scheduler of a batch processor has to schedule jobs with random capacity requirements and rewards as they arrive over time. Fixed schedules or customer commitments lead to deadlines. The pressure to increase the utilization of equipment and labor, and to maintain a high level of customer service, lead to a waiting cost per unit time. The cost of running the batch processor may depend on the number of jobs in the batch.
3. A real estate agent selling new condominiums receives

Subject classifications: Transportation, dynamic and stochastic loading. Resource allocation: on-line in a probabilistic environment.
Area of review: DISTRIBUTION.

0030-364X/98/4617-0017 $05.00

offers stochastically over time and may want to sell the assets before winter or before the new tax year. Hence, the agent faces a deadline, possibly with a salvage value for the unsold assets. There is also an opportunity cost associated with the capital tied into the unsold assets, and property taxes, which cause a waiting cost per unit time to exist.

4. An investor who wishes to invest a certain amount of funds faces a similar problem. The investor is presented with investment projects with random arrival times, funding requirements, and returns. The opportunity cost of unutilized capital is represented by a waiting cost per unit time, and the objective is to maximize the expected value earned from investing the funds.

These problems are characterized by the allocation of limited resources to competing items that arrive randomly over time. Items are associated with resource requirements as well as rewards, which may include any item specific costs incurred. Usually the arrival times, resource requirements, and rewards are unknown before arrival, and become known upon arrival. Arriving items can be either accepted or rejected. Incentives such as a deadline after which arriving items cannot be accepted, discounting, and a waiting cost per unit time serve to encourage the timely acceptance of items. The problem can be stopped at any time before or at the deadline. There may also be a cost associated with the group of accepted items as a whole, or a salvage value for unused resources, which may depend on the amount of resources allocated to the accepted items. A typical objective is to maximize the expected total value (rewards minus costs). We call problems of this general nature the *Dynamic and Stochastic Knapsack Problem* (DSKP).

In this paper different versions of this problem are formulated and analyzed for the case where all items have equal size, for both the infinite and finite horizon cases. (The case where items have random sizes is analyzed in Kleywegt and Papastavrou 1995.) We show that an optimal acceptance rule is given by a simple threshold rule. It is also shown how to find an optimal stopping time. We derive structural characteristics of the optimal value function and the optimal acceptance threshold, and propose recursive algorithms to compute optimal solutions. Closed-form solutions are obtained for some cases.

In Section 1, previous research on similar problems is reviewed. In Section 2, the DSKP is defined and notation is introduced, and general results are derived in Section 3. The DSKP without a deadline is considered in Section 4, and the DSKP with a deadline is considered in Section 5. Our concluding remarks follow in Section 6.

1. RELATED RESEARCH

Stochastic versions of the knapsack problem can be classified as either static or dynamic. In static stochastic knapsack problems the set of items is given, but the rewards and/or sizes are unknown. Steinberg and Parks (1979) proposed a preference order dynamic programming algorithm for the knapsack problem with random rewards. Sniedovich (1980, 1981) further investigated preference order dynamic programming and pointed out that the preference relations used by Steinberg and Parks may lead to suboptimal solutions. Other preference relations may lead to the failure of an optimal solution to exist, or to a trivial optimal solution. Henig (1990) combined dynamic programming and a search procedure to solve stochastic knapsack problems where the items have known sizes and independent normally distributed rewards. Carraway et al. (1993) proposed a hybrid dynamic programming/branch-and-bound algorithm for a stochastic knapsack problem similar to that of Henig, with an objective that maximizes the probability of target achievement.

In dynamic stochastic knapsack problems the items arrive over time, and the rewards and/or sizes are unknown before arrival. Decisions are made sequentially as items arrive.

Some stopping time problems and best choice (optimal selection) problems are similar to the DSKP. A well-known example is the *secretary problem*, where candidates arrive over time. The objective is to maximize the probability of choosing the best candidate or k best candidates from a given, or random, number of candidates, or to maximize the expected value of the chosen candidates. These problems have been studied by Presman and Sonin (1972), Stewart (1981), Freeman (1983), Yasuda (1984), Bruss (1984), Nakai (1986a), Sakaguchi (1986), and Tamaki (1986a, 1986b).

The problem of selling a single asset, where offers arrive periodically (Rosenfield et al. 1983), or according to a renewal process (Mamer 1986), with a fixed waiting cost, with or without a deadline, has also been studied. Albright (1977) studied a house selling problem where a given number, n, of offers are received, and $k \leqslant n$ houses are to be sold. Asymptotic properties of an optimal policy for the house selling problem with discrete time periods, as the deadline and number of houses become large, were derived by Saario (1985).

A more general 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 and Derman (1972), Albright (1974), Sakaguchi (1984a and 1984b), Nakai (1986b and 1986c), and Kennedy (1986).

Some resource allocation problems are similar to the DSKP. Mendelson et al. (1980) investigated the problem of allocating a given amount of resource to a set of activity classes, with demands following a known probability distribution, and arriving according to a renewal process. The objective is to maximize the expected time until the resource allocated to an activity is depleted. Righter (1989) studied a resource allocation problem that is an extension of the SSAP.

Many investment problems can be regarded as DSKPs. 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 in discrete time. We incorporate a waiting cost, in addition to the issues taken into account by Prastacos, and arrivals occur either according to a geometric process in discrete time, or according to a Poisson process in continuous time. Also, Prastacos assumed that each investment opportunity was large enough to absorb all the available capital, but in our problem the sizes of the investment opportunities are given, and cannot be chosen.

Other versions of the DSKP have been studied for communications applications by Kaufman (1981), Ross and Tsang (1989), and Ross and Yao (1990). Papastavrou et al. (1995) studied a version of DSKP similar to that in this paper, with different sized items, with arrivals occurring periodically in discrete time, and without waiting costs.

A class of problems similar to the DSKP have been termed *Perishable Asset Revenue Management* (PARM) problems by Weatherford and Bodily (1992) and Weatherford et al. (1993). These problems are often called *yield management* problems, and have been studied extensively, with specific application to airline seat inventory control and hotel yield management by Rothstein (1971, 1974, 1985), Shlifer and Vardi (1975), Alstrup et al. (1986), Belobaba (1987, 1989), Dror et al. (1988), Curry (1990), Brumelle et al. (1990), Brumelle and McGill (1993), Wollmer (1992), Lee and Hersh (1993), and Robinson (1995). In most of these problems there are a number of different fare classes, which are usually assumed to be given due to competition. The objective is to dynamically assign the available capacity to the different fare classes to maximize expected revenues. In the DSKP, the available capacity is dynamically assigned to arriving demands with random rewards and random resource requirements. Another type of PARM problem, in which an inventory has to be sold before a deadline, has been studied by Kincaid and Darling (1963), Stadje (1990), and Gallego and Van Ryzin (1994). In their problems customers arrive according to a Poisson process with price-dependent probability of purchasing. The major differences 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 the selling of real estate; whereas in the models of Stadje and Gallego and Van Ryzin, prices are set and all demands are accepted as long as supplies last, which is typical in retail. Also, we incorporate a waiting cost and an option to stop before the deadline.

2. PROBLEM DEFINITION

Items arrive according to a Poisson process in time. Each item has an associated reward. The reward of an item is unknown prior to arrival, and becomes known upon arrival. The distribution of the rewards is known and is independent of the arrival time and of the rewards of other arrivals. In this paper it is assumed that items have equal capacity requirements (sizes). Without loss of generality, let the size of each item be 1, and the known initial capacity be integer. The items are to be included in a knapsack of known capacity. Each arriving item can be either accepted or rejected. If an item is accepted, the reward associated with the item is received, and if the item is rejected, a penalty is incurred. Once an item is rejected, it cannot be recalled.

There is a known deadline (possibly infinite) after which items can no longer be accepted. It is allowed to stop waiting for arrivals before the capacity is exhausted or the deadline is reached (for example, when a vehicle is dispatched without filling it to capacity and before the deadline is reached). There is a waiting cost per unit time that depends on the number of items already accepted, or equivalently, on the remaining capacity. A terminal value is earned that depends on the remaining capacity at the stopping time. Rewards and costs may be discounted. The objective is to determine a policy for accepting items and for stopping that maximizes the expected total (discounted) value (rewards minus costs) accumulated.

Let $\{A_i\}_{i=1}^{\infty}$ denote the arrival times of a Poisson process on $(0, \infty)$ with rate $\lambda \in (0, \infty)$. Let R_i denote the reward of arrival i, and assume that $\{R_i\}_{i=1}^{\infty}$ is an i.i.d. sequence, independent of $\{A_i\}_{i=1}^{\infty}$. Let F_R denote the probability distribution of R, and assume that $E[R] < \infty$. Let $(\Omega, \mathcal{F}, P)$ be a probability space satisfying these assumptions. Let N_0 denote the initial capacity, and let $\mathcal{N} \equiv \{0, 1, \ldots, N_0\}$. Let $T \in (0, \infty]$ denote the deadline for accepting items, and let $\mathcal{T} \in [0, T]$ denote the stopping time. Let D_i denote the decision whether to accept or reject arrival i, defined as follows:

$$D_i \equiv \begin{cases} 1 & \text{if arrival } i \text{ is accepted,} \\ 0 & \text{if arrival } i \text{ is rejected.} \end{cases}$$

Let $\mathcal{I}_s$ denote the set of all unit step functions $f_s : (0, T] \mapsto \{0, 1\}$ of the form

$$f_s(t) \equiv \begin{cases} 1 & \text{if } t \in (0, \tau], \\ 0 & \text{if } t \in (\tau, T], \end{cases}$$

for some $\tau \in [0, T]$.

The class $\Pi_{\text{DSKP}}^{\text{HD}}$ of history-dependent deterministic policies for the DSKP is defined as follows. For any $t \in [0, \infty)$, let $\mathcal{H}_t$ be the history of the process $\{A_i, R_i\}$ up to time t (i.e., the σ-algebra generated by $\{(A_i, R_i) : A_i \leq t\}$), denoted $\mathcal{H}_t \equiv \sigma(\{(A_i, R_i) : A_i \leq t\})$. Let $\mathcal{H}_{t^-} \equiv \sigma(\{(A_i, R_i) : A_i < t\})$, let $\mathcal{H}_\infty \equiv \sigma(\{(A_i, R_i)\}_{i=1}^{\infty})$, $\mathcal{H}_\infty \subset \mathcal{F}$, and let $\mathcal{H}_{A_i} \equiv \{B \in \mathcal{H}_\infty : B \cap \{A_i \leq t\} \in \mathcal{H}_t \forall t \in [0, \infty]\}$. Let $\mathcal{A} \equiv \{\{A_i\}_{i=1}^{\infty} : 0 < A_1 < A_2 < \cdots < \infty\}$, let $\mathcal{R} \equiv \{\{R_i\}_{i=1}^{\infty} :$

$R_i \in \Re\}$, and let $\mathcal{D} \equiv \{\{D_i\}_{i=1}^{\infty} : D_i \in \{0, 1\}\}$. Define $\Pi_{\text{DSKP}}^{\text{HD}}$ as the set of all Borel-measurable functions $\pi : \mathcal{A} \times \mathcal{R} \mapsto \mathcal{D} \times \mathcal{I}_s$ which satisfy the conditions

$$D_i^\pi \text{ is } \mathcal{H}_{A_i} \text{ measurable, i.e., } \{D_i^\pi = 1\} \in \mathcal{H}_{A_i} \text{ for all } i \in \{1, 2, \cdots\},$$

$$I^\pi(t) \text{ is } \mathcal{H}_{t^-} \text{ measurable, i.e., } \{I^\pi(t) = 1\} \in \mathcal{H}_{t^-} \text{ for all } t \in (0, T],$$

$$\sum_{\{i: A_i \leqslant \mathcal{T}^\pi\}} D_i^\pi \leqslant N_0,$$

where $(\{D_i^\pi\}, I^\pi) \equiv \pi(\{A_i\}, \{R_i\})$, and the stopping time $\mathcal{T}^\pi$ is given by

$$I^\pi(t) = \begin{cases} 1 & \text{if } t \in (0, \mathcal{T}^\pi], \\ 0 & \text{if } t \in (\mathcal{T}^\pi, T]. \end{cases}$$

Let $N^\pi(t)$ denote the remaining capacity under policy π at time t, where N^π is defined to be left-continuous, i.e.,

$$N^\pi(t) \equiv N_0 - \sum_{\{i: A_i < t\}} D_i^\pi I^\pi(A_i), \tag{1}$$

and let

$$N^\pi(t^+) \equiv N_0 - \sum_{\{i: A_i \leqslant t\}} D_i^\pi I^\pi(A_i). \tag{2}$$

Let $c(n)$ denote the waiting cost per unit time while the remaining capacity is n. Let p denote the penalty that is incurred if an item is rejected. Let $v(n)$ denote the terminal value that is earned at time $\mathcal{T}^\pi$ if the remaining capacity $N^\pi(\mathcal{T}^{\pi+}) = n$. Let α be the discount rate; if $T = \infty$, we require that $\alpha > 0$.

Let V_{DSKP}^π denote the expected total discounted value under policy $\pi \in \Pi_{\text{DSKP}}^{\text{HD}}$, i.e.,

$$\begin{aligned} V_{\text{DSKP}}^\pi \equiv E\Bigg[& \sum_{\{i: A_i \leqslant \mathcal{T}^\pi\}} e^{-\alpha A_i}[D_i^\pi R_i - (1 - D_i^\pi)p] \\ & - \int_0^{\mathcal{T}^\pi} e^{-\alpha\tau} c(N^\pi(\tau))\, d\tau \\ & + e^{-\alpha \mathcal{T}^\pi} v(N^\pi(\mathcal{T}^{\pi+})) \Bigg| N^\pi(0^+) = N_0 \Bigg] \\ = E\Bigg[& \sum_{\{i: A_i \leqslant T\}} e^{-\alpha A_i}[D_i^\pi R_i - (1 - D_i^\pi)p] I^\pi(A_i) \\ & + \int_0^T e^{-\alpha\tau}[-c(N^\pi(\tau)) I^\pi(\tau) \\ & \qquad + \alpha v(N^\pi(\tau))(1 - I^\pi(\tau))]\, d\tau \\ & + e^{-\alpha T} v(N^\pi(T^+)) \Bigg| N^\pi(0^+) = N_0 \Bigg]. \end{aligned}$$

The objective is to find the optimal expected value V_{DSKP}^*, i.e.,

$$V_{\text{DSKP}}^* \equiv \sup_{\pi \in \Pi_{\text{DSKP}}^{\text{HD}}} V_{\text{DSKP}}^\pi,$$

and to find an optimal policy $\pi^* \in \Pi_{\text{DSKP}}^{\text{HD}}$ that achieves this optimal value, if such a policy exists.

A summary of the most important notation is given in Table I.

Table I
Summary of Notation

λ	item arrival rate, $\lambda \in (0, \infty)$
A_i	arrival time of item i
R_i, r	item reward
F_R	probability distribution of item rewards R, $F_R(0) < 1$
T	deadline
$\mathcal{T}$	stopping time
n	remaining capacity
t	time
$N(t)$	remaining capacity at time t
$c(n)$	waiting cost per unit time while $N(t) = n$
p	penalty for rejecting an item
$v(n)$	terminal value with $N(\mathcal{T}^+) = n$
α	discount rate
π	policy
$D^\pi(n, t, r)$	acceptance decision rule of policy π
$I^\pi(n, t)$	stopping decision rule of policy π
$V^\pi(n, t)$	expected value of policy π
$x^\pi(n, t)$	acceptance threshold used by policy π

3. GENERAL RESULTS

The relation between the *Dynamic and Stochastic Knapsack Problem* (DSKP) and a closely related continuous time Markov Decision Process (MDP) is investigated. The option of choosing a stopping time for the DSKP introduces a complexity into the DSKP that is not modeled in a straightforward way by an MDP, unless we introduce a *stopped* state, with an infinite rate transition to this state as soon as the decision is made to stop. However, for most of the useful results for MDPs, transition rates have to be bounded. We therefore study an MDP which is a relaxation of the DSKP, in that the MDP can switch *off* and *on* multiple times, instead of stopping only once, which can be modeled with bounded transition rates. We show that there exists an optimal policy for the MDP which stops only once, and hence which is admissible and optimal for the DSKP.

The MDP has state space $\mathcal{N}$. The policy spaces $\Pi_{\text{MDP}}^{\text{HD}}$, $\Pi_{\text{MDP}}^{\text{MD}}$, and $\Pi_{\text{MDP}}^{\text{SD}}$, are defined hereafter, where superscript HD denotes history-dependent deterministic policies, MD denotes memoryless deterministic policies, and SD denotes stationary deterministic policies. Let $\mathcal{I}_I$ denote the set of all Borel-measurable functions $f_I : (0, T] \mapsto \{0, 1\}$. The class $\Pi_{\text{MDP}}^{\text{HD}}$ is defined as the set of all Borel-measurable functions $\pi : \mathcal{A} \times \mathcal{R} \mapsto \mathcal{D} \times \mathcal{I}_I$ which satisfy the conditions

$$D_i^\pi \text{ is } \mathcal{H}_{A_i} \text{ measurable for all } i \in \{1, 2, \cdots\},$$

$$I^\pi(t) \text{ is } \mathcal{H}_{t^-} \text{ measurable for all } t \in (0, T],$$

$$\sum_{\{i: A_i \leqslant T\}} D_i^\pi I^\pi(A_i) \leqslant N_0,$$

where $(\{D_i^\pi\}, I^\pi) \equiv \pi(\{A_i\}, \{R_i\})$. Note that the MDP is allowed to switch *on* $(I^\pi(t) = 1)$ and *off* $(I^\pi(t) = 0)$ multiple times, in contrast with the DSKP, which has to remain *off* once it stops. Hence, with V_{MDP}^π properly defined, the MDP is a relaxation of the DSKP. The optimal expected value of the MDP is therefore at least as good as that of the DSKP. This is the result of Lemma 1, which follows after the definitions of V_{MDP}^π and V_{MDP}^*.

V_{MDP}^π denotes the expected total discounted value under policy $\pi \in \Pi_{\text{MDP}}^{\text{HD}}$, given by

$$
\begin{aligned}
V_{\text{MDP}}^\pi \equiv E\Bigg[& \sum_{\{i:\, A_i \leqslant T\}} e^{-\alpha A_i}[D_i^\pi R_i - (1 - D_i^\pi) p] I^\pi(A_i) \\
& + \int_0^T e^{-\alpha\tau}[-c(N^\pi(\tau)) I^\pi(\tau) \\
& + \alpha v(N^\pi(\tau))(1 - I^\pi(\tau))]\, d\tau \\
& + e^{-\alpha T} v(N^\pi(T^+)) \Big| N^\pi(0^+) = N_0 \Bigg].
\end{aligned}
$$

The optimal expected value V_{MDP}^* is given by

$$V_{\text{MDP}}^* \equiv \sup_{\pi \in \Pi_{\text{MDP}}^{\text{HD}}} V_{\text{MDP}}^\pi$$

Lemma 1. $V_{\text{MDP}}^* \geqslant V_{\text{DSKP}}^*$.

Proof. $\Pi_{\text{DSKP}}^{\text{HD}} \subset \Pi_{\text{MDP}}^{\text{HD}}$ because $\mathcal{I}_s \subset \mathcal{I}_I$. For any $\pi \in \Pi_{\text{DSKP}}^{\text{HD}}$, $V_{\text{MDP}}^\pi = V_{\text{DSKP}}^\pi$. Hence, $V_{\text{MDP}}^* \equiv \sup_{\pi \in \Pi_{\text{MDP}}^{\text{HD}}} V_{\text{MDP}}^\pi \geqslant \sup_{\pi \in \Pi_{\text{DSKP}}^{\text{HD}}} V_{\text{MDP}}^\pi = \sup_{\pi \in \Pi_{\text{DSKP}}^{\text{HD}}} V_{\text{DSKP}}^\pi \equiv V_{\text{DSKP}}^*$. $\square$

Let $\mathcal{I}_R$ denote the set of all Borel-measurable functions $f_R : \Re \mapsto \{0, 1\}$. The class $\Pi_{\text{MDP}}^{\text{MD}}$ of memoryless deterministic policies for the MDP is defined as the set of all Borel-measurable functions $\pi : \mathcal{N} \times (0, T] \mapsto \mathcal{I}_R \times \{0, 1\}$, where $\pi \equiv (D^\pi, I^\pi)$, and D^π and I^π are as follows. $D^\pi(n, t, r)$ denotes the decision under policy π whether to accept or reject an arrival i at time $A_i = t$ with reward $R_i = r$ if the remaining capacity $N^\pi(t) = n$, defined as follows:

$$D^\pi(n, t, r) \equiv \begin{cases} 1 & \text{if } n > 0 \text{ and arrival } i \text{ is accepted,} \\ 0 & \text{if } n = 0 \text{ or arrival } i \text{ is rejected.} \end{cases}$$

Let the acceptance set for policy π be denoted by $\mathcal{R}_1^\pi(n, t) \equiv \{r \in \Re : D^\pi(n, t, r) = 1\}$, and the rejection set be denoted by $\mathcal{R}_0^\pi(n, t) \equiv \{r \in \Re : D^\pi(n, t, r) = 0\}$. $I^\pi(n, t)$ denotes the decision under policy π whether to be switched *on* or *off* at time t if the remaining capacity $N^\pi(t) = n$, defined as follows:

$$I^\pi(n, t) \equiv \begin{cases} 1 & \text{if switched } on \text{ at time } t, \\ 0 & \text{if switched } off \text{ at time } t. \end{cases}$$

It is easy to show that $\Pi_{\text{MDP}}^{\text{MD}} \subset \Pi_{\text{MDP}}^{\text{HD}}$.

The remaining capacity corresponding to policy π is given by

$$N^\pi(t) = N_0 - \sum_{\{i:\, A_i < t\}} D^\pi(N^\pi(A_i), A_i, R_i) I^\pi(N^\pi(A_i), A_i).$$

The MDP can be modeled with transition rates

$$\lambda(\pi(n, t)) \equiv \lambda I^\pi(n, t),$$

and transition probabilities

$$P[n | n, \pi(n, t)] \equiv \int_{\mathcal{R}_0^\pi(n, t)} dF_R(r),$$

$$P[n - 1 | n, \pi(n, t)] \equiv \int_{\mathcal{R}_1^\pi(n, t)} dF_R(r).$$

Let $V_{\text{MDP}}^\pi(n, t)$ be the expected total discounted value under policy $\pi \in \Pi_{\text{MDP}}^{\text{MD}}$ from time t until time T, if the remaining capacity $N^\pi(t^+) = n$, i.e.,

$$
\begin{aligned}
V_{\text{MDP}}^\pi(n, t) \equiv E\Bigg[& \sum_{\{i:\, A_i \in (t, T]\}} e^{-\alpha(A_i - t)} \\
& \cdot [D^\pi(N^\pi(A_i), A_i, R_i) R_i \\
& \quad - (1 - D^\pi(N^\pi(A_i), A_i, R_i)) p] \\
& \cdot I^\pi(N^\pi(A_i), A_i) + \int_t^T e^{-\alpha(\tau - t)} \\
& \cdot [-c(N^\pi(\tau)) I^\pi(N^\pi(\tau), \tau) \\
& + \alpha v(N^\pi(\tau))(1 - I^\pi(N^\pi(\tau), \tau))]\, d\tau \\
& + e^{-\alpha(T - t)} v(N^\pi(T^+)) \Big| N^\pi(t^+) = n \Bigg] \\
= E\Bigg[\int_t^T & e^{-\alpha(\tau - t)} \Bigg\{ \lambda \Bigg[\int_{\mathcal{R}_1^\pi(N^\pi(\tau), \tau)} dF_R(r) \\
& - p \int_{\mathcal{R}_0^\pi(N^\pi(\tau), \tau)} dF_R(r) \Bigg] \\
& \cdot I^\pi(N^\pi(\tau), \tau) - c(N^\pi(\tau)) I^\pi(N^\pi(\tau), \tau) \\
& + \alpha v(N^\pi(\tau))(1 - I^\pi(N^\pi(\tau), \tau)) \Bigg\}\, d\tau \\
& + e^{-\alpha(T - t)} v(N^\pi(T^+)) \Big| N^\pi(t^+) = n \Bigg]. \quad (3)
\end{aligned}
$$

The equality follows from an integration theorem for point processes; see for example Brémaud (1981, Theorem II.T8). Let $V_{\text{MDP}}^*(n, t)$ be the corresponding optimal expected value, i.e.,

$$V_{\text{MDP}}^*(n, t) \equiv \sup_{\pi \in \Pi_{\text{MDP}}^{\text{MD}}} V_{\text{MDP}}^\pi(n, t).$$

Note that $V_{\text{MDP}}^*(n, t) \geqslant v(n)$ for all n and t, because the policy $\pi \in \Pi_{\text{MDP}}^{\text{MD}}$ with $I^\pi = 0$ has $V_{\text{MDP}}^\pi(n, t) = v(n)$ for all n and t.

Intuitively we would expect V_{MDP}^* to decrease as the deadline comes closer. This is the result of Proposition 1. In Proposition 2 it is shown that V_{MDP}^* is nondecreasing in n if c is nonincreasing and v is nondecreasing. These are the conditions that usually hold in applications. It is typical for the waiting cost to increase as the number of accepted customers increases, and for the terminal value to decrease (for example, for the dispatching and transportation cost to increase) as the final number of customers increases. Proofs can be found in Kleywegt and Papastavrou (1995).

Proposition 1. *For any $n \in \mathcal{N}$, $V^*_{\text{MDP}}(n, t)$ is nonincreasing in t on* $[0, T]$.

Proposition 2. *If c is nonincreasing and v is nondecreasing, then for any $t \in [0, T]$, $V^*_{\text{MDP}}(n, t)$ is nondecreasing in n on $\mathcal{N}$.*

As for policies $\pi \in \Pi^{\text{HD}}_{\text{MDP}}$, policies $\pi \in \Pi^{\text{MD}}_{\text{MDP}}$ are allowed to switch *on* and *off* multiple times. However, consider policies $\pi \in \Pi^{\text{MD}}_{\text{MDP}}$ with stopping rules $I^\pi(n, \cdot) \in \mathcal{I}_s$ for each $n \in \mathcal{N}$, i.e., $I^\pi(n, \cdot)$ is a unit step function of t of the form

$$I^{\pi}(n, t) \equiv \begin{cases} 1 & \text{if } t \in (0, \tau^{\pi}(n)], \\ 0 & \text{if } t \in (\tau^{\pi}(n), T], \end{cases}$$

for some $\tau^\pi(n) \in [0, T]$ for each $n \in \mathcal{N}$. Such policies π are admissible for the DSKP ($\pi \in \Pi^{\text{HD}}_{\text{DSKP}}$), because once the process switches *off*, it remains stopped. For each such policy π, the sample path $N^\pi(\omega)$ is the same for the DSKP and the MDP for each $\omega \in \Omega$, and $V^\pi_{\text{DSKP}} = V^\pi_{\text{MDP}}$. Intuitively we expect that there is an optimal policy $\pi^* \in \Pi^{\text{MD}}_{\text{MDP}}$ with such a unit step function stopping rule I^*, for the following reason. For any $t_1 \in [0, T]$ such that $V^*_{\text{MDP}}(n, t_1) > v(n)$, it holds that $V^*_{\text{MDP}}(n, t) > v(n)$ for all $t \in [0, t_1]$, because V^*_{MDP} is nonincreasing in t from Proposition 1. Hence, if the remaining capacity is n, it is optimal to continue waiting (i.e., $I^*(n, t) = 1$) for all $t \in (0, t_1]$. Similarly, for any $t_1 \in [0, T]$ such that $V^*_{\text{MDP}}(n, t_1) = v(n)$, it holds that $V^*_{\text{MDP}}(n, t) = v(n)$ and it is optimal to stop (i.e., $I^*(n, t) = 0$) for all $t \in [t_1, T]$. It is shown that there is a policy $\pi^* \in \Pi^{\text{MD}}_{\text{MDP}}$ that has such a unit step function stopping rule I^*, and that is optimal among all policies $\pi \in \Pi^{\text{HD}}_{\text{MDP}}$. From this it follows that π^* is also an optimal policy for the DSKP among all policies $\pi \in \Pi^{\text{HD}}_{\text{DSKP}}$.

A policy $\pi \in \Pi^{\text{MD}}_{\text{MDP}}$ is said to be a threshold policy if it has a threshold acceptance rule D^π with a reward threshold $x^\pi : \mathcal{N}\backslash\{0\} \times (0, T] \mapsto \Re$. If the reward r of an item arriving at time t when the remaining capacity $N^\pi(t) = n > 0$, is greater than $x^\pi(n, t)$, then the item is accepted; otherwise the item is rejected. That is,

$$D^{\pi}(n, t, r) \equiv \begin{cases} 1 & \text{if } n > 0 \text{ and } r > x^{\pi}(n, t), \\ 0 & \text{if } n = 0 \text{ or } r \leq x^{\pi}(n, t). \end{cases}$$

The following argument suggests that threshold $x^*(n, t) = V^*(n, t) - V^*(n-1, t) - p$ gives an optimal acceptance rule. Suppose an item with reward r arrives at time t when the remaining capacity $N^*(t) = n > 0$, and $I^*(n, t) = 1$. If the item is accepted, the optimal expected value from then on is $r + V^*(n-1, t)$. If the item is rejected, the optimal expected value from then on is $V^*(n, t) - p$. Hence, the item is accepted if $r + V^*(n-1, t) > V^*(n, t) - p$, i.e., if $r > V^*(n, t) - V^*(n-1, t) - p$; otherwise the item is rejected. It is shown that there is a threshold policy π^* with threshold $x^*(n, t) = V^*(n, t) - V^*(n-1, t) - p$ that is optimal among all policies $\pi \in \Pi^{\text{HD}}_{\text{MDP}}$.

The class $\Pi^{\text{SD}}_{\text{MDP}}$ of stationary deterministic policies for the MDP is the subset of $\Pi^{\text{MD}}_{\text{MDP}}$ of policies π that do not depend on t. Stationary policies have unit step function stopping rules with $\tau^\pi(n) = T$ if $I^\pi(n) = 1$, and $\tau^\pi(n) = 0$ if $I^\pi(n) = 0$. Therefore, for any $\pi \in \Pi^{\text{SD}}_{\text{MDP}}$, $V^\pi_{\text{DSKP}} = V^\pi_{\text{MDP}}$.

In order to derive some characteristics of π^* and V^*, consider the function $f : \Re \mapsto \Re$ defined by

$$f(y) \equiv \int_y^\infty (r - y)\, dF_R(r) = \int_y^\infty (1 - F_R(r))\, dr$$

The function f can be interpreted as $f(y) = P[R > y]\, E[R - y | R > y]$.

Lemma 2.

1. *f satisfies the following Lipschitz condition*:

$$|f(y_2) - f(y_1)| \leq |y_2 - y_1| \qquad \text{for all } y_1, y_2 \in \Re.$$

2. *f is absolutely continuous on $\Re$.*
3. *f is nonincreasing on $\Re$, and strictly decreasing on $\{y \in \Re : F_R(y) < 1\}$.*
4. *For any $\epsilon > 0$, there exists a y_1 such that $f(y_1) > \epsilon$.*
5. *For any $\epsilon > 0$, there exists a y_2 such that $f(y_2) < \epsilon$.*
6.

$$f(y) = \sup_{B \in \mathcal{B}} \int_B (r - y)\, dF_R(r),$$

where $\mathcal{B}$ is the Borel sets on $\Re$.

Proofs can be found in Kleywegt and Papastavrou (1995).

4. THE INFINITE HORIZON DSKP

It was shown by Yushkevic and Fainberg (1979, Theorem 2) for the MDP with an infinite horizon that if $\alpha > 0$, then for any $\epsilon > 0$ there is a stationary deterministic policy $\pi \in \Pi^{\text{SD}}_{\text{MDP}}$ that is ϵ-optimal among all history-dependent deterministic policies $\pi \in \Pi^{\text{HD}}_{\text{MDP}}$. Therefore, we restrict attention to the class of policies $\Pi^{\text{SD}}_{\text{MDP}}$. Because $V^\pi_{\text{DSKP}} = V^\pi_{\text{MDP}}$ for any $\pi \in \Pi^{\text{SD}}_{\text{MDP}}$, we will drop the subscripts of V in this section. For $\pi \in \Pi^{\text{SD}}_{\text{MDP}}$, D^π is a function of n and r only, and I^π and V^π are functions of n only. This also means that stopping times are restricted to the starting time and the times when the remaining capacity changes, i.e., those arrival times when items are accepted, and that a stopping capacity m^π can be derived for a policy $\pi \in \Pi^{\text{SD}}_{\text{MDP}}$ from its stopping rule I^π as follows:

$$m^{\pi} \equiv \max\{n \in \mathcal{N} : I^{\pi}(n) = 0\}.$$

If $I^\pi(0) = 0$, then $V^\pi(0) = v(0)$. If $I^\pi(0) = 1$, then

$$\alpha V^{\pi}(0) = -[\lambda p + c(0)]. \tag{4}$$

For $n > 0$, if $I^\pi(n) = 0$, then $V^\pi(n) = v(n)$. If $I^\pi(n) = 1$, then by conditioning on the arrival time A_k and the reward R_k of the first arrival k after time t, it follows that

$$
\begin{aligned}
V^{\pi}(n) = V^{\pi}(n, t) = {} & -\frac{1}{\alpha+\lambda}\, c(n) + \frac{\lambda}{\alpha+\lambda} \\
& \cdot \left[-p \int_{\Re_0^{\bar\pi}(n)} dF_R(r_k) + \int_{\Re_1^{\bar\pi}(n)} r_k \; dF_R(r_k) \right] \\
& + \int_{\Re_0^{\bar\pi}(n)} \int_t^{\infty} \lambda e^{-(\alpha+\lambda)(a_k - t)} \\
& \cdot E\Bigg[\sum_{\{i: A_i \in (a_k, \mathcal{T}^{\pi}]\}} e^{-\alpha(A_i - a_k)} [D^{\pi}(N^{\pi}(A_i), R_i) R_i \\
& \qquad - (1 - D^{\pi}(N^{\pi}(A_i), R_i)) p] \\
& \quad - \int_{a_k}^{\mathcal{T}^{\pi}} e^{-\alpha(\tau - a_k)} c(N^{\pi}(\tau)) \; d\tau \\
& \quad + e^{-\alpha(\mathcal{T}^{\pi} - a_k)} v(N^{\pi}(\mathcal{T}^{\pi+})) \\
& \quad \cdot \Bigg| A_k = a_k, \; R_k = r_k \Bigg] \; da_k \; dF_R(r_k) \\
& + \int_{\Re_1^{\bar\pi}(n)} \int_t^{\infty} \lambda e^{-(\alpha+\lambda)(a_k - t)} \\
& \cdot E\Bigg[\sum_{\{i: A_i \in (a_k, \mathcal{T}^{\pi}]\}} e^{-\alpha(A_i - a_k)} [D^{\pi}(N^{\pi}(A_i), R_i) R_i \\
& \qquad - (1 - D^{\pi}(N^{\pi}(A_i), R_i)) p] \\
& \quad - \int_{a_k}^{\mathcal{T}^{\pi}} e^{-\alpha(\tau - a_k)} c(N^{\pi}(\tau)) \; d\tau \\
& \quad + e^{-\alpha(\mathcal{T}^{\pi} - a_k)} v(N^{\pi}(\mathcal{T}^{\pi+})) \\
& \quad \cdot \Bigg| A_k = a_k, \; R_k = r_k \Bigg] \; da_k \; dF_R(r_k).
\end{aligned}
$$

From Equation (3), independence of $\{A_i\}$ and $\{R_i\}$, and the memoryless arrival process, it follows that for $\pi \in \Pi^{\mathrm{SD}}_{\mathrm{MDP}}$

$$
\begin{aligned}
& \int_{\Re_0^{\bar\pi}(n)} \int_t^{\infty} \lambda e^{-(\alpha+\lambda)(a_k - t)} \\
& \cdot E\Bigg[\sum_{\{i: A_i \in (a_k, \mathcal{T}^{\pi}]\}} e^{-\alpha(A_i - a_k)} [D^{\pi}(N^{\pi}(A_i), R_i) R_i \\
& \qquad - (1 - D^{\pi}(N^{\pi}(A_i), R_i)) p] \\
& \quad - \int_{a_k}^{\mathcal{T}^{\pi}} e^{-\alpha(\tau - a_k)} c(N^{\pi}(\tau)) \; d\tau \\
& \quad + e^{-\alpha(\mathcal{T}^{\pi} - a_k)} v(N^{\pi}(\mathcal{T}^{\pi+})) \Bigg| A_k = a_k, \; R_k = r_k \Bigg] \\
& \quad \cdot \; da_k \; dF_R(r_k) \\
& = \frac{\lambda}{\alpha+\lambda} V^{\pi}(n) \int_{\Re_0^{\bar\pi}(n)} dF_R(r_k),
\end{aligned}
$$

and

$$
\begin{aligned}
& \int_{\Re_1^{\bar\pi}(n)} \int_t^{\infty} \lambda e^{-(\alpha+\lambda)(a_k - t)} \\
& \cdot E\Bigg[\sum_{\{i: A_i \in (a_k, \mathcal{T}^{\pi}]\}} e^{-\alpha(A_i - a_k)} [D^{\pi}(N^{\pi}(A_i), R_i) R_i \\
& \qquad - (1 - D^{\pi}(N^{\pi}(A_i), R_i)) p] \\
& \quad - \int_{a_k}^{\mathcal{T}^{\pi}} e^{-\alpha(\tau - a_k)} c(N^{\pi}(\tau)) \; d\tau \\
& \quad + e^{-\alpha(\mathcal{T}^{\pi} - a_k)} v(N^{\pi}(\mathcal{T}^{\pi+})) \Bigg| A_k = a_k, \; R_k = r_k \Bigg] \\
& \cdot \; da_k \; dF_R(r_k) \\
& = \frac{\lambda}{\alpha+\lambda} V^{\pi}(n-1) \int_{\Re_1^{\bar\pi}(n)} dF_R(r_k).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
V^{\pi}(n) = {} & -\frac{1}{\alpha+\lambda}\, c(n) + \frac{\lambda}{\alpha+\lambda} \\
& \cdot \left[-p \int_{\Re_0^{\bar\pi}(n)} dF_R(r) + \int_{\Re_1^{\bar\pi}(n)} r \; dF_R(r) \right] \\
& + \frac{\lambda}{\alpha+\lambda} \left[V^{\pi}(n) \int_{\Re_0^{\bar\pi}(n)} dF_R(r) \right. \\
& \qquad \left. + V^{\pi}(n-1) \int_{\Re_1^{\bar\pi}(n)} dF_R(r) \right]
\end{aligned}
$$

$$
\begin{aligned}
\text{f } \alpha V^{\pi}(n) = {} & \lambda \int_{\Re_1^{\bar\pi}(n)} \{ r - [V^{\pi}(n) - V^{\pi}(n-1) - p] \} \\
& \cdot \; dF_R(r) - [\lambda p + c(n)]. \qquad (5)
\end{aligned}
$$

It also follows that

$$
V^{\pi}(n) = \frac{\lambda \int_{\Re_1^{\bar\pi}(n)} [r + V^{\pi}(n-1) + p] \; dF_R(r) - [\lambda p + c(n)]}{\alpha + \lambda \int_{\Re_1^{\bar\pi}(n)} dF_R(r)}.
$$

Consider the following equation:

$$
\begin{aligned}
\alpha y & = \lambda \int_{y - y_3}^{\infty} [r - (y - y_3)] \; dF_R(r) - y_4 \\
& = \lambda f(y - y_3) - y_4. \qquad (6)
\end{aligned}
$$

Lemma 3. *For any given y_3 and y_4, if $\alpha > 0$ or if $\alpha = 0$ and $y_4 > 0$, then Equation* (6) *has a unique solution y.*

Proof.

Case 1: $\alpha > 0$. Then αy is strictly increasing in y, and takes on all values in $\Re$. Also, from Lemma 2, $\lambda f(y - y_3) - y_4$ is nonincreasing and continuous in y. Therefore, from the intermediate value theorem, there is a unique value y such that $\alpha y = \lambda f(y - y_3) - y_4$.

Case 2: $\alpha = 0$. From Lemma 2, for any $y_4 > 0$, there exists a y_1 such that $\lambda f(y_1 - y_3) > y_4$, and a y_2 such that $\lambda f(y_2 - y_3) < y_4$. Hence, from the continuity of f and the intermediate value theorem, there is at least one value y such that $\lambda f(y - y_3) = y_4 > 0$. For any such y, $F_R(y - y_3) < 1$. Thus, from Lemma 2, f is strictly decreasing at y, and nonincreasing everywhere. Therefore, there is a unique value y such that $\lambda f(y - y_3) - y_4 = 0 = \alpha y$. □

Inductively define the sequence of threshold policies $\{\psi(n)\}_{n=0}^{N_0}$ as follows. $D^{\psi(0)}(0,\cdot) = 0$, $I^{\psi(0)}(0) = 1$. $V^{\psi(0)}(0)$ is given by Equation (4). Let $\psi(n-1)$ and $V^{\psi(n-1)}(n-1)$ be defined, and let $\hat{V}(n)$ be the unique solution to

$$\begin{aligned}\alpha \hat{V}(n) &= \lambda \int_{\hat{V}(n)-\max\{v(n-1),V^{\psi(n-1)}(n-1)\}-p}^{\infty} \\ &\quad \cdot \{r - [\hat{V}(n) - \max\{v(n-1), V^{\psi(n-1)}(n-1)\} \\ &\quad - p]\}\ dF_R(r) - [\lambda p + c(n)] \\ &= \lambda f(\hat{V}(n) - \max\{v(n-1),\ V^{\psi(n-1)}(n-1)\} - p) \\ &\quad - [\lambda p + c(n)], \end{aligned} \tag{7}$$

which exists by Lemma 3. Let $I^{\psi(n)}(n) = 1$, and $x^{\psi(n)}(n) = \hat{V}(n) - \max\{v(n-1), V^{\psi(n-1)}(n-1)] - p$. Let $D^{\psi(n)}(n', \cdot) = D^{\psi(n-1)}(n', \cdot)$ and $I^{\psi(n)}(n') = I^{\psi(n-1)}(n')$ for all $n' \in \{0, 1, \ldots, n-1\}$, except that

$$I^{\psi(n)}(n-1) = \begin{cases} 1 & \text{if } V^{\psi(n-1)}(n-1) > v(n-1), \\ 0 & \text{if } V^{\psi(n-1)}(n-1) \leqslant v(n-1) \end{cases}$$

Hence, $V^{\psi(n)}(n-1) = \max\{v(n-1), V^{\psi(n-1)}(n-1)\}$. It follows from Equation (5) that $V^{\psi(n)}(n)$ satisfies

$$\begin{aligned}\alpha V^{\psi(n)}(n) &= \lambda \int_{\hat{V}(n)-\max\{v(n-1),\ V^{\psi(n-1)}(n-1)\}-p}^{\infty} \\ &\quad \cdot\ \{r - [V^{\psi(n)}(n) - \max\{v(n-1), \\ &\quad \cdot V^{\psi(n-1)}(n-1)\} - p]\} dF_R(r) \\ &\quad \cdot - [\lambda p + c(n)]. \end{aligned} \tag{8}$$

From Equation (7), Equation (8) has a solution $V^{\psi(n)}(n) = \hat{V}(n)$, and it can easily be shown that this is the unique solution of Equation (8). Therefore,

$$\begin{aligned}\alpha V^{\psi(n)}(n) &= \lambda \int_{V^{\psi(n)}(n)-\max\{v(n-1),V^{\psi(n-1)}(n-1)\}-p}^{\infty} \\ &\quad \cdot\ \{r - [V^{\psi(n)}(n) - \max\{v(n-1), \\ &\quad V^{\psi(n-1)}(n-1)\} - p]\}\ dF_R(r) \\ &\quad - [\lambda p + c(n)] \\ &= \lambda f(x^{\psi(n)}(n)) - [\lambda p + c(n)]. \end{aligned} \tag{9}$$

By definition, $V^*(n) \geqslant \max\{v(n), V^{\psi(n)}(n)\}$ for all n. Theorem 1 shows that $V^*(n) = \max\{v(n), V^{\psi(n)}(n)\}$ for all n. Therefore an optimal policy is as follows. For each n, if $V^{\psi(n)}(n) > v(n)$, then continue (i.e., $I^*(n) = 1$), using threshold $x^*(n) = x^{\psi(n)}(n) = V^{\psi(n)}(n) - \max\{v(n-1), V^{\psi(n-1)}(n-1)\} - p = V^*(n) - V^*(n-1) - p$, else stop (i.e., $I^*(n) = 0$), and collect $v(n)$. This result is useful, not only because it gives a clear, intuitive characterization of an optimal policy and the optimal expected value, but also because it provides a straightforward method for computing the optimal expected value V^* and optimal threshold x^*.

Theorem 1. *The optimal expected value V^* satisfies $V^*(n) = \max\{v(n), V^{\psi(n)}(n)\}$ for all $n \in \{0, 1, \ldots, N_0\}$.*

Proof. By induction on n. For $n = 0$ it is clear that $V^*(0) = \max\{v(0), V^{\psi(0)}(0)\}$. Suppose $V^*(n-1) = \max\{v(n-1), V^{\psi(n-1)}(n-1)\}$. Hence, $V^{\psi(n)}(n-1) = \max\{v(n-1), V^{\psi(n-1)}(n-1)\} = V^*(n-1)$.

Case 1. $V^\pi(n) > v(n)$ for some policy $\pi \in \Pi^{SD}_{MDP}$: Consider any such policy π. Then $I^\pi(n) = 1$, and $V^\pi(n-1) \leqslant V^*(n-1) = V^{\psi(n)}(n-1)$. It is shown by contradiction that $V^\pi(n) \leqslant V^{\psi(n)}(n)$. Suppose $V^\pi(n) > V^{\psi(n)}(n)$. From Equation (5) and Lemma 2

$$\begin{aligned}\alpha V^\pi(n) &\leqslant \lambda \int_{V^\pi(n)-V^\pi(n-1)-p}^{\infty} \\ &\quad \{r - [V^\pi(n) - V^\pi(n-1) - p]\}\ dF_R(r) \\ &\quad - [\lambda p + c(n)] \\ &\leqslant \lambda \int_{V^{\psi(n)}(n)-V^{\psi(n)}(n-1)-p}^{\infty} \\ &\quad \{r - [V^{\psi(n)}(n) - V^{\psi(n)}(n-1) - p]\}\ dF_R(r) \\ &\quad - [\lambda p + c(n)] \\ &= \alpha V^{\psi(n)}(n), \end{aligned}$$

which contradicts the assumption.

Case 2. $V^\pi(n) \leqslant v(n)$ for every policy $\pi \in \Pi^{SD}_{MDP}$: Then $V^*(n) = v(n) = \max\{v(n), V^{\psi(n)}(n)\}$. □

Theorem 2. *The following stationary deterministic threshold policy π^* is an optimal policy among all history-dependent deterministic policies for the MDP and DSKP. An optimal stopping rule is*

$$I^*(n) = \begin{cases} 1 & \text{if } V^{\psi(n)}(n) > v(n), \\ 0 & \text{if } V^{\psi(n)}(n) \leqslant v(n). \end{cases}$$

An optimal acceptance rule for $n > 0$ is

$$D^*(n, r) = \begin{cases} 1 & \text{if } r > V^*(n) - V^*(n-1) - p, \\ 0 & \text{if } r \leqslant V^*(n) - V^*(n-1) - p. \end{cases}$$

Proof. From Theorem 1, π^* is optimal among all $\pi \in \Pi^{SD}_{MDP}$. From Yushkevic and Fainberg (1979, Theorem 2), for any $\epsilon > 0$, there is a $\pi \in \Pi^{SD}_{MDP}$ that is ϵ-optimal among all $\pi \in \Pi^{HD}_{MDP}$. Hence, π^* is optimal for MDP among all π

$\in \Pi^{\text{HD}}_{\text{MDP}}$. From Lemma 1, $V^*_{\text{MDP}} \geqslant V^*_{\text{DSKP}}$. But π^* is admissible for DSKP ($\pi^* \in \Pi^{\text{HD}}_{\text{DSKP}}$), and $V^{\pi^*}_{\text{DSKP}} = V^{\pi^*}_{\text{MDP}} = V^*_{\text{MDP}} \geqslant V^*_{\text{DSKP}}$. Therefore, π^* is optimal for DSKP among all $\pi \in \Pi^{\text{HD}}_{\text{DSKP}}$. $\square$

An optimal policy π^* is not a unique optimal policy, because $D^{\pi^*}(n, \cdot)$ can be modified on any set with F_R-measure 0, without changing the expected value V^{π^*}. Hence, there exist optimal policies which are not threshold policies. The threshold $x^*(n) = V^*(n) - V^*(n-1) - p$ is the unique optimal threshold if and only if for all $n > 0$ and for every $a < x^*(n)$ and for every $b > x^*(n)$, $P[a < R < x^*(n)] > 0$ and $P[x^*(n) < R < b] > 0$.

4.1. Choice of Initial Capacity

Suppose the initial capacity M can be chosen from a set $\mathcal{M}$ of available capacities, $\mathcal{M} \subseteq \{0, \ldots, N_0\}$. This is typical when an initial size is chosen for a ship, a truck, or a batch processor, from a number of sizes available in the market. Once in operation, the option exists to stop waiting and transport or process the items, even if the capacity of the vehicle or processor is not exhausted. An optimal initial capacity M^* is given by

$$M^* \in \operatorname{argmax}_{M \in \mathcal{M}} \{V^*(M)\},$$

and an optimal stopping capacity m^* is then given by

$$m^* = \max\{m \in \{0, 1, \ldots, M^*\} : V^{\psi(m)}(m) \leqslant v(m)\}.$$

The following algorithm computes the optimal expected value V^*, an optimal threshold x^*, an optimal initial capacity M^*, and an optimal stopping capacity m^* in $\Theta(N_0)$ time, if solving Equation (9) is counted as an operation for each value of n.

Algorithm Infinite-Horizon-Knapsack

compute $V^{\psi(0)}(0)$ from Equation (4);
if $V^{\psi(0)}(0) > v(0)$ **then**
    $V^*(0) = V^{\psi(0)}(0)$; $m = -1$; $m^* = -1$;
else
    $V^*(0) = v(0)$; $m = 0$; $m^* = 0$;
endif;
$M^* = 0$;
for $n = 1$ **to** N_0
    solve Equation (9) for $V^{\psi(n)}(n)$;
    if $V^{\psi(n)}(n) > v(n)$ **then**
      $V^*(n) = V^{\psi(n)}(n)$; $x^*(n) = V^*(n) - V^*(n-1) - p$;
    else
      $V^*(n) = v(n)$; $m = n$;
    endif;
    if $V^*(n) > V^*(M^*)$ **and** $n \in \mathcal{M}$ **then**
        $M^* = n$; $m^* = m$;
    endif;
endfor:

4.2. Example

From Lemma 3, $V^{\psi(n)}$ is well defined if $\alpha = 0$ and $\lambda p + c(n) > 0$ for all n. An exponential reward distribution may be appealing in light of the rule of thumb known as the 80–20 principle ("ABC analysis" in inventory management; Zenz 1987). If the rewards are exponentially distributed with mean $1/\mu$, $\alpha = 0$, c and v are constant, and $\lambda p + c > 0$, then

$$x^{\psi(n)}(n) = V^{\psi(1)}(1) - v - p = \frac{1}{\mu} \ln\left[\frac{\lambda}{\mu(\lambda p + c)}\right],$$

for all $n > 0$, and

$$\begin{aligned} V^{\psi(n)}(n) &= nV^{\psi(1)}(1) - (n-1)v \\ &= \frac{n}{\mu} \ln\left[\frac{\lambda}{\mu(\lambda p + c)}\right] + np + v. \end{aligned}$$

5. THE FINITE HORIZON DSKP

In this section V^π denotes V^π_{MDP}, unless noted otherwise. It will be shown that $V^*_{\text{DSKP}} = V^*_{\text{MDP}}$. A differential equation satisfied by the expected value $V^\pi(n, t)$ under a policy $\pi \in \Pi^{\text{MD}}_{\text{MDP}}$ can be derived intuitively as follows. If $I^\pi(n, t) = 1$, then by conditioning on whether an arrival takes place in the next Δt time units, and on the reward r of the item if there is an arrival, we obtain

$$\begin{aligned} V^\pi(n, t) &= (1-\alpha\Delta t)\Bigg\{\lambda\Delta t\Bigg[\int_{\mathcal{R}_1^\pi(n,t)} [r + V^\pi(n-1, t+\Delta t)]\, dF_R(r) \\ &\qquad + \int_{\mathcal{R}_0^\pi(n,t)} [V^\pi(n, t+\Delta t) - p]\, dF_R(r)\Bigg] \\ &\qquad + (1 - \lambda\Delta t)V^\pi(n, t+\Delta t) - c(n)\Delta t\Bigg\} \\ &\qquad + o(\Delta t) \end{aligned}$$

$$\begin{aligned} \text{f} \quad \frac{V^\pi(n, t) - V^\pi(n, t+\Delta t)}{\Delta t} &= (1 - \alpha\Delta t)\lambda\Bigg[\int_{\mathcal{R}_1^\pi(n,t)} [r + V^\pi(n-1, t+\Delta t)]\, dF_R(r) \\ &\qquad + [V^\pi(n, t+\Delta t) - p] \cdot \int_{\mathcal{R}_0^\pi(n,t)} dF_R(r)\Bigg] \\ &\quad + (-\alpha - \lambda + \alpha\lambda\Delta t)V^\pi(n, t+\Delta t) \\ &\quad - (1 - \alpha\Delta t)c(n) + \frac{o(\Delta t)}{\Delta t}, \end{aligned}$$

where $o(\Delta t)/\Delta t$ 3 0 as Δt 3 0, from the corresponding property for the Poisson process, and $E[R] < \infty$. Letting Δt 3 0,

$$\frac{\partial V^{\pi}(n,\,t)}{\partial t} = -\lambda\left[\int_{\mathcal{R}_1^{\pi}(n,t)} [r + V^{\pi}(n-1,\,t)]\,dF_R(r) + [V^{\pi}(n,\,t) - p]\int_{\mathcal{R}_0^{\pi}(n,t)} dF_R(r)\right] - (-\alpha - \lambda)V^{\pi}(n,\,t) + c(n)$$

$$= -\lambda\int_{\mathcal{R}_1^{\pi}(n,t)} \{r - [V^{\pi}(n,\,t) - V^{\pi}(n-1,\,t) - p]\} \cdot dF_R(r) + \alpha V^{\pi}(n,\,t) + \lambda p + c(n). \qquad (10)$$

If $I^{\pi}(n,\,t) = 0$, then

$$V^{\pi}(n,t) = (1 - \alpha\Delta t)V^{\pi}(n, t+\Delta t) + \alpha\Delta t v(n)$$

$$\text{f} \quad \frac{V^{\pi}(n,t) - V^{\pi}(n,t+\Delta t)}{\Delta t} = -\alpha V^{\pi}(n, t+\Delta t) + \alpha v(n)$$

$$\text{f} \quad \frac{\partial V^{\pi}(n,t)}{\partial t} = \alpha V^{\pi}(n,t) - \alpha v(n).$$

The boundary condition is $V^{\pi}(n,\,T) = v(n)$.

These differential equations can also be derived from the results in Pliska (1975), where the existence of a unique absolutely continuous solution for each policy $\pi \in \Pi_{\mathrm{MDP}}^{\mathrm{MD}}$ is shown, or in Brémaud (1981), where these equations are called the Hamilton-Jacobi equations. Note that if $I^{\pi}(n,\,t) = 0$ for all $t \in (t_1,\,T)$, then $V^{\pi}(n,\,t) = v(n)$ for all $t \in [t_1,\,T]$, as for the DSKP.

From the results in Pliska (1975), Yushkevic and Fainberg (1979), or Brémaud (1981), it follows that the optimal expected value V^* is the unique absolutely continuous solution of

$$-\frac{\partial V^*(n,\,t)}{\partial t} = \sup_{(D(n,t,\cdot),I(n,t))}\left\{\left[\lambda\int_{\mathcal{R}_1(n,t)} \{r - [V^*(n,t) - V^*(n-1,t) - p]\}\,dF_R(r) - \alpha V^*(n,t) - \lambda p - c(n)\right] I(n,t) + [-\alpha V^*(n,t) + \alpha v(n)](1 - I(n,t))\right\}, \qquad (11)$$

with boundary condition $V^*(n,\,T) = v(n)$.

Consider the threshold policy $\pi^* \in \Pi_{\mathrm{MDP}}^{\mathrm{MD}}$ with threshold x^* for $n > 0$ given by

$$x^*(n,\,t) = V^*(n,\,t) - V^*(n-1,\,t) - p.$$

The stopping rule is given by

$$I^*(0,\,t) = \begin{cases} 0 & \text{if } -\lambda p - c(0) < \alpha v(0), \\ 1 & \text{if } -\lambda p - c(0) \geqslant \alpha v(0), \end{cases}$$

and for $n > 0$

$$I^*(n,\,t) = \begin{cases} 0 & \text{if } \lambda f(x^*(n,\,t)) - \lambda p - c(n) < \alpha v(n), \\ 1 & \text{if } \lambda f(x^*(n,\,t)) - \lambda p - c(n) \geqslant \alpha v(n). \end{cases}$$

Theorem 3. *The memoryless deterministic threshold policy π^* is an optimal policy among all history-dependent deterministic policies for MDP.*

Proof. From Equation (11) and Lemma 2

$$-\frac{\partial V^*(n,\,t)}{\partial t} = \max\left\{\sup_{\mathcal{R}_1(n,t)\in\mathcal{B}} \lambda\int_{\mathcal{R}_1(n,t)} \{r - [V^*(n,\,t) - V^*(n-1,\,t) - p]\}\,dF_R(r) - \alpha V^*(n,\,t) - \lambda p - c(n),\; -\alpha V^*(n,\,t) + \alpha v(n)\right\}$$

$$= \max\left\{\lambda\int_{V^*(n,t)-V^*(n-1,t)-p}^{\infty} \{r - [V^*(n,\,t) - V^*(n-1,\,t) - p]\}\,dF_R(r) - \alpha V^*(n,\,t) - \lambda p - c(n),\; -\alpha V^*(n,\,t) + \alpha v(n)\right\},$$

or

$$\frac{\partial V^*(n,\,t)}{\partial t} = \min\{-\lambda f(x^*(n,\,t)) + \alpha V^*(n,\,t) + \lambda p + c(n),\; \alpha V^*(n,\,t) - \alpha v(n)\}. \qquad (12)$$

Hence, the sup in the expression for $\partial V^*(n,\,t)/\partial t$ is attained by policy π^*. Thus, V^{π^*} and V^* satisfy the same differential equation with the same boundary condition. Therefore, $V^{\pi^*} = V^*$, and policy π^* is optimal among all memoryless deterministic policies for MDP. From Yushkevic and Fainberg (1979, Theorem 1), for any $\epsilon > 0$, there exists a memoryless deterministic policy that is ϵ-optimal among all history-dependent deterministic policies. Hence, policy π^* is also optimal among all history-dependent deterministic policies for MDP. $\square$

Proposition 3. *For each n, $I^*(n,\,\cdot)$ is a unit step function of the form*

$$I^*(n,\,t) = \begin{cases} 1 & \text{if } t \in (0,\,\tau^*(n)], \\ 0 & \text{if } t \in (\tau^*(n),\,T], \end{cases}$$

where $\tau^*(n) \in [0,\,T]$.

Proof. For $n = 0$, $I^*(0,\,t)$ is independent of t. If $I^*(0,\,t) = 1$ for all $t \in (0,\,T]$, then $\tau^*(0) = T$. Else, if $I^*(0,\,t) = 0$ for all $t \in (0,\,T]$, then $\tau^*(0) = 0$. For $n > 0$, consider the following two cases.

Case 1. $\alpha > 0$: Let $t_1 \equiv \sup\{t \in [0,\,T] : V^*(n,\,t) > v(n)\}$. From Proposition 1, V^* is nonincreasing in t, hence $V^*(n,\,t) > v(n)$ for all $t \in [0,\,t_1)$, and $V^*(n,\,t) = v(n)$ for all $t \in [t_1,\,T]$. Thus, for all $t \in [0,\,t_1)$, $\alpha V^*(n,\,t) - \alpha v(n) > 0$. But from Proposition 1, $\partial V^*(n,\,t)/\partial t \leqslant 0$, hence $\partial V^*(n,\,t)/\partial t = -\lambda f(x^*(n,\,t)) + \alpha V^*(n,\,t) + \lambda p + c(n) < \alpha V^*(n,$

$t) - \alpha v(n)$. Thus $\lambda f(x^*(n, t)) - \lambda p - c(n) > \alpha v(n)$, and $I^*(n, t) = 1$ for all $t \in [0, t_1)$. If $t_1 = T$, then $\lambda f(x^*(n, T)) - \lambda p - c(n) \geqslant \alpha v(n)$, from the continuity of f and V^* in t, hence $I^*(n, T) = 1$ and $\tau^*(n) = T$. If $t_1 < T$, then for $t \in (t_1, T)$, $\partial V^*(n, t)/\partial t = 0 = \alpha V^*(n, t) - \alpha v(n) \leqslant -\lambda f(x^*(n, t)) + \alpha V^*(n, t) + \lambda p + c(n)$, and $\lambda f(x^*(n, t_1)) - \lambda p - c(n) = \alpha v(n)$ from the continuity of f and V^* in t. For $t \in [t_1, T]$, $f(x^*(n, t)) = f(V^*(n, t) - V^*(n-1, t) - p) = f(v(n) - V^*(n-1, t) - p)$, and is nonincreasing in t, since V^* is nonincreasing in t and f is nonincreasing. Let $t_2 \equiv \sup\{t \in [t_1, T] : \lambda f(x^*(n, t)) - \lambda p - c(n) = \alpha v(n)\}$. Then $\lambda f(x^*(n, t)) - \lambda p - c(n) \geqslant \alpha v(n)$ and $I^*(n, t) = 1$ for all $t \in [0, t_2]$, and $\lambda f(x^*(n, t)) - \lambda p - c(n) < \alpha v(n)$ and $I^*(n, t) = 0$ for all $t \in (t_2, T]$. Therefore, $\tau^*(n) = t_2$, and the result holds.

Case 2. $\alpha = 0$: By contradiction. Suppose there exists an $n > 0$ and $0 < t_s < t_c \leqslant T$ such that $I^*(n, t_s) = 0$ and $I^*(n, t_c) = 1$, i.e., $\lambda f(x^*(n, t_s)) - \lambda p - c(n) < 0$ and $\lambda f(x^*(n, t_c)) - \lambda p - c(n) \geqslant 0$. From the continuity of f and V^* in t, it follows that there exists $t_b \in (t_s, t_c]$ such that $\lambda f(x^*(n, t)) - \lambda p - c(n) < 0$ for all $t \in [t_s, t_b)$, and $\lambda f(x^*(n, t_b)) - \lambda p - c(n) = 0$. Then $\partial V^*(n, t)/\partial t = 0$ for all $t \in [t_s, t_b)$, and V^* is absolutely continuous in t, hence $V^*(n, t_s) = V^*(n, t_b)$. Because V^* is nonincreasing in t, $V^*(n-1, t_s) \geqslant V^*(n-1, t_b)$. Thus, $\lambda f(x^*(n, t_s)) - \lambda p - c(n) = \lambda f(V^*(n, t_s) - V^*(n-1, t_s) - p) - \lambda p - c(n) \geqslant \lambda f(V^*(n, t_b) - V^*(n-1, t_b) - p) - \lambda p - c(n) = 0$. But this contradicts the assumption that $\lambda f(x^*(n, t_s)) - \lambda p - c(n) < 0$. Therefore, $\lambda f(x^*(n, t)) - \lambda p - c(n) \geqslant 0$ and $I^*(n, t) = 1$ for all $t \in (0, \tau^*(n)]$, and $\lambda f(x^*(n, t)) - \lambda p - c(n) < 0$ and $I^*(n, t) = 0$ for all $t \in (\tau^*(n), T]$ for some $\tau^*(n) \in [0, T]$. □

Theorem 4. *The memoryless deterministic threshold policy π^* is an optimal policy among all history-dependent deterministic policies for the DSKP.*

Proof. From Theorem 3, π^* is optimal for the MDP among all $\pi \in \Pi^{\mathrm{HD}}_{\mathrm{MDP}}$. From Proposition 3, π^* satisfies $I^*(n, \cdot) \in \mathcal{I}_s$ for all n; hence π^* is admissible for the DSKP ($\pi^* \in \Pi^{\mathrm{HD}}_{\mathrm{DSKP}}$), and $V^{\pi^*}_{\mathrm{DSKP}} = V^{\pi^*}_{\mathrm{MDP}}$. From Lemma 1, $V^*_{\mathrm{MDP}} \geqslant V^*_{\mathrm{DSKP}}$. Therefore, $V^{\pi^*}_{\mathrm{DSKP}} = V^{\pi^*}_{\mathrm{MDP}} = V^*_{\mathrm{MDP}} \geqslant V^*_{\mathrm{DSKP}}$, and π^* is optimal for the DSKP among all $\pi \in \Pi^{\mathrm{HD}}_{\mathrm{DSKP}}$. □

5.1. Structural Characteristics

A number of interesting structural characteristics of the optimal expected value V^* and optimal threshold x^* are derived in this section. First a characterization is given of an optimal policy and the optimal expected value that holds under typical conditions. This characterization is useful because it gives a simple, intuitive recipe for following an optimal policy, and it simplifies computation of V^* and x^*. Thereafter, some monotonicity and concavity properties are shown to hold. Also interesting are the counterintuitive cases where certain properties do not hold, which can be found in Kleywegt and Papastavrou (1995).

The properties of V^* depend to a large extent on the relative magnitudes of $\lambda f(-p)$ and $\lambda p + c(n) + \alpha v(n)$. The importance of these quantities makes intuitive sense, by noting that $\lambda[f(-p) - p] = \lambda[\int_{-p}^{\infty} (r+p)\, dF_R(r) - p] = \lambda[\int_{-p}^{\infty} r\, dF_R(r) - p\int_{-\infty}^{-p} dF_R(r)] = \lambda(P[R > -p]\, E[R \mid R > -p] - pP[R \leqslant -p])$, and by interpreting $\lambda P[R > -p]E[R \mid R > -p]$ as the effective reward rate while we continue to wait, and comparing it with $\lambda p P[R \leqslant -p] + c(n) + \alpha v(n)$, the rate at which (opportunity) cost is incurred while we continue to wait.

Proposition 4. *If c is nonincreasing, v is nondecreasing, and $\lambda f(-p) \leqslant \lambda p + c(n) + \alpha v(n)$ for an $n > 0$, then $V^*(n, t) = v(n)$ for all $t \in [0, T]$.*

Proof. From Proposition 2, if c is nonincreasing and v is nondecreasing, then V^* is nondecreasing in n. Hence, $V^*(n, t) - V^*(n-1, t) \geqslant 0$ for all $t \in [0, T]$. Thus

$$\lambda f(V^*(n, t) - V^*(n-1, t) - p) - \alpha V^*(n, t) - \lambda p - c(n)$$
$$\leqslant \lambda f(-p) - \alpha V^*(n, t) - \lambda p - c(n) \leqslant -\alpha V^*(n, t) + \alpha v(n).$$

Therefore, $\partial V^*(n, t)/\partial t = \alpha V^*(n, t) - \alpha v(n)$, and $V^*(n, t) = v(n)$ for all $t \in [0, T]$. □

Similar to the infinite horizon case, inductively define the sequence of threshold policies $\{\psi(n)\}_{n=0}^{N_0}$. $D^{\psi(0)}(0, t) = 0$, $I^{\psi(0)}(0, t) = 1$ for all $t \in (0, T]$. Then

$$V^{\psi(0)}(0, t) = e^{-\alpha(T-t)} v(0) - \frac{\lambda p + c(0)}{\alpha}\left(1 - e^{-\alpha(T-t)}\right),$$

if $\alpha > 0$, and $V^{\psi(0)}(0, t) = -[\lambda p + c(0)](T - t) + v(0)$ if $\alpha = 0$. Let $\psi(n-1)$ and $V^{\psi(n-1)}(n-1, \cdot)$ be defined, and let $\hat{V}(n, \cdot)$ satisfy

$$\begin{aligned}
\frac{\partial \hat{V}(n, t)}{\partial t}
&= -\lambda \int_{\hat{V}(n,t) - \max\{v(n-1), V^{\psi(n-1)}(n-1,t)\} - p}^{\infty} \\
&\qquad \{r - [\hat{V}(n, t) - \max\{v(n-1), V^{\psi(n-1)}(n-1, t)\} \\
&\qquad - p]\}\, dF_R(r) + \alpha \hat{V}(n, t) + \lambda p + c(n) \\
&= -\lambda f(\hat{V}(n, t) - \max\{v(n-1), V^{\psi(n-1)}(n-1, t)\} - p) \\
&\qquad + \alpha \hat{V}(n, t) + \lambda p + c(n),
\end{aligned} \tag{13}$$

for $t \in (0, T)$ with boundary condition $\hat{V}(n, T) = v(n)$. It is shown in Kleywegt and Papastavrou (1995) that Equation (13) has a unique absolutely continuous solution $\hat{V}(n, \cdot)$. Let $I^{\psi(n)}(n, t) = 1$, and $x^{\psi(n)}(n, t) = \hat{V}(n, t) - \max\{v(n-1), V^{\psi(n-1)}(n-1, t)\} - p$. Let $D^{\psi(n)}(n', t, \cdot) = D^{\psi(n-1)}(n', t, \cdot)$ and $I^{\psi(n)}(n', t) = I^{\psi(n-1)}(n', t)$ for all $n' \in \{0, 1, \ldots, n-1\}$ and all $t \in (0, T]$, except that

$$I^{\psi(n)}(n-1, t) = \begin{cases} 1 & \text{if } V^{\psi(n-1)}(n-1, t) > v(n-1), \\ 0 & \text{if } V^{\psi(n-1)}(n-1, t) \leqslant v(n-1). \end{cases}$$

Lemma 4. *For all $n > 0$ and all $t \in [0, T]$,*

$V^{\psi(n)}(n-1,\,t) = \max\{v(n-1),\ V^{\psi(n-1)}(n-1,\,t)\}$.

The proof can be found in Kleywegt and Papastavrou (1995).

It follows from Equation (10) that $V^{\psi(n)}(n, t)$ satisfies

$$\begin{aligned}\frac{\partial V^{\psi(n)}(n,\,t)}{\partial t} &= -\lambda \int_{\hat{V}(n,t)-\max\{v(n-1),V^{\psi(n-1)}(n-1,t)\}-p}^{\infty} \\ &\quad \{r-[V^{\psi(n)}(n,t)-\max\{v(n-1),\,V^{\psi(n-1)}(n-1,t)\}-p]\}\,dF_R(r) \\ &\quad +\alpha V^{\psi(n)}(n,t)+\lambda p+c(n), \qquad (14)\end{aligned}$$

for $t \in (0, T)$ with boundary condition $V^{\psi(n)}(n, T) = v(n)$. From Equation (13), Equation (14) has a solution $V^{\psi(n)}(n, \cdot) = \hat{V}(n, \cdot)$, and it can easily be shown that this is the unique solution of Equation (14). Therefore,

$$\begin{aligned}\frac{\partial V^{\psi(n)}(n,\,t)}{\partial t} &= -\lambda f(x^{\psi(n)}(n,\,t)) \\ &\quad + \alpha V^{\psi(n)}(n,\,t) + \lambda p + c(n), \qquad (15)\end{aligned}$$

with $x^{\psi(n)}(n, t) = V^{\psi(n)}(n, t) - \max\{v(n-1), V^{\psi(n-1)}(n-1, t)\} - p$.

Proposition 5. *If v is nonincreasing and $\lambda f(-p) > \lambda p + c(n) + \alpha v(n)$ for an $n > 0$, then $V^{\psi(n)}(n, t) > v(n)$ for all $t \in [0, T)$.*

Proof. By contradiction. Suppose there exists $t_1 \in [0, T)$ such that $V^{\psi(n)}(n, t_1) \leqslant v(n)$. Then $x^{\psi(n)}(n, t_1) = V^{\psi(n)}(n, t_1) - \max\{v(n-1), V^{\psi(n-1)}(n-1, t_1)\} - p \leqslant -p$. From Lemma 2, $-\lambda f(x^{\psi(n)}(n, t_1)) + \alpha V^{\psi(n)}(n, t_1) + \lambda p + c(n) \leqslant -\lambda f(-p) + \alpha v(n) + \lambda p + c(n) < 0$. Then, from the continuity of f and $V^{\psi(n)}$ in t, there exists a neighborhood $(t_0, t_2) \subseteq (0, T)$ of t_1 such that $-\lambda f(x^{\psi(n)}(n, t)) + \alpha V^{\psi(n)}(n, t) + \lambda p + c(n) < 0$ for all $t \in (t_0, t_2)$. Then from Equation (15)

$$\begin{aligned}\frac{\partial V^{\psi(n)}(n,\,t)}{\partial t} &= -\lambda f(x^{\psi(n)}(n,\,t)) \\ &\quad + \alpha V^{\psi(n)}(n,\,t) + \lambda p + c(n) < 0,\end{aligned}$$

for all $t \in (t_0, t_2)$. Thus $V^{\psi(n)}(n, t)$ is strictly decreasing on $[t_1, T]$. This implies that $V^{\psi(n)}(n, T) < V^{\psi(n)}(n, t_1) \leqslant v(n)$, which violates the boundary condition $V^{\psi(n)}\,(n, T) = v(n)$. Therefore, $V^{\psi(n)}(n, t) > v(n)$ for all$t \in [0, T)$. □

Corollary 1. *If v is nonincreasing and $\lambda f(-p) > \lambda p + c(n) + \alpha v(n)$ for an $n > 0$, then $V^*(n, t) > v(n)$, and it is optimal to continue ($I^*(n, t) = 1$) for all $t \in [0, T)$.*

By definition, $V^*(n, t) \geqslant \max\{v(n), V^{\psi(n)}(n, t)\}$ for all n and t. As noted before, it is typical in applications for c to be nonincreasing. It is also not unusual for v to not vary much with n, for example the dispatching cost of a vehicle or batch processor does not depend very much on the number of loads. It is shown that if c is nonincreasing and v is constant, then $V^*(n, t) = \max\{v, V^{\psi(n)}(n, t)\}$ for all n and t.

Theorem 5. *If c is nonincreasing and v is constant, then $V^*(n, t) = \max\{v, V^{\psi(n)}(n, t)\}$ for all n and t.*

Proof. By induction on n. For $n = 0$, if $-\lambda p - c(0) > \alpha v$, then $I^*(0, t) = 1$ and $\partial V^*(0, t)/\partial t = \alpha V^*(0, t) + \lambda p + c(0)$ for all $t \in (0, T)$. Then $V^*(0, t) = e^{-\alpha(T-t)}v - (1 - e^{-\alpha(T-t)})(\lambda p + c(0))/\alpha = V^{\psi(0)}(0, t) > v$ for all $t \in [0, T)$. Else, if $-\lambda p - c(0) \leqslant \alpha v$, then $I^*(0, t) = 0$ and $\partial V^*(0, t)/\partial t = \alpha V^*(0, t) - \alpha v$ for all $t \in (0, T)$. Then $V^*(0, t) = v \geqslant V^{\psi(0)}(0, t)$ for all $t \in [0, T]$.

Suppose $V^*(n-1, t) = \max\{v, V^{\psi(n-1)}(n-1, t)\}$ for all $t \in [0, T]$. For $n > 0$, consider the following two cases.

Case 1. $\lambda f(-p) \leqslant \lambda p + c(n) + \alpha v$: Then, from Proposition 4, $V^*(n, t) = v$ for all $t \in [0, T]$, and because $V^{\psi(n)}(n, t) \leqslant V^*(n, t)$, $V^*(n, t) = \max\{v, V^{\psi(n)}(n, t)\}$ for all $t \in [0, T]$.

Case 2. $\lambda f(-p) > \lambda p + c(n) + \alpha v$: Then, from Corollary 1, $V^*(n, t) > v$ and $I^*(n, t) = 1$ for all $t \in [0, T)$. Then $V^*(n, \cdot)$ satisfies

$$\begin{aligned}\frac{\partial V^*(n,\,t)}{\partial t} &= -\lambda \int_{V^*(n,t)-V^*(n-1,t)-p}^{\infty} \\ &\quad \{r - [V^*(n,\,t) - V^*(n-1,\,t) - p]\}\,dF_R(r) \\ &\quad + \alpha V^*(n,\,t) + \lambda p + c(n),\end{aligned}$$

for $t \in (0, T)$ with boundary condition $V^*(n, T) = v(n)$. $V^{\psi(n)}(n, \cdot)$ satisfies

$$\begin{aligned}\frac{\partial V^{\psi(n)}(n,\,t)}{\partial t} &= -\lambda \int_{V^{\psi(n)}(n,t)-\max\{v,V^{\psi(n-1)}(n-1,\,t)\}-p}^{\infty} \\ &\quad \{r-[V^{\psi(n)}(n,\,t)-\max\{v,\,V^{\psi(n-1)}(n-1,\,t)\}-p]\}\,dF_R(r) \\ &\quad + \alpha V^{\psi(n)}(n,\,t) + \lambda p + c(n) \\ &= -\lambda \int_{V^{\psi(n)}(n,t)-V^*(n-1,t)-p}^{\infty} \\ &\quad \{r - [V^{\psi(n)}(n,\,t) - V^*(n-1,\,t) - p]\}\,dF_R(r) \\ &\quad + \alpha V^{\psi(n)}(n,\,t) + \lambda p + c(n),\end{aligned}$$

for $t \in (0, T)$ with boundary condition $V^{\psi(n)}(n, T) = v(n)$. Hence, $V^*(n, \cdot)$ and $V^{\psi(n)}(n, \cdot)$ satisfy the same differential equation with the same boundary condition. Therefore, $V^*(n, t) = V^{\psi(n)}(n, t) \geqslant v$, and $V^*(n, t) = \max\{v, V^{\psi(n)}(n, t)\}$ for all $t \in [0, T]$. □

If the conditions of Theorem 5 hold, then an optimal policy π^* has the following convenient form. If $-\lambda p - c(0) > \alpha v$, then $\lambda f(-p) > \lambda p + c(n) + \alpha v$ for all $n \in \mathcal{N}$, because $\lambda f(-p) \geqslant 0$ and c is nonincreasing. Then $V^*(n, t) = V^{\psi(n)}(n, t)$ and $I^*(n, t) = 1$ for all n and t. Else, if $-\lambda p - c(0) \leqslant \alpha v$, then let $m^* = \max\{0, \max\{n \in \mathcal{N}\backslash[0\} : \lambda f(-p) \leqslant \lambda p + c(n) + \alpha v\}\}$. Then $\lambda f(-p) \leqslant \lambda p + c(n) + \alpha v$ for all $n \leqslant m^*$, because c is nonincreasing. Then

$V^*(n, t) = v$ and $I^*(n, t) = 0$ for all t. Also, $\lambda f(-p) > \lambda p + c(n) + \alpha v$ for all $n > m^*$, and $V^*(n, t) = V^{\psi(n)}(n, t)$ and $I^*(n, t) = 1$ for all t. Hence, as long as $t < T$ and $N^*(t) > m^*$, it is optimal to continue, using threshold $x^*(n, t) = x^{\psi(n)}(n, t) = V^{\psi(n)}(n, t) - \max\{v, V^{\psi(n-1)}(n - 1, t)\} - p = V^*(n, t) - V^*(n - 1, t) - p$. It is optimal to stop and collect v as soon as $N^*(t)$ reaches m^*. This result characterizes an optimal policy and the optimal expected value in a simple, intuitive way, and also leads to an easy method for computing the optimal expected value V^* and optimal threshold x^*.

A number of monotonicity and concavity results for V^* and x^* are derived next.

Theorem 6. *If $\alpha = 0$, c and v are constant, and $\lambda f(-p) \geqslant \lambda p + c$, then the following results hold.*

(*i*) *$\partial V^*(n, t)/\partial t \leqslant \partial V^*(n - 1, t)/\partial t$ for all $n \in \{1, \ldots, N_0\}$ and all $t \in (0, T)$ (the marginal optimal expected value of remaining time, $-\partial V^*(n, t)/\partial t$, is nondecreasing in remaining capacity).*

(*ii*) *$\partial x^*(n, t)/\partial t \leqslant 0$ for all $n \in \{1, \ldots, N_0\}$ and all $t \in (0, T)$ (the optimal threshold is nonincreasing in time).*

(*iii*) *$\partial V^*(n, t_2)/\partial t \leqslant \partial V^*(n, t_1)/\partial t$ for all $n \in \{0, \ldots, N_0\}$ and all $0 < t_1 \leqslant t_2 < T$ ($\partial V^*(n, t)/\partial t$ is nonincreasing in time, or the optimal expected value is concave in time).*

(*iv*) *$x^*(n + 1, t) \leqslant x^*(n, t)$ for all $n \in \{1, \ldots, N_0 - 1\}$ and all $t \in [0, T]$ (the optimal threshold is nonincreasing in remaining capacity).*

(*v*) *$V^*(n + 1, t) - V^*(n, t) \leqslant V^*(n, t) - V^*(n - 1, t)$ for all $n \in \{1, \ldots, N_0 - 1\}$ and all $t \in [0, T]$ (the optimal expected value is concave in remaining capacity).*

Proof. Similar to Corollary 1, because v is constant and $\lambda f(-p) \geqslant \lambda p + c$, it follows that $I^*(n, t) = 1$ and $\partial V^*(n, t)/\partial t = -\lambda f(x^*(n, t)) + \lambda p + c$ for all $n > 0$ and all $t \in (0, T)$. First it is shown that all the results are equivalent, and then it is shown that (i) and (iii) hold.

(i) N (ii):

$$\frac{\partial V^*(n, t)}{\partial t} \leqslant \frac{\partial V^*(n-1, t)}{\partial t}$$

$$\mathrm{N} \quad \frac{\partial x^*(n, t)}{\partial t} = \frac{\partial[V^*(n, t) - V^*(n-1, t) - p]}{\partial t} \leqslant 0.$$

(ii) N (iii): For $n > 0$, $\partial V^*/\partial t$ is nonincreasing in t if and only if x^* is nonincreasing in t. Also, V^* is concave in t if and only if V^* is continuous in t and $\partial V^*/\partial t$ is nonincreasing in t.

(i) N (iv):

$$\frac{\partial V^*(n+1, t)}{\partial t} \leqslant \frac{\partial V^*(n, t)}{\partial t}$$

$$\mathrm{N} -\lambda f(x^*(n+1, t)) + \lambda p + c \leqslant -\lambda f(x^*(n, t)) + \lambda p + c$$

$$\mathrm{N}\, x^*(n+1, t) \leqslant x^*(n, t).$$

(iv) N (v):

$$x^*(n+1, t) \leqslant x^*(n, t)$$

$$\mathrm{N}\;\; V^*(n+1, t) - V^*(n, t) - p \leqslant V^*(n, t) - V^*(n-1, t) - p.$$

For $n > 0$, $\partial V^*(n, t)/\partial t = -\lambda f(V^*(n, t) - V^*(n - 1, t) - p) + \lambda p + c$ for all $t \in (0, T)$. From the continuity of V^* in t, $V^*(n, t)\ 3\ v$ as $t\ 3\ T$. Hence, from the continuity of f, $\partial V^*(n, t)/\partial t\ 3\ -\lambda f(-p) + \lambda p + c$ as $t\ 3\ T$ for all $n > 0$.

It is shown by induction on n that (i) and (iii) hold. For $n = 0$, if $\lambda p + c \leqslant 0$, then $\partial V^*(0, t)/\partial t = \lambda p + c$ for all $t \in (0, T)$. Else, if $\lambda p + c > 0$, then $\partial V^*(0, t)/\partial t = 0$ for all $t \in (0, T)$. Hence, $\partial V^*(0, t)/\partial t = \min\{0, \lambda p + c\}$ for all $t \in (0, T)$. For $n = 1$, it is shown by contradiction that (i) holds. Suppose there exists $t_1 \in (0, T)$ such that $\partial V^*(1, t_1)/\partial t > \partial V^*(0, t_1)/\partial t$. From the continuity of $\partial V^*/\partial t$ in t, there exists a neighborhood $(t_0, t_2) \subseteq (0, T)$ of t_1 such that $\partial V^*(1, t)/\partial t > \partial V^*(0, t)/\partial t$ for all $t \in (t_0, t_2)$. Then for all $t \in (t_1, t_2]$

$$\int_{t_1}^{t} \frac{\partial V^*(1, \tau)}{\partial \tau}\, d\tau > \int_{t_1}^{t} \frac{\partial V^*(0, \tau)}{\partial \tau}\, d\tau$$

$$\mathrm{f}\;\; V^*(1, t) - V^*(1, t_1) > V^*(0, t) - V^*(0, t_1)$$

$$\mathrm{f}\;\; V^*(1, t) - V^*(0, t) > V^*(1, t_1) - V^*(0, t_1)$$

$$\mathrm{f}\;\; -\lambda f(V^*(1, t) - V^*(0, t) - p) + \lambda p + c \geqslant -\lambda f(V^*(1, t_1) - V^*(0, t_1) - p) + \lambda p + c$$

$$\mathrm{f}\;\; \frac{\partial V^*(1, t)}{\partial t} \geqslant \frac{\partial V^*(1, t_1)}{\partial t}.$$

Thus $\partial V^*(1, t)/\partial t$ is nondecreasing on $[t_1, T)$. But $\partial V^*(1, t)/\partial t\ 3\ -\lambda f(-p) + \lambda p + c$ as $t\ 3\ T$, and $f \geqslant 0$ f $-\lambda f(-p) + \lambda p + c \leqslant \lambda p + c$, and $-\lambda f(-p) + \lambda p + c \leqslant 0$ from the assumptions. Hence $\lim_{t\,3\,T} \partial V^*(1, t)/\partial t \leqslant \min\{0, \lambda p + c\} = \partial V^*(0, t)/\partial t$, which contradicts $\partial V^*(1, t_1)/\partial t > \partial V^*(0, t_1)/\partial t$, $\partial V^*(1, t)/\partial t$ nondecreasing on $[t_1, T)$, and $\partial V^*(0, t)/\partial t$ constant on $(0, T)$. Therefore, $\partial V^*(1, t)/\partial t \leqslant \partial V^*(0, t)/\partial t$, $\partial x^*(1, t)/\partial t \leqslant 0$, and $\partial V^*(1, t)/\partial t$ is nonincreasing on $(0, T)$.

For $n > 1$, suppose that $\partial V^*(n - 1, t)/\partial t$ is nonincreasing on $(0, T)$. Similar to the case for $n = 1$, it is shown by contradiction that (i) holds. Suppose there exists $t_1 \in (0, T)$ such that $\partial V^*(n, t_1)/\partial t > \partial V^*(n - 1, t_1)/\partial t$. In the same way as for $n = 1$, it follows that $\partial V^*(n, t)/\partial t$ is nondecreasing on $[t_1, T)$. But $\lim_{t\,3\,T} \partial V^*(n, t)/\partial t = -\lambda f(-p) + \lambda p + c = \lim_{t\,3\,T} \partial V^*(n - 1, t)/\partial t$, which contradicts $\partial V^*(n, t_1)/\partial t > \partial V^*(n - 1, t_1)/\partial t$, $\partial V^*(n, t)/\partial t$ nondecreasing on $[t_1, T)$, and $\partial V^*(n - 1, t)/\partial t$ nonincreasing on $(0, T)$. Therefore, $\partial V^*(n, t)/\partial t \leqslant \partial V^*(n - 1, t)/\partial t$, $\partial x^*(n, t)/\partial t \leqslant 0$, and $\partial V^*(n, t)/\partial t$ is nonincreasing on $(0, T)$. □

5.2. Examples

Closed-form solutions for the optimal expected value $V^*(n, t)$ and the optimal threshold $x^*(n, t)$ can be obtained for some reward distributions F_R. Let $\alpha = 0$, and let the rewards be exponentially distributed with mean $1/\mu$. (Kincaid and Darling 1963, Stadje 1990, and Gallego and Van Ryzin 1994 considered examples of a pricing problem where the maximum price a customer is willing to pay, or the arrival rate of buying customers, is exponentially distributed.) If $\lambda p + c(0) \geqslant 0$, then $V^*(0, t) = v(0)$ for all t; otherwise $V^*(0, t) = v(0) - (\lambda p + c(0))(T - t)$. Suppose $p = 0$, $c(0) \geqslant 0$, v is constant, and $\lambda/\mu > c(1) > 0$. Then $\lambda f(-p) = \lambda/\mu > c(1)$. Hence, from Corollary 1, it is optimal to continue if $n = 1$ for all $t < T$. Then it follows from Equation (12) that

$$\frac{\partial V^*(1, t)}{\partial t} = -\frac{\lambda}{\mu} e^{-\mu[V^*(1,t) - v]} + c(1).$$

The solution is

$$V^*(1, t) = \frac{1}{\mu} \ln\left[e^{\mu v} e^{-\mu c(1)(T - t)} + \frac{\lambda e^{\mu v}}{\mu c(1)} \left[1 - e^{-\mu c(1)(T - t)}\right]\right].$$

Solutions for $n > 1$ were obtained numerically. Computation times were less than a second on a SunSPARC2 workstation. Figure 1(a) shows the optimal expected value $V^*(n, t)$ as a function of time t for different values of the remaining capacity n, for arrival rate $\lambda = 1$, deadline $T = 100$, mean reward $1/\mu = 10$, penalty $p = 0$, waiting cost per unit time $c(n) = 10 - n/10$, terminal value $v = 0$, and discount rate $\alpha = 0$. Figure 1(b) shows the optimal threshold $x^*(n, t)$ versus t for different n.

Figure 2(a) shows the optimal expected value $V^*(n, t)$ as a function of time t for different values of the remaining capacity n, for arrival rate $\lambda = 1$, deadline $T = 100$, exponentially distributed rewards with mean $1/\mu = 10$, penalty $p = 0$, constant waiting cost per unit time $c = 5$, terminal value $v = 0$, and discount rate $\alpha = 0$. Figure 2(b) shows the optimal threshold $x^*(n, t)$ versus t for different n. Note that the optimal expected value is decreasing and concave in time, and the optimal threshold is decreasing in time, as stated in Theorem 6. The shape of the optimal expected value curve is similar to that in Figure 1(a) for a decreasing waiting cost. However, the optimal threshold curves are very different for the different waiting cost structures. This suggests that an optimal policy is quite sensitive with respect to the cost structure.

Figure 3(a) shows the optimal expected value $V^*(n, t)$ as a function of time t for different values of the remaining capacity n, for arrival rate $\lambda = 1$, deadline $T = 100$, uniform reward distribution $u(0, 20)$, penalty $p = 0$, constant waiting cost per unit time $c = 5$, terminal value $v = 0$, and discount rate $\alpha = 0$. Figure 3(b) shows the optimal threshold $x^*(n, t)$ versus time for different n. The curves are similar to those of Figure 2 for the case of exponential rewards. This and other experimentation suggest that an optimal policy is not very sensitive with respect to the reward distribution.

Let $\alpha = 0$, $p = 0$ and $c = 0$. If the rewards are exponentially distributed with mean $1/\mu$, then

$$\frac{\partial V^*(n, t)}{\partial t} = -\frac{\lambda}{\mu} e^{-\mu[V^*(n,t) - V^*(n-1,t)]}.$$

It can be shown by induction on n that

$$V^*(n, t) = \frac{1}{\mu} \ln\left[\sum_{i=0}^{n} \frac{\lambda^i (T - t)^i}{i!} e^{\mu v(n-i)} \right].$$

If v is constant, then

$$V^*(n, t) = \frac{1}{\mu} \ln\left[\sum_{i=0}^{n} \frac{\lambda^i (T - t)^i}{i!} \right] + v.$$

It is interesting to note that, due to the continuity of the ln function,

$$\lim_{n \to \infty} V^*(n, t) = \frac{1}{\mu} \ln\left[\lim_{n \to \infty} \sum_{i=0}^{n} \frac{\lambda^i (T - t)^i}{i!} \right] + v$$
$$= \frac{\lambda}{\mu}(T - t) + v$$

$$\lim_{n \to \infty} x^*(n, t) = 0.$$

This result is intuitive, because if the remaining capacity is very large, it is optimal to accept all arrivals, and the optimal threshold $x^*(n, t) = 0$ for all t. From Wald's equation the expected value is the expected number of arrivals in the remaining time, $\lambda(T - t)$, times the expected reward per arrival, $1/\mu$, plus the terminal value v.

6. CONCLUDING REMARKS

The Dynamic and Stochastic Knapsack Problem (DSKP) was defined and analyzed. For the infinite horizon case it was shown that a stationary deterministic threshold policy is optimal among all history-dependent deterministic policies. For the finite horizon case it was shown that a memoryless deterministic threshold policy is optimal among all history-dependent deterministic policies. General characteristics of the optimal policies and optimal expected values were derived for different cases. Optimal solutions can be computed recursively with very little computational effort. Closed-form solutions were obtained for special cases.

An interesting extension to the DSKP with equal sized items is the case where items have random sizes. This problem is the topic of a separate paper (Kleywegt and Papastavrou 1995), in which some counter-intuitive properties of optimal policies are pointed out. Another useful extension to the DSKP considers the case where items as well as knapsacks arrive according to some stochastic process, and the objective is to find an optimal acceptance policy for items, and an optimal dispatching policy for knapsacks.

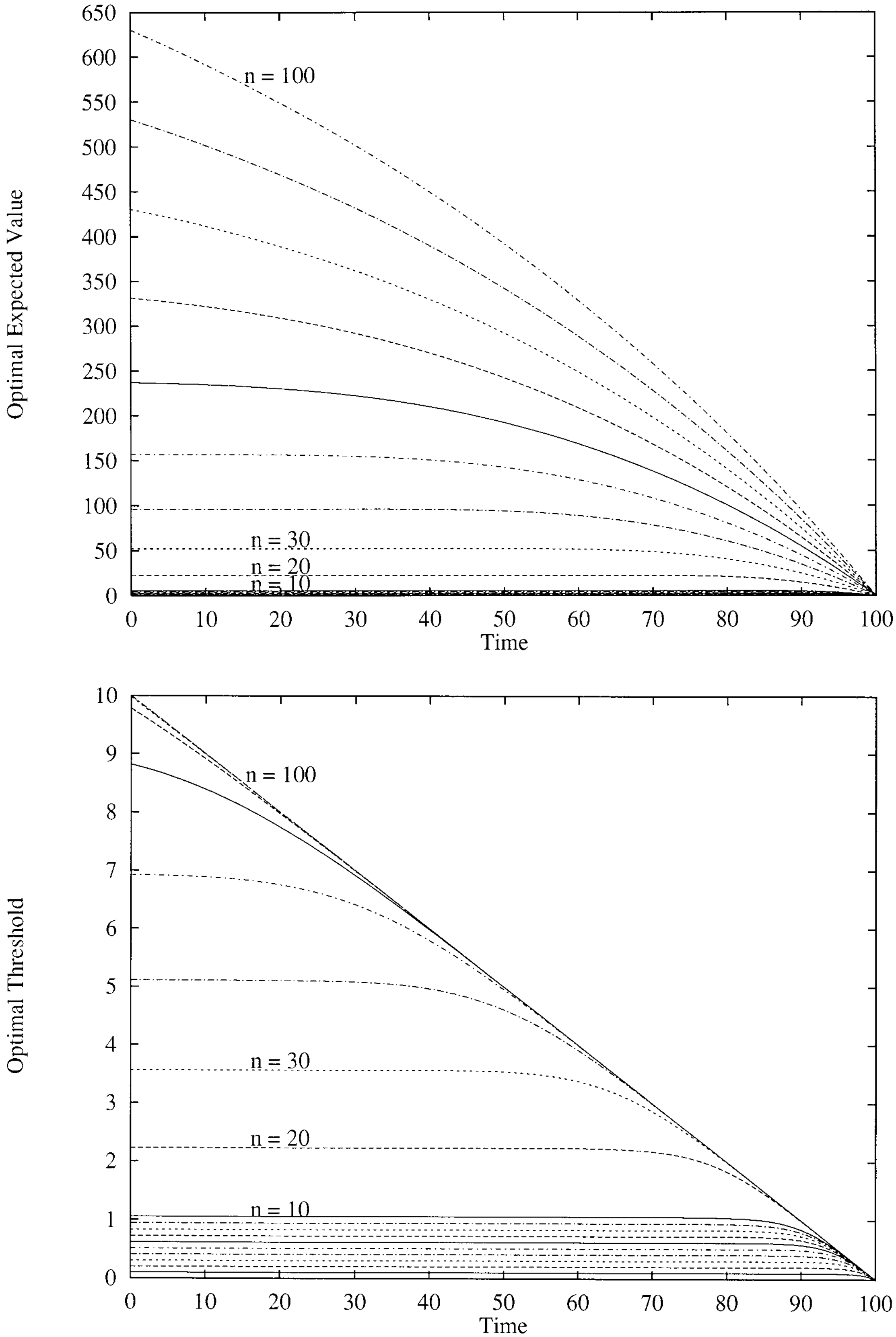

Figure 1. Poisson arrivals with rate $\lambda = 1$, deadline $T = 100$, exponential rewards with mean $1/\mu = 10$, penalty $p = 0$, variable waiting cost per unit time $c(n) = 10 - n/10$, terminal value $v = 0$, discount rate $\alpha = 0$. (a) Optimal expected value versus time for different remaining capacities n. (b) Optimal threshold versus time for different remaining capacities n.

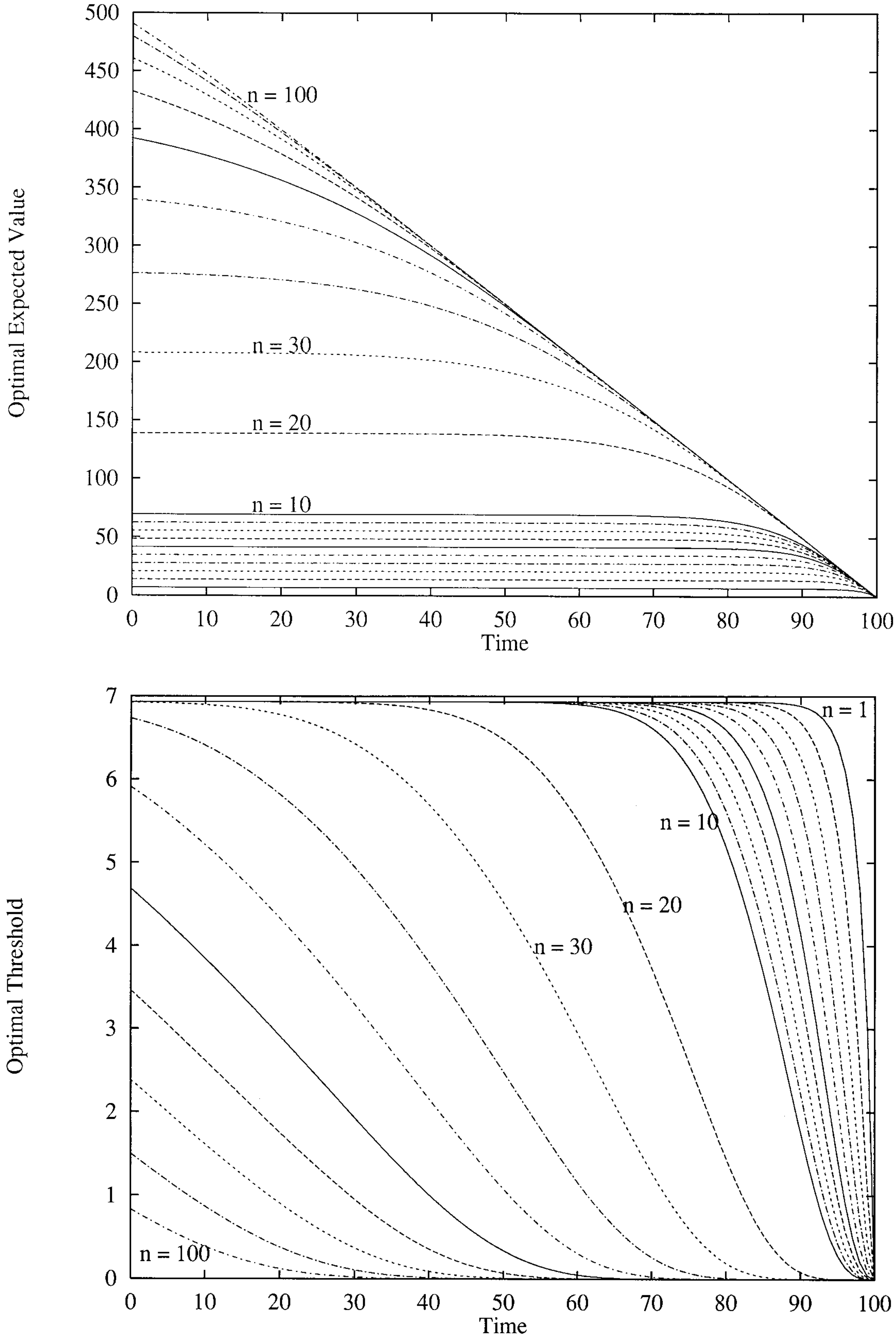

Figure 2. Poisson arrivals with rate $\lambda = 1$, deadline $T = 100$, exponential rewards with mean $1/\mu = 10$, penalty $p = 0$, constant waiting cost per unit time $c = 5$, terminal value $v = 0$, discount rate $\alpha = 0$. (a) Optimal expected value versus time for different remaining capacities n. (b) Optimal threshold versus time for different remaining capacities n.

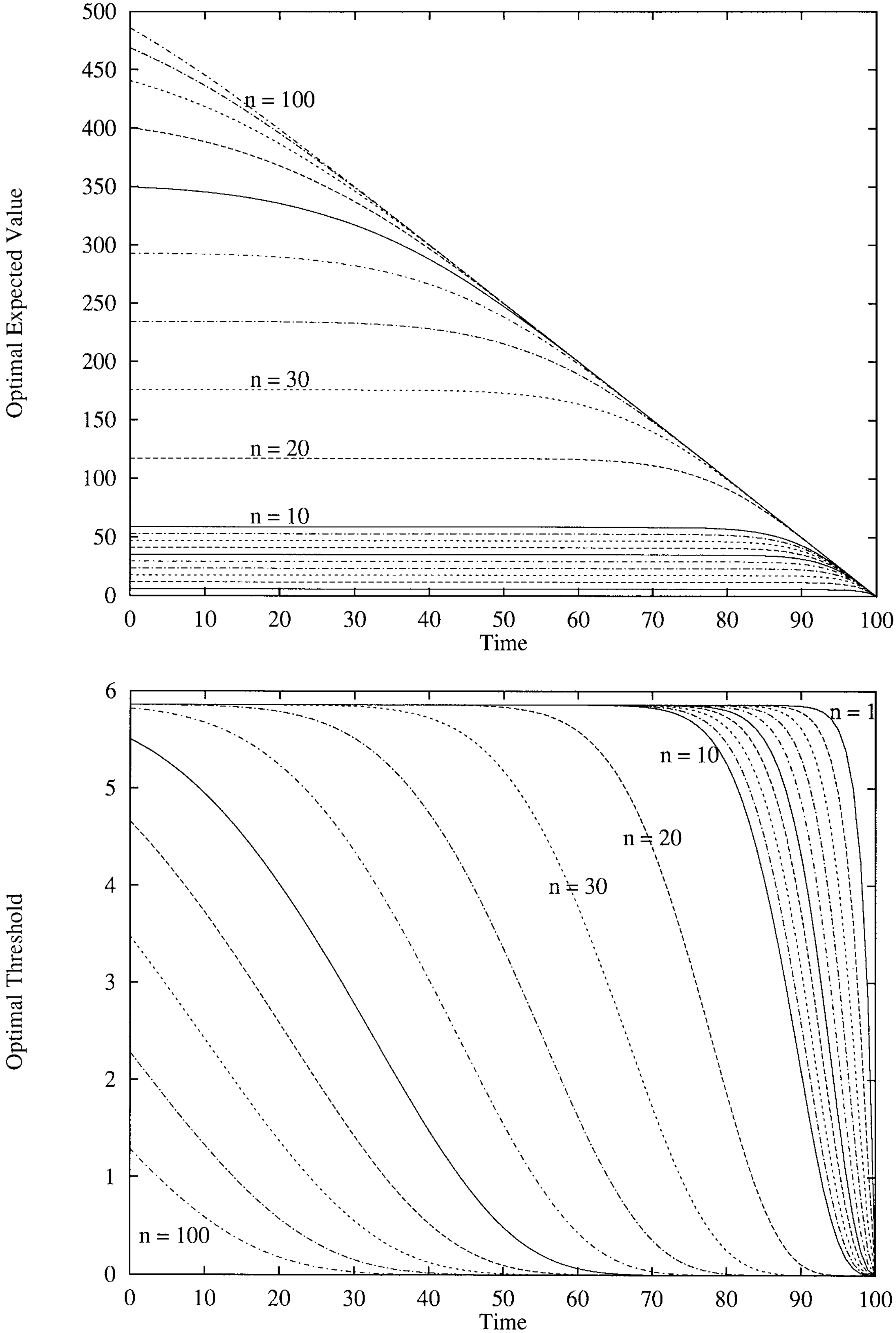

Figure 3. Poisson arrivals with rate $\lambda = 1$, deadline $T = 100$, uniform $u(0, 20)$ rewards, penalty $p = 0$, constant waiting cost per unit time $c = 5$, terminal value $v = 0$, discount rate $\alpha = 0$. (a) Optimal expected value versus time for different remaining capacities n. (b) Optimal threshold versus time for different remaining capacities n.

ACKNOWLEDGMENT

We thank Colm O'Cinneide, Tom Sellke, and the anonymous referees for helpful suggestions. This research was supported by the National Science Foundation under grant DDM-9309579.

APPENDIX

The appendix can be found at the *Operations Research* Home Page: http://opim.wharton.upenn.edu/~harker/opsresearch.html in the Online Collection.

REFERENCES

ALBRIGHT, C. AND C. DERMAN. 1972. Asymptotic Optimal Policies for the Stochastic Sequential Assignment Problem. *Mgmt. Sci.* **19,** 46–51.

ALBRIGHT, S. C. 1974. Optimal Sequential Assignments with Random Arrival Times. *Mgmt. Sci.* **21,** 60–67.

ALBRIGHT, S. C. 1977. A Bayesian Approach to a General House Selling Problem. *Mgmt. Sci.* **24,** 432–440.

ALSTRUP, J., S. BOAS, O. B. G. MADSEN, AND R. V. V. VIDAL. 1986. Booking Policy for Flights with Two Types of Passengers. *Eur. J. Opnl. Res.* **27,** 274–288.

BELOBABA, P. P. 1987. Airline Yield Management. An Overview of Seat Inventory Control. *Trans. Sci.* **21,** 63–73.

BELOBABA, P. P. 1989. Application of a Probabilistic Decision Model to Airline Seat Inventory Control. *Opns. Res.* **37,** 183–197.

BRÉMAUD, P. 1981. *Point Processes and Queues, Martingale Dynamics*. Springer-Verlag, New York, NY.

BRUMELLE, S. L. AND J. I. MCGILL. 1993. Airline Seat Allocation with Multiple Nested Fare Classes. *Opns. Res.* **41,** 127–137.

BRUMELLE, S. L., J. I. MCGILL, T. H. OUM, K. SAWAKI, AND M. W. TRETHEWAY. 1990. Allocation of Airline Seats Between Stochastically Dependent Demands. *Transp. Sci.* **24,** 183–192.

BRUSS, F. T. 1984. A Unified Approach to a Class of Best Choice Problems with an Unknown Number of Options. *Annals of Probability*, **12,** 882–889.

CARRAWAY, R. L., R. L. SCHMIDT, AND L. R. WEATHERFORD. 1993. An Algorithm for Maximizing Target Achievement in the Stochastic Knapsack Problem with Normal Returns. *Naval Res. Logist.* **40,** 161–173.

CURRY, R. E. 1990. Optimum Airline Seat Allocation with Fare Classes Nested by Origins and Destinations. *Trans. Sci.* **24,** 193–204.

DERMAN, C., G. J. LIEBERMAN, AND S. M. ROSS. 1972. A Sequential Stochastic Assignment Problem. *Mgmt. Sci.* **18,** 349–355.

DROR, M., P. TRUDEAU, AND S. P. LADANY. 1988. Network Models for Seat Allocation of Flights. *Trans. Res.* **22B,** 239–250.

FREEMAN, P. R. 1983. The Secretary Problem and Its Extensions: A Review. *International Statistical Review*, **51,** 189–206.

GALLEGO, G. AND G. VAN RYZIN. 1994. Optimal Dynamic Pricing of Inventories with Stochastic Demand over Finite Horizons. *Mgmt. Sci.* **40,** 999–1020.

HENIG, M. I. 1990. Risk Criteria in a Stochastic Knapsack Problem. *Opns. Res.* **38,** 820–825.

KAUFMAN, J. F. 1981. Blocking in a Shared Resource Environment. *IEEE Trans. Communications*, **29,** 1474–1481.

KENNEDY, D. P. 1986. Optimal Sequential Assignment. *Math. O. R.* **11,** 619–626.

KINCAID, W. M. AND D. A. DARLING. 1963. An Inventory Pricing Problem. *J. Math. Analysis and Appl.* **7,** 183–208.

KLEYWEGT, A. J. AND J. D. PAPASTAVROU. 1995. The Dynamic and Stochastic Knapsack Problem. Technical Report 95-17, School of Industrial Engineering, Purdue University, West Lafayette, IN.

LEE, T. C. AND M. HERSH. 1993. A Model for Dynamic Airline Seat Inventory Control with Multiple Seat Bookings. *Transp. Sci.* **27,** 252–265.

MAMER, J. W. 1986. Successive Approximations for Finite Horizon, Semi-Markov Decision Processes with Application to Asset Liquidation. *Opns. Res.* **34,** 638–644.

MARTELLO, S. AND P. TOTH. 1990. *Knapsack Problems. Algorithms and Computer Implementations*. John Wiley & Sons, West Sussex, England.

MENDELSON, H., J. S. PLISKIN, AND U. YECHIALI. 1980. A Stochastic Allocation Problem. *Opns. Res.* **28,** 687–693.

NAKAI, T. 1986a. An Optimal Selection Problem for a Sequence with a Random Number of Applicants per Period. *Opns. Res.* **34,** 478–485.

NAKAI, T. 1986b. A Sequential Stochastic Assignment Problem in a Partially Observable Markov Chain. *Math. O. R.* **11,** 230–240.

NAKAI, T. 1986c. A Sequential Stochastic Assignment Problem in a Stationary Markov Chain. *Math. Japonica*, **31,** 741–757.

PAPASTAVROU, J. D., S. RAJAGOPALAN, AND A. J. KLEYWEGT. 1995. A Stochastic Model for the Knapsack Problem with a Deadline. Technical Report, School of Industrial Engineering, Purdue University, West Lafayette, IN.

PLISKA, S. R. 1975. Controlled Jump Processes. *Stochastic Processes and Their Applications*, **3,** 259–282.

PRASTACOS, G. P. 1983. Optimal Sequential Investment Decisions under Conditions of Uncertainty. *Mgmt. Sci.* **29,** 118–134.

PRESMAN, E. L. AND I. M. SONIN. 1972. The Best Choice Problem for a Random Number of Objects. *Theory of Probability and Its Applications*, **17,** 657–668.

RIGHTER, R. 1989. A Resource Allocation Problem in a Random Environment. *Opns. Res.* **37,** 329–338.

ROBINSON, L. W. 1995. Optimal and approximate control policies for airline booking with sequential nonmonotonic fare classes. *Opns. Res.* **43,** 252–263.

ROSENFIELD, D. B., R. D. SHAPIRO, AND D. A. BUTLER. 1983. Optimal Strategies for Selling an Asset. *Mgmt. Sci.* **29,** 1051–1061.

ROSS, K. W. AND D. H. K. TSANG. 1980. The Stochastic Knapsack Problem. *IEEE Trans. Communications*, **37,** 740–747.

ROSS, K. W. AND D. D. YAO. 1990. Monotonicity Properties for the Stochastic Knapsack. *IEEE Trans. Information Theory*, **36,** 1173–1179.

ROTHSTEIN, M. 1971. An Airline Overbooking Model. *Transp. Sci.* **5,** 180–192.

ROTHSTEIN, M. 1974. Hotel Overbooking as a Markovian Sequential Decision Process. *Decision Sci.* **5,** 389–404.

Rothstein, M. 1985. OR and the Airline Overbooking Problem. *Opns. Res.* **33,** 237–248.

Saario, V. 1985. Limiting Properties of the Discounted House-selling Problem. *European J. Opnl. Res.* **20,** 206–210.

Sakaguchi, M. 1986. Best Choice Problems for Randomly Arriving Offers During a Random Lifetime. *Math. Japonica,* **31,** 107–117.

Sakaguchi, M. 1984a. A Sequential Stochastic Assignment Problem Associated with a Non-homogeneous Markov Process. *Math. Japonica,* **29,** 13–22.

Sakaguchi, M. 1984b. A Sequential Stochastic Assignment Problem with an Unknown Number of Jobs. *Math. Japonica,* **29,** 141–152.

Shlifer, E. and Y. Vardi. 1975. An Airline Overbooking Policy. *Transp. Sci.* **9,** 101–114.

Sniedovich, M. 1980. Preference Order Stochastic Knapsack Problems: Methodological Issues. *J. Opnl. Res. Soc.* **31,** 1025–1032.

Sniedovich, M. 1981. Some Comments on Preference Order Dynamic Programming Models. *J. Math. Analysis and Appl.* **79,** 489–501.

Stadje, W. 1990. A Full Information Pricing Problem for the Sale of Several Identical Commodities. *Zeitschrift für Opns. Res.* **34,** 161–181.

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

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

Tamaki, M. 1986a. A Full-information Best-choice Problem with Finite Memory. *J. Appl. Probability,* **23,** 718–735.

Tamaki, M. 1986b. A Generalized Problem of Optimal Selection and Assignment. *Opns. Res.* **34,** 486–493.

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

Weatherford, L. R., S. E. Bodily, and P. E. Pfeifer. 1993. Modeling the Customer Arrival Process and Comparing Decision Rules in Perishable Asset Revenue Management Situations. *Transp. Sci.* **27,** 239–251.

Wollmer, R. D. 1992. An Airline Seat Management Model for a Single Leg Route When Lower Fare Classes Book First. *Opns. Res.* **40,** 26–37.

Yasuda, M. 1984. Asymptotic Results for the Best-choice Problem with a Random Number of Objects. *J. Appl. Probability,* **21,** 521–536.

Yushkevic, A. A. and E. A. Fainberg. 1979. On Homogeneous Markov Models with Continuous Time and Finite or Countable State Space. *Theory of Probability and Its Applications,* **24,** 156–161.

Zenz, G. J. 1987. *Purchasing and the Management of Materials.* John Wiley & Sons, New York, NY.