AN (S − 1, S) INVENTORY SYSTEM WITH FIXED SHELF LIFE AND CONSTANT LEAD TIMES

DAVID PERRY
University of Haifa, Haifa, Israel

M. J. M. POSNER
University of Toronto, Toronto, Ontario, Canada

(Received July 1993; revisions received July 1994, December 1995; accepted August 1996)

We consider an (S − 1, S) type perishable inventory system in which the maximum shelf life of each item is fixed. An order for an item is placed at each demand time as well as at each time that the maximum shelf life of an item is reached. The order lead times are constant, and the demand process for items is Poisson. Although the resulting process is ostensibly nonregenerative, we adapt level-crossing theory for the case of an S-dimensional Markov process to obtain its stationary law. Within this framework a number of model variants are solved.

In some recent papers, Kaspi and Perry (1984) and Perry and Posner (1989, 1990), examine an inventory system of perishable items in which the arrival process of items and the demand process for those items are random, and mutually independent. In particular, when both processes are Poisson, explicit results can be obtained that have utility for optimization purposes. In this paper we will study a somewhat different perishable inventory system in which the arrival process is actually determined by the demand process. Specifically, we consider a version of a perishable inventory system of the so-called (S − 1, S) type, in which an order for exactly one item is placed at each time that a demand is satisfied as well as at each occurrence of an outdating of an item, i.e., when its fixed shelf life has been reached. The resulting order replenishment lead time is fixed at τ, and the shelf life of an item is a constant m. The customer demand process is Poisson.

We distinguish between the two phrases, "number of items in the system" and "number of available items on the shelf" in the sense that there are always S items in the system of which only K_t (= 1, . . . , S) are available on the shelf at time t. Since the lead time is constant, every customer that arrives to find an empty shelf can "see" exactly how much time, if any, he must wait for the arrival of the next available item. We assume that such customers are willing to wait a random time Y, where the Y_is are i.i.d. random variables having distribution H. We develop the model for general H and present its solution in terms of a Volterra integral. We then consider four special examples of H for which explicit solutions are obtained.

The first waiting time policy is of the form:

$$H(y) = q + p1_{\{y \geqslant \tau\}}, \quad y \geqslant 0, \tag{0.1}$$

where $0 \leqslant p \leqslant 1$, $q = 1 - p$, and $1_{\{A\}}$ designate the indicator function.

Under this waiting time policy, an arbitrary demand that encounters an empty shelf is willing to wait for the next item with probability p, and will leave unsatisfied with probability q. Three other waiting time policy models that will also be introduced to illustrate the broad scope of applicability of the model to be developed will include H of the forms exponential, Rayleigh, and uniform distributions.

Schmidt and Nahmias (1985) have already considered a special case of waiting time policy (0.1). Namely, their model is the same (S − 1, S) as ours but is valid only under the restriction $p = 0$. In practice, however, their restrictive model is not really natural since the lead time is constant and every arriving customer demand can find out the detailed history of the process. In other words, customers that arrive to find an empty shelf can know exactly how much time they must wait for the next available item. But, in the special case $p = 0$ they are required to leave w.p.1 even though a replacement item arrival may be imminent. On the other hand, the waiting time policy (0.1) with $0 < p \leqslant 1$ is more general since it takes the option of waiting into account. Consequently, some customers will choose to wait, while others will leave unsatisfied. Furthermore, as will be seen below, our approach is more elegant and sophisticated, and lends itself readily to a host of productive generalizations.

The class of so-called (S − 1, S) inventory models represents the basis for much of the multiechelon models such as METRIC (see Sherbrooke 1968) that have application to control replenishment for high value spare parts, particularly in the military. However, most models operating with (S − 1, S) policies that have appeared in the literature assume unlimited shelf life (Smith 1977, Moinzadeh and Schmidt 1991, and Graves 1982). The only related

Subject classifications: Inventory, stochastic: perishable items. Lead items: probabalistic propensity for customer waiting.
Area of review: MANUFACTURING OPERATIONS.

0030-364X/98/4603-S065 $05.00

work that discusses the $(S - 1, S)$ model under the assumption of limited shelf life is that of Schmidt and Nahmias (1985).

The approach in this study is based on the concepts used in Perry and Posner (1990), namely, on an analogy between our model and some queueing model with impatient customers. (See also Perry and Asmussen 1995.) Graves (1982) was probably the first to consider a continuous-review inventory model as an application of the theory of queues with impatient customers. However, our model differs considerably from his.

The paper is organized as follows: In Section 1 we introduce the general framework and dynamics of the model. In Section 2 we formulate the methodology, and develop the general solution in Section 3 based on the steady state law of a non-Markov process in terms of a Volterra integral. In Section 4 an explicit solution is obtained for the waiting time policy (0.1), while in Section 5, the distribution of the number of items on the shelf as well as other useful system measures are obtained. Finally, in Section 6 we demonstrate the versatility of the model by generating solutions for models incorporating a variety of waiting time policies.

1. THE MODEL DYNAMICS

The model is characterized by the condition that a replenishment order is placed each time an item is removed from the shelf; namely, at a moment either of a satisfied demand or of an outdating. The requirement that both the shelf life and the leadtime are constants guarantees that the items arrive onto the shelf according to a first-ordered-first-enter discipline.

A natural criterion for controlling this system should be linked with the three processes: (i) the unsatisfied demand process, (ii) the outdating process, and (iii) $K = \{K_t : t \geq 0\}$, the number of items on the shelf. Clearly, the point processes (i) and (ii) are not renewal processes, and (iii) is not even a Markov process. Thus, in order to analyze the $(S - 1, S)$ inventory system we must define a state space for the model that characterizes a Markov process. Now, the number of items in the system is always S, and recall that we would know the remaining times to outdating of each of them if the demand process were stopped. An appropriate state space is then the set of possible values of:

$$W = \{W_1(t), W_2(t), \ldots, W_s(t) : t \geq 0\}, \tag{1.1}$$

where $W_i(t)$ $(i = 1, \ldots, S)$ is the time to outdating of the ith *youngest* item if the demand process were stopped at t. Thus, $m + \tau \geq W_1(t) > \ldots > W_s(t) \geq 0$.

Note that we emphasize the word "youngest," because, if $W_i(t) > m$, then the ith youngest item is still unavailable on the shelf. Also note that the "age" $m + \tau - W_i(t)$ is, in fact, the time elapsed since the ith youngest item was *ordered*, and not from the time it arrived on the shelf.

We specify the S-dimensional Markov process W as the collection of the Virtual Outdating Times (VOTs) process, and the one-dimensional component W_s as the minimal VOT (MVOT). Our analysis is based on the MVOT, even though it is not a Markov process.

To characterize the dynamics governing the evolution of W we first introduce the following processes.

Let Z_n, $n = 1, 2, \ldots$, be the time of the nth demand arrival, and let Y_n be the corresponding amount of time it is willing to wait. Then:

$$W_1(Z_n^+) = \begin{cases} m + \tau, & 0 < W_s(Z_n^-) < m, \\ (m + \tau)1_{\{Y_n \geq W_s(Z_n^-) - m\}} \\ \quad + W_1(Z_n^-)1_{\{Y_n < W_s(Z_n^-) - m\}}, & m \leq W_s(Z_n^-) < m + \tau, \end{cases} \tag{1.2}$$

and for $i = 1, 2, \ldots, S - 1$:

$$W_{i+1}(Z_n^+) = \begin{cases} W_i(Z_n^-), & 0 < W_s(Z_n^-) < m, \\ W_i(Z_n^-)1_{\{Y_n \geq W_s(Z_n^-) - m\}} \\ \quad + W_{i+1}(Z_n^-)1_{\{Y_n < W_s(Z_n^-) - m\}}, & m \leq W_s(Z_n^-) < m + \tau. \end{cases} \tag{1.3}$$

Equation (1.2) says simply that immediately after a demand arrival, two events may occur:

(i) If, just before that arrival, the youngest item is available on the shelf, then the demand is satisfied and a new order is placed. Thus, the "age" of the *new* youngest item is now 0, and consequently, its remaining potential lifetime is $m + \tau$.

(ii) If, just before that arrival the youngest item is still unavailable on the shelf, then the demand will be either satisfied (if it is willing to wait), or unsatisfied (if it is not willing to wait). If the demand is willing to wait, case (i) now applies, and if it is not willing to wait, no order has been placed and so $W_1(Z_n^+) = W_1(Z_n^-)$.

The idea behind (1.3) is similar to that of (1.2). Figure 1 describes the dynamics of the model in case of a satisfied demand (1(a)) and in case of an unsatisfied demand (1(b)) for $S = 3$.

Next, let T_k, $k = 1, 2, \ldots$, be the time of the kth outdating. Then,

$$T_{k+1} = \inf\{t > T_k : W_s(t) = 0\}, \tag{1.4}$$

and as a result, we have (i) $W_s(T_k^-) = 0$, (ii) $\{W_1(T_k^+) = m + \tau\}$, and (iii) $\{W_{i+1}(T_k^+) = W_i(T_k^-)\}$ for all $i = 1, \ldots, S - 1$, and $k = 1, 2, \ldots$.

A demand can be either satisfied or unsatisfied. The composition of the satisfied demand process with the outdating process generates a jump process associated with each component W_i $(i = 1, \ldots, S)$ of W. Let J_n be the time of the nth jump of W_i. Then, for an arbitrary time $0 < t < J_{n+1} - J_n$:

$$W_i(J_n + t) = W_i(J_n^+) - t, \quad i = 1, \ldots, S.$$

A typical realization of the model is depicted in Figure 2

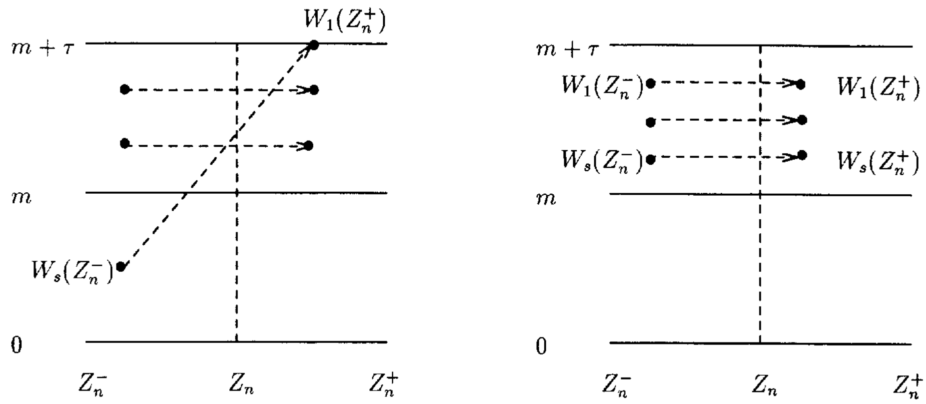

Figure 1. (a) A satisfied demand. (b) An unsatisfied demand.

for the case $S = 3$. In that realization, the shelf is empty of items in the interval $[A_3, E_1)$; B_1 is a time of a satisfied demand since the customer was willing to wait, but C_1 is a time of an unsatisfied demand since the customer was not willing to wait. As time progresses, the VOTs decrease with slope 1 between jumps, and so forth.

2. MODEL FORMULATION

As was already mentioned, $K = \{K_t : t \geq 0\}$, defined as the number of items on the shelf, is not a Markov process. However, the following relationship between K and W holds:

$$\{K_t = k\} = \{W_{S-k+1}(t) < m \leq W_{S-k}(t)\}, \quad k = 0, 1, \ldots S, \tag{2.1}$$

in which, by definition, $W_{S+1}(t) \equiv 0$ and $W_0(t) \equiv m + \tau$. It follows from (2.1) that $W = \{W_1(t), \ldots, W_s(t) : t \geq 0\}$ has a stationary law if and only if K has. It is evident that K has a limiting law since, by definition, $0 \leq K_t \leq S$. Let K_∞, W_{S-k+1}, and W_{S-k} be random variables such that:

$$\begin{aligned} P(K_\infty = k) &= \lim_{t\to\infty} P(K_t = k) \\ &= \lim_{t\to\infty} P(W_{S-k+1}(t) < m \leq W_{S-k}(t)) \\ &= P(W_{S-k+1} < m \leq W_{S-k}). \end{aligned} \tag{2.2}$$

Then also, for the moments, $EK_\infty^n = \lim_{t\to\infty} EK_t^n$, for all $n = 1, 2, \ldots$, by dominated convergence.

According to (2.2) the stationary law of K can be analyzed through that of the appropriate components of W.

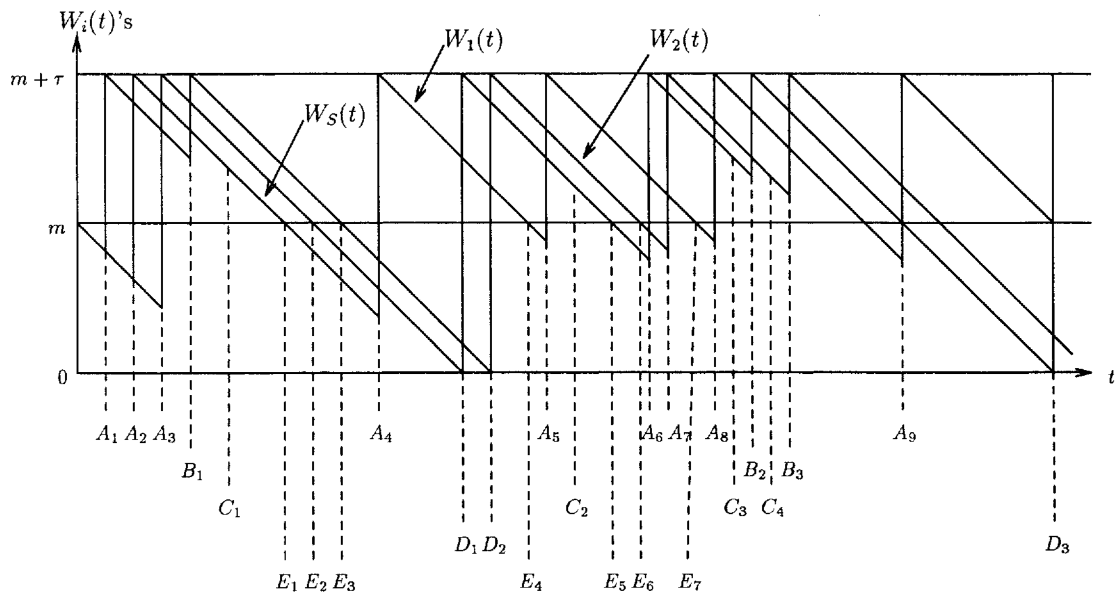

Figure 2. The VOT process with $S = 3$.
Legend: $A_1, A_2, \ldots$ satisfied demands without waiting.
$B_1, B_2, \ldots$ satisfied demands that are willing to wait.
$C_1, C_2, \ldots$ unsatisfied demands. (Arrival points of customers t hat are *not* willing to wait.)
$D_1, D_2, \ldots$ points of outdating.
$E_1, E_2, \ldots$ arrivals of fresh available items onto the shelf.

Indeed, every subset of W is not a Markov process. Thus, in order to utilize the Markov property, we should use the entire S-dimensional process $W = (W_1, \ldots, W_s)$, and then compute the marginal laws of its individual components.

Our developmental method is based on level-crossing theory (Perry and Posner 1989, 1990). We arbitrarily fix $x > 0$ and partition the state space of the stationary process $W = (W_1, \ldots, W_s)$ into two disjoint subsets, I_x and I_x^c. Then we equate the steady state transition rates from one set to the other.

Let $I_x^c = \{W_1 > W_2 > \ldots > W_s > x\}$. Since, by definition, $W_1 > W_2 > \ldots > W_s$, it follows that $I_x = \{W_s \leqslant x\}$. In steady state, we have:

$$\{\text{rate of } I_x \to I_x^c \text{ transitions}\}$$
$$= \{\text{rate of } I_x^c \to I_x \text{ transitions}\}. \tag{2.3}$$

In our model, transitions $I_x^c \to I_x$ occur only when the MVOT W_s hits level x. By level-crossing theory (e.g., Perry and Posner 1985), the $I_x^c \to I_x$ transition rate is the marginal density of W_s, and the right-hand side of (2.3) is:

$$\int_{\omega_{s-1}=x}^{m+0} \cdots \int_{\omega_2=\omega_3}^{m+\tau} \cdots \int_{\omega_1=\omega_2}^{m+\tau} \cdot f(\omega_1, \omega_2, \ldots, \omega_{s-1}, x)\, d\omega_1\, d\omega_2 \ldots d\omega_{s-1} \equiv f_{W_s}(x), \tag{2.4}$$

where $f(\cdot, \cdot, \ldots, \cdot)$ is the S-dimensional density of $(W_1, \ldots, W_s)$, and $f_{W_s}(\cdot)$ is the marginal density of W_s.

To better understand the assertion leading to (2.4), recall that an $I_x^c \to I_x$ transition occurs each time the sample path of W_s downcrosses level x. The total amount of time W_s stays within $[x, x + h)$ during the time interval $[t_1, t_2)$ (or $[0, t)$, for $t = t_2 - t_1$, since W_s is stationary) is:

$$\int_0^t 1_{\{x \leqslant W_s(u) < x+h\}}\, du. \tag{2.5}$$

To evaluate (2.5) let $D_x(t)$ be the number of downcrossings of level $x > 0$ during $[0, t)$. Now, W_s decreases with slope 1. Thus, for small h, each downcrossing of level $x + h$ is also a downcrossing of level x, and vice versa. Therefore, (2.5) can be evaluated approximately by $D_x(t) \cdot h$. As $h \downarrow 0$, we get w.p.1:

$$\frac{D_x(t)}{t} = \frac{d}{dx} \frac{1}{t} \int_0^t 1_{\{W_s(u) \leqslant x\}}\, du. \tag{2.6}$$

Letting $t \to \infty$ in (2.6), we obtain from the strong law of large numbers that the left-hand side of (2.6) is the average number of downcrossings of level x by W_s per unit time. Similarly, the right-hand side of (2.6) is the derivative of the limiting proportion of time W_s stays below level x; this is precisely the stationary density of W_s.

Remark. The interchange of limits $h \to 0$ and $t \to \infty$ in obtaining (2.6) should be generally justified more rigorously. In fact, this is the basis for the level-crossing theory that we have used in some of our previous work (Perry and Posner 1989, 1990).

To compute the left-hand side of (2.3), note that transitions $I_x \to I_x^c$ occur only at jump epochs, namely, either at a satisfied demand time or at an outdating time. However, not every jump is a transition. In order that a jump be an $I_x \to I_x^c$ transition it is necessary that the state just before the jump be $\{W_s \leqslant x\} \cap \{W_{s-1} > x\} = \{W_s < x \leqslant W_{s-1}\} = \{W_{s-1} - W_s \geqslant x - W_s > 0\}$. In words, the difference between the ages of the second oldest and oldest items is greater than the level x minus the MVOT.

Due to the Markov property of the VOT process, the probability that a jump occurs in the small time interval $[t, t + h)$ is comprised of the probability of the union of two (almost) disjoint events; a demand arrival with probability $\lambda h + o(h)$, and an outdating with probability $f_{W_s}(0) \cdot h + o(h)$, where the probability of the intersection is $o(h)$.

We distinguish between two cases:

(i) $x < m$. In this case, the $I_x \to I_x^c$ transition rate is:

$$\lambda P(0 < W_s < x < W_{s-1}) + f_{W_s}(0) P(W_{s-1} > x), \tag{2.7}$$

(ii) $x > m$. In this case, the $I_x \to I_x^c$ transition rate is:

$$\lambda P(0 < W_s < m;\ W_{s-1} > x) + \lambda P(m < W_s < x \leqslant W_{s-1}) P(Y \geqslant W_s - m) + f_{W_s}(0) P(W_{s-1} \geqslant x). \tag{2.8}$$

Case (ii) differs from case (i) in that the waiting time is optional. That is, a customer may arrive at the system and find an empty shelf. Then, he will decide to wait with conditional probability $1 - H(W_s - m)$ given W_s. This factor enters into the second expression of (2.8). In fact, such a customer is not satisfied immediately, but only at the end of the leadtime. However, this does not affect the VOT process since the shelf is already empty.

Implementing now the concept (2.3) for the Markov process W, we equate (2.4) with (2.7) and (2.8) for cases (i) and (ii). Then, by conditioning on W_s we finally obtain:

$$f(x) = \begin{cases} f(0) P(W_{s-1} \geqslant x) \\ \quad + \lambda \displaystyle\int_0^x P(W_{s-1} \geqslant x - \omega) f(\omega)\, d\omega, & x < m, \\ f(0) P(W_{s-1} \geqslant x) \\ \quad + \lambda \displaystyle\int_0^x [1 - H(\omega - m)] \\ \quad \cdot P(W_{s-1} \geqslant x - \omega) f(\omega)\, d\omega, & x \geqslant m, \end{cases} \tag{2.9}$$

where f is used for the marginal density, f_{W_s}, of the MVOT process in steady state.

3. THE MODEL SOLUTION

The integral Equation (2.9) is still unsolvable since the probability of the event $\{W_{s-1} \geqslant x\}$ is as yet unknown. The next lemma will provide the necessary tool to compute this probability. We assume that the process is sufficiently

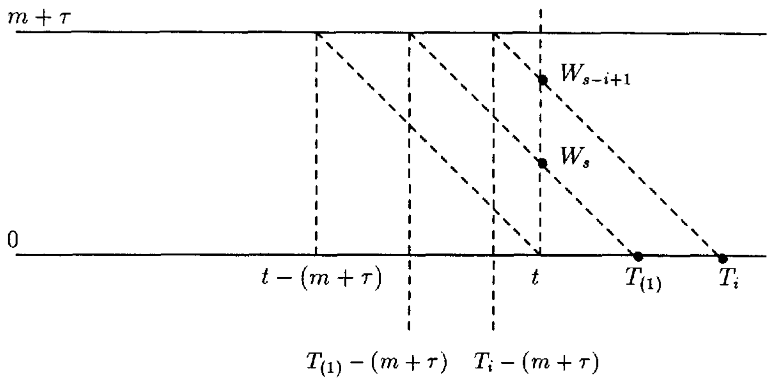

Figure 3. A realization of ordered outdating times.

old so that all items have distinct ages. Then, for arbitrary time t, we assume that the demand process is stopped.

Lemma 1. *Let $T_{(i)}$ be the ith ordered outdating time after t, for $i = 1, \ldots, S$. Then, $T_{(2)}, \ldots, T_{(s)}$ have the same common law as the $S-1$ order statistics taken from a uniform distribution on $[T_{(1)}, t+m+\tau)$.*

Proof. The demand process is stopped at t. Thus, exactly S outdatings are anticipated in $[t, t+m+\tau)$, and consequently $S-1$ outdatings are anticipated in $[T_{(1)}, t+m+\tau)$. Arbitrarily select one of the $S-1$ units and observe its outdating time, T_i. That is, we select the ith item (among the $S-1$) with probability $1/(S-1)$. We wish to show that T_i is uniformly distributed on $[T_{(1)}, t+m+\tau)$. Now, if T_i is an outdating time, then $T_i-(m+\tau)$ is a departure time representing either a demand arrival or an outdating, so that $T_i \epsilon [T_{(1)}, t+m+\tau)$ implies $T_{(1)}-(m+\tau) \leqslant T_i-(m+\tau) < t$. If $T_i-(m+\tau)$ is a demand arrival time, then it is uniformly distributed on $[T_{(1)}-(m+\tau), t)$ because the demand process is Poisson. If $T_i-(m+\tau)$ is not a demand arrival time then it must be an outdating time. In this case, $T_i-2(m+\tau)$ is also a departure time that is either a demand time or an outdating time. Again, $T_{(1)}-2(m+\tau) \leqslant T_i-2(m+\tau) < t-(m+\tau)$. By repeating the same argument we conclude that for some $n = 1, 2, \ldots$:

$$T_{(1)} - n(m+\tau) \leqslant T_i - n(m+\tau) < t-(n-1)(m+\tau),$$

so that for some n, $T_i - n(m+\tau)$ is a demand arrival time w.p.l. Since the process is *old*, t is far enough away from the origin, and thus $T_i - n(m+\tau)$ is uniformly distributed on $[T_{(1)} - n(m+\tau), t-(n-1)(m+\tau))$. By shifting this whole interval (containing point $T_i - n(m+\tau)$) by $n(m+\tau)$ time units to the right, we obtain that T_i is uniformly distributed on $[T_{(1)}, t+m+\tau)$. By repeating this argument, we get that $T_j - n(m+\tau)$ is a time of a satisfied demand for some $n = 1, 2, \ldots$ and all $j = 1, \ldots, S-1$, and the lemma is an application of a well-known property associated with the Poisson process. □

A typical realization of the above description is given in Figure 3.

As an immediate consequence of Lemma 1 we have the following:

Theorem 1. *The difference $W_{s-1} - W_s$ has the same conditional law (given W_s) as that of the minimal order statistic (among $s-1$) taken from a uniform distribution on $[W_s, m+\tau)$.*

Substituting this into (2.9) yields:

$$f(x) = \begin{cases} f(0)\left(\dfrac{m+\tau-x}{m+\tau}\right)^{s-1} + \lambda \displaystyle\int_0^x \left(\dfrac{m+\tau-x}{m+\tau-\omega}\right)^{s-1} f(\omega)\, d\omega, \\ \qquad 0 < x < m, \\ f(0)\left(\dfrac{m+\tau-x}{m+\tau}\right)^{s-1} + \lambda \displaystyle\int_0^m \left(\dfrac{m+\tau-x}{m+\tau-\omega}\right)^{s-1} f(\omega)\, d\omega \\ \quad + \lambda p \displaystyle\int_m^x [1-H(\omega-m)]\left(\dfrac{m+\tau-x}{m+\tau-\omega}\right)^{s-1} f(\omega)\, d\omega, \\ \qquad x \geqslant m. \end{cases} \tag{3.1}$$

4. THE SPECIAL CASE (0.1)

In this section, we explore the particular application of H of the form (0.1). Substituting (0.1) into (3.1) we have:

$$f(x) = \begin{cases} f(0)\left(\dfrac{m+\tau-x}{m+\tau}\right)^{s-1} + \lambda \displaystyle\int_0^x \left(\dfrac{m+\tau-x}{m+\tau-\omega}\right)^{s-1} f(\omega)\, d\omega, \\ \qquad 0 < x < m, \\ f(0)\left(\dfrac{m+\tau-m}{m+\tau}\right)^{s-1} \\ \quad + \lambda \displaystyle\int_0^m \left(\dfrac{m+\tau-x}{m+\tau-\omega}\right)^{s-1} f(\omega)\, d\omega \\ \quad + \lambda p \displaystyle\int_m^x \left(\dfrac{m+\tau-x}{m+\tau-\omega}\right)^{s-1} f(\omega)\, d\omega, \\ m \leqslant x < m+\tau. \end{cases} \tag{4.1}$$

Set $g(x) = f(x)/(m + \tau - x)^{S-1}$. We then get from (4.1) that:

$$g(x) = \begin{cases} \dfrac{f(0)}{(m+\tau)^{s-1}} + \lambda G(x), & x < m, \\ \dfrac{f(0)}{(m+\tau)^{s-1}} + \lambda G(m) + \lambda p[G(x) - G(m)], & \\ \qquad m \leqslant x < m + \tau, \end{cases} \tag{4.2}$$

where $G(x) = \int_0^x g(\omega)\, d\omega$. Solving for g in (4.2) and using the obvious continuity condition $g(m^+) = g(m^-)$ (see Perry and Posner 1990), we obtain:

$$g(x) = \begin{cases} \alpha e^{\lambda x}, & 0 < x < m, \\ \alpha e^{\lambda q m + \lambda p x}, & m \leqslant x < m + \tau, \end{cases} \tag{4.3}$$

for some positive α. From (4.3) we then have:

$$f(x) = \begin{cases} \beta\lambda\phi(S-1, m+\tau-x), & 0 < x < m, \\ \beta\lambda e^{-\lambda q(x-m)}\phi(S-1, m+\tau-x), & m \leqslant x < m+\tau \end{cases} \tag{4.4}$$

for some positive β, where:

$$\phi(j, \theta) = \frac{e^{-\lambda\theta}(\lambda\theta)^j}{j!}, \quad j = 0, 1, \ldots, .$$

Note that, for $p > 0$, (4.4) can also be written:

$$f(x) = \begin{cases} \beta\lambda\phi(S-1, m+\tau-x), & 0 < x < m, \\ \dfrac{\beta\lambda e^{-\lambda q\tau}}{p^{S-1}}\phi(S-1, p(m+\tau-x)), & m \leqslant x < m+\tau. \end{cases} \tag{4.5}$$

From the normalization condition $\int_0^{m+\tau} f(x)\, dx = 1$ we then obtain:

$$\beta^{-1} = \frac{e^{-\lambda q\tau}}{p^S}\left(1 - \sum_{r=0}^{S-1}\phi(r, p\tau)\right) + \sum_{r=0}^{S-1}(\phi(r, \tau) - \phi(r, m+\tau)). \tag{4.6}$$

The constant $\beta \equiv \beta(p)$ is determined according to the parameter p. Now, let $\beta_0 = \lim_{p\downarrow 0}\beta(p)$ and $\beta_1 = \lim_{p\uparrow 1}\beta(p)$. It follows easily from (4.6) that:

$$\beta_1^{-1} = 1 - \sum_{r=0}^{S-1}\phi(r, m+\tau), \tag{4.7}$$

and for β_0 it can also be shown that:

$$\beta_0^{-1} = \phi(S, \tau) + \sum_{r=0}^{S-1}(\phi(r, \tau) - \phi(r, m+\tau)).$$

5. NUMBER OF ITEMS ON THE SHELF

Let $P_j (j = 0, 1, \ldots, S)$ be the steady state probability that j items are on the shelf. To compute P_j, we note that j items are available if and only if m is somewhere between the $(j-1)$th and the jth order statistics. Thus, we have, for $j = 1, \ldots, S$, that:

$$P_j = \int_0^m \binom{S-1}{j-1}\left(\frac{m-\omega}{m+\tau-\omega}\right)^{j-1} \cdot \left(\frac{\tau}{m+\tau-\omega}\right)^{S-j} f(\omega)\, d\omega, \tag{5.1}$$

and,

$$P_0 = Pr(W_s \geqslant m) = \int_m^{m+\tau} f(\omega)\, d\omega. \tag{5.2}$$

Other useful measures can also be obtained once f is known. Invoking level-crossing theory, it readily follows that the rate of the outdating process is $f(0)$, and the rate of replenishment item arrivals to an empty shelf is $f(m)$. In addition, the rate of the unsatisfied demand process is:

$$\lambda\int_m^{m+\tau} H(\omega - m) f(\omega)\, d\omega, \tag{5.3}$$

by PASTA (Wolff 1988).

For general H, we may compute all these measures numerically. However, for the special case of (0.1) we will present an analytic solution. To compute P_j for $j = 1, 2, \ldots, S$, substitute (4.5) into (5.1), yielding:

$$P_j = \beta\phi(S-j, \tau)\left[1 - \sum_{r=0}^{j-1}\phi(r, m)\right], \tag{5.4}$$

and

$$P_0 = \frac{\beta e^{-\lambda q\tau}}{p^{S-1}}\left[1 - \sum_{r=0}^{S-1}\phi(r, p\tau)\right]. \tag{5.5}$$

Furthermore, the rate of the outdating process is

$$f(0) = \beta\lambda\phi(S-1, m+\tau). \tag{5.6}$$

Now, the extreme substitution of $\beta = \beta_1$ into (5.4), (5.5), and (5.6) is appropriate for the model in which customer demands that arrive to find an empty shelf are *always* willing to wait. Similarly, the extreme substitution $\beta = \beta_0$ is appropriate for the model in which such customers *never* wait. This latter substitution is, in fact, the identical result obtained by Schmidt and Nahmias (1985).

6. OTHER WAITING TIME POLICIES

In order to demonstrate the versatility of our model, we now choose several examples of H for which numerical work is required to obtain an explicit solution.

6.1. Exponential Waiting Time

Under this policy we substitute $H(y) = 1 - e^{-\eta y}$ and obtain:

$$f(x) = \begin{cases} \alpha e^{\lambda x}(m+\tau-x)^{s-1}, & 0 \leqslant x < m, \\ \alpha e^{\lambda m - \lambda/2(1-e^{-\eta(x-m)})}[m+\tau-x], & m \leqslant x < m+\tau, \end{cases} \tag{6.1}$$

where α is obtained numerically from the normalizing condition.

6.2. Rayleigh Waiting Time

Under this policy we substitute $H(y) = 1 - e^{-y^2/2\sigma^2}$ and get:

$$f(x) = \begin{cases} \alpha e^{\lambda x}(m + \tau - x)^{S-1}, & 0 \leqslant x < m, \\ \alpha e^{\lambda m - \lambda\sigma\sqrt{2\pi}\left[\Phi\left(\frac{x-m}{\sigma}\right)-\frac{1}{2}\right]} \cdot (m + \tau - x), & m \leqslant x < m + \tau, \end{cases} \tag{6.2}$$

where $\Phi(\cdot)$ designates the standard normal distribution.

Note that the Rayleigh distribution has the unique feature that its hazard rate is linear. This would be natural and useful in cases where customers can observe their waiting times before making a decision.

6.3. Uniform Waiting Time

Under this policy we substitute $H(y) = (y - m)/\tau$, $(m < y < m + \tau)$, and obtain:

$$f(x) = \begin{cases} \alpha e^{\lambda x}(m + \tau - x)^{S-1}, & 0 \leqslant x < m, \\ \alpha e^{\lambda m + 1/2\left(\frac{m-\lambda(m+\tau)}{1/\sqrt{\tau}}\right)^2 - \frac{1}{2}\left(\frac{x-\lambda(m+\tau)}{1/\sqrt{\tau}}\right)^2} \cdot (m + \tau - x)^{S-1}, & m \leqslant x < m + \tau. \end{cases} \tag{6.3}$$

ACKNOWLEDGMENT

We are grateful to the anonymous referee for his helpful suggestions that enabled us to model more general classes of waiting time policies.

REFERENCES

GRAVES, S. C. 1982. The Applications of Queueing Theory to Continuous Perishable Inventory Systems. *Mgmt. Sci.* **28,** 400–406.

KASPI, H. AND D. PERRY. 1983. Inventory Systems of Perishable Commodities. *Adv. Appl. Probab.* **15,** 674–685.

KASPI, H. AND D. PERRY. 1984. Inventory Systems for Perishable Commodities with Renewal Input and Poisson Output. *Adv. Appl. Probab.* **16,** 402–421.

MOINZADEH, K. AND C. P. SCHMIDT. 1991. An $(S - 1, S)$ Inventory System with Emergency Orders. *Opns. Res.* **39,** 308–321.

PERRY, D. AND S. ASMUSSEN. 1995. Rejection Rules in the *M/G/*1 Queue. *Questa,* **19,** 105–130.

PERRY, D. AND M. J. M. POSNER. 1989. Control Policies for Two Classes of Inventory Systems via a Duality Equivalence Relationship. *Prob. Eng. Inf. Sci.* **3,** 561–579.

PERRY, D. AND M. J. M. POSNER. 1990. Control of Input and Demand Rates in Inventory Systems of Perishable Commodities. *Naval Res. Logist.* **37,** 85–97.

SCHMIDT, C. P. AND S. NAHMIAS. 1985. $(S - 1, S)$ Policies for Perishable Inventory. *Mgmt. Sci.* **31,** 719–728.

SHERBROOKE, C. C. 1968. METRIC: A Multi-echelon Technique for Recoverable Item Inventory Control. *Opns. Res.* **16,** 122–141.

SMITH, S. A. 1977. Optimal Inventories for an $(S - 1, S)$ System With No Backorders. *Mgmt. Sci.* **23,** 522–528.

WOLFF, R. 1988. *Stochastic Modeling and the Theory of Queues.* Academic Press, New York.