Discussion Paper No. 214

OPTIMAL CONTROL OF SOME MARKOV PROCESSES WITH APPLICATIONS TO BATCH QUEUEING AND CONTINUOUS REVIEW INVENTORY SYSTEMS

by

Howard J. Weiss[1]

and

Stanley R. Pliska[2]

April 1976

This research was partially supported by Research and Training Grant IL-11-0005 of the Urban Mass Transportation Administration.

[1] Department of Quantitative and Information Science, College of Business, Western Illinois University, Macomb, Illinois 61455.

[2] Department of Industrial Engineering and Management Sciences, The Technological Institute, Northwestern University, Evanston, Illinois 60201.

ABSTRACT

This paper analyses a class of Markov process control problems where the objective is to minimize the expected cost per unit time. Using some renewal and Markov process theory, it is shown that there exists an optimal policy which is an element of the class of what are called derivative policies. An algorithm is provided for computing such an optimal policy.

The model is then applied to the problems of controlling (1) a queue with batch service, (2) a continuously reviewed inventory system, and (3) a pestilent population. The first two applications are noteworthy, because in the first case one can allow for a customer's waiting cost to be a nonlinear function of his waiting time and in the second case one can allow the shortage cost per unit of excess demand to be a nonlinear function of the shortage time. It is also shown that with linear costs the optimal derivative policy becomes identical to a control-limit policy in the queueing case and to an (s, S) policy in the inventory case. However, with other, more general cost structures, control-limit policies will not be optimal for the batch queueing case and (s, S) policies will not be optimal for the continuous review inventory case.

In this paper, we study a class of Markov process control problems that have application to some controlled, batch service queues, to some continuous review inventory models, and to some controlled population processes. For ease of exposition, we shall present our main results in the queueing context, postponing the discussion of the other two applications until the final two sections.

Consider a batch service queueing system where the number of customers waiting for the first service is described by a nondecreasing stationary Markov process $\{N(t) : t \geq 0\}$ which has state space the nonnegative integers. For example, the arrival process for the queueing system might be Poisson, in which case $\{N(t) : t \geq 0\}$ would be too. Alternatively, $\{N(t) : t \geq 0\}$ could be such that the arrival rate depends on the number of customers present.

At each epoch S_n, $n = 1, 2, \ldots$, a batch of customers is served, thereby interrupting the process $\{N(t) : t \geq 0\}$ and causing it to resume from some new state $N(S_n)$, where $N(S_n^-) - N(S_n)$ equals the number of customers served in the nth batch. For example, if all of the customers present are served, then $\{N(t) : t \geq 0\}$ resumes starting at zero.

The service times comprise a sequence $\{D_n\}$ of independent and identically distributed random variables. The queueing system has a single, infinite capacity server, the service times are independent of the batch size, and we require $S_{n+1} \geq S_n + D_n$, $n = 1, 2, \ldots$ We also assume $E[D_n] < \infty$.

The queueing process is continuously reviewed, and we wish to find the service policy that minimizes the long run (infinite horizon) expected cost per unit time for the following cost structure. Each time a batch of customers is served, the queueing system incurs a fixed cost K, a positive constant independent of batch size. In addition, the queueing system incurs a customer waiting cost.

The most noteworthy feature of our model is the general form of the customer waiting cost. Let the random variable $\mathcal{C}(t)$ equal the total waiting cost that is incurred by the queueing system during the first t time-units of its operation. Of course, $\mathcal{C}(\cdot)$ is a functional of the process $\{N(t) : t \geq 0\}$, for $\mathcal{C}(t)$ is a function of the history of the process up through epoch t. We shall make the following

Assumptions: For any realization of the process,

(i) $\mathcal{C}(t) \geq 0, \ t \geq 0$,

(ii) the right-hand derivative $\mathcal{C}'(t)$ exists and is nonnegative and nondecreasing between successive service epochs, and

(iii) for any $t_1 < t_2$, $\mathcal{C}(t_2) - \mathcal{C}(t_1)$ is nondecreasing with respect to $N(t_1)$.

(iv) Suppose $N(S_n) = 0$ for some service starting epoch S_n and $S_{n+1} = \infty$. Then $\mathcal{C}(\tau + S_n) - \mathcal{C}(S_n)$ is independent of S_n as well as the history of the process before S_n for all $\tau \geq 0$.

Roughly speaking, the second assumption says that the marginal system waiting cost (the waiting cost rate) is nondecreasing with respect to time, and the third assumption says that the more customers there are, the greater the cost. Our model clearly allows for the standard linear case in the literature, namely

$$\mathcal{C}(t) = h \int_0^t N(s)ds \,, \tag{1}$$

but we can also be more general by specifying

$$\mathcal{C}(t) = \mathcal{C}(S_n) + \int_{S_n}^{t} h\Big(N(s)\Big)ds\ , \quad S_n \le t < S_{n+1}, \quad n = 0, 1, \ldots,$$

where $S_0 \equiv 0$ and h is a nonnegative, nondecreasing function on $[0, \infty)$. This generalization allows for the customers to become more aggravated as the waiting facility becomes more crowded. In addition, the waiting cost in our model could include a term like

$$\sum_{i=N(S_n)+1}^{N(t)} h(W_i)\ , \qquad S_n \le t < S_{n+1}\ , \quad n = 0, 1, \ldots, \tag{2}$$

where W_i is the length of time that customer i has waited for service and h is a nonnegative, nondecreasing, convex function on $[0, \infty)$. In other words, the system incurs a cost for each customer according to how long the customer must wait, and we can allow for the customers to become more aggravated as the wait becomes longer.

Of the related models studied in the literature, ours most closely resembles a model studied by Deb and Serfozo [1]. In their model the server could have a finite or infinite capacity, but the waiting cost is linear as in (1) above. They show that a control-limit policy is optimal, that is, for some number M^*, serve if and only if the server is available and there are M^* or more customers waiting for service. In contrast, we shall prove that for more general waiting costs a control-limit policy is not optimal, and, in fact, we shall present the optimal policy. Our results are applicable to several models, most of which, due to the assumption of linear waiting costs, were originally presented as Markov or semi-Markov decision processes. These models include the post office example in Ross [4, p. 164], the machine repair example in Ross [4, p. 129], the custodian problem of Kosten [3], and the infinite capacity shuttle of Ignall and Kolesar [2].

In section 1 we present our main results and identify a class of policies which contains an optimal policy. We have termed these policies derivative policies for reasons which will become obvious. Essentially, an optimal policy is of the following form: begin service if and only if the server is available and the marginal cost of not beginning service is greater than or equal to the optimal long average cost.

Derivative policies are only slightly more complex than control-limit policies, and they have some nice features. First, in the case of linear waiting costs as in (1) above, the two kinds of policies are equivalent. Secondly, even though a derivative policy is an optimal continuous review policy, the times at which we need to review the system are the same as with control-limit policies, namely, the service completion and customer arrival epochs. Finally, and most importantly, optimal derivative policies are easy to compute. More will be said about computing and implementing optimal policies in section 2.

Although our main results are stated in the context of a controlled queueing problem, our model can be applied to a variety of other applications. In section 3 we illustrate the generality of our model by applying it to the problem of controlling a pestilent population, where the population grows according to a branching process.

We conclude, in section 4, by applying our results to some continuous review inventory models of the type studied by Sivazlian [5]. In his model, the holding and shortage costs are linear, but we extend these assumptions to convex holding and shortage costs. We prove that the optimal ordering epochs are given by a derivative policy which, in the special case of linear holding and shortage costs, is equivalent to an (s, S) policy.

1. DERIVATIVE POLICIES

Our objective is to determine a policy which minimizes the expected cost per unit time

$$\lim_{t\to\infty} \frac{E[\mathcal{C}(t)]}{t} .$$

We shall consider all policies that specify the service starting epochs as well as which customers are served at each service epoch.

In view of Assumption (iii), however, it should be clear that under any optimal policy, every customer present at a service epoch would be included in the corresponding service batch. This can be demonstrated formally by taking an arbitrary policy and constructing a new policy with exactly the same service epochs but where all of the customers present are served. Then by taking an arbitrary realization of the process $\{N(t) : t \geq 0\}$, Assumption (iii) implies $\mathcal{C}(t)$ under the new policy is less than or equal to what it is under the original policy, in which case the same thing can be said for the expected cost per unit time (see [6]).

Hence it suffices to confine our attentions to the choice of the service starting epochs $\{S_n\}$. In particular, at each such epoch, the process $\{N(t) : t \geq 0\}$ resumes from the same fixed state, namely state zero. This and Assumption (iv) imply that the optimal choice of $S_{n+1} - S_n$ is independent of the history of the process before epoch S_n for each $n = 1, 2, \ldots$ Hence under any optimal policy, the sequence $\{S_{n+1} - S_n\}$ will be independent and identically distributed, and, by standard results in renewal theory (e.g., Ross [4]), the expected cost per unit time will be equal to the expected cost incurred between service starting epochs divided by the expected time between service starting epochs.

It should now be apparent that we have reduced our queueing control problem to an optimal stopping problem. Suppose $N(0) = 0$. Let $\mathcal{T}$ denote the set of all stopping rules T for the process $\{N(t) : t \geq 0\}$ which evolves as if the customers are never served. For each $T \in \mathcal{T}$, we also specify $T \geq D$ (recall D is the service time) and $E[T] < \infty$, and we denote

$$g(T) \equiv \frac{K + E[\mathcal{C}(T)]}{E[T]} .$$

Thus if $T \in \mathcal{T}$ is used to define the queueing control policy where the sequence $\{S_{n+1} - S_n\}$ is independent and identically distributed as T, then the corresponding expected cost per unit time will equal g(T). Moreover,

$$c^* = \inf_{T \in \mathcal{T}} g(T)$$

will equal the minimum expected cost per unit time for the queueing control problem, and the policy corresponding to $T \in \mathcal{T}$ will be optimal if $g(T) = c^*$.

We shall now define what we call derivative policies:

<u>Definition</u>: A <u>derivative policy</u> is the policy corresponding to some real number c and the stopping rule T defined by

(3) $$T = \min\{t \geq D : \mathcal{C}'(t) \geq c\} .$$

In words, a derivative policy is a continuous review control policy for batch service queues such that service is initiated (for all the customers present) at time t if and only if the server is available and the marginal system waiting cost at time t is greater than or equal to some specified positive constant.

For the remainder of this section, we shall concentrate on the optimal stopping problem. For any real number c, let T_c denote the corresponding stopping rule defined by (3). If T, $T' \in \mathcal{T}$, then we can define a new stopping rule

$$T \vee T' \equiv \max\{t \geq 0 : t \geq T \text{ and } t \geq T'\} .$$

Clearly $T \vee T' \in \mathcal{T}$, because $E[T \vee T'] \leq E[T + T'] < \infty$. Similarly, $T \wedge T' = \min\{t \geq 0 : t \geq T \text{ or } t \geq T'\} \in \mathcal{T}$.

At this point, a technicality needs to be taken care of, because there is no guarantee, for some particular number c, that $E[T_c] < \infty$ and $T_c \in \mathcal{T}$. We could ensure this with some conditions on the functional $\mathcal{C}$, but we find it more convenient to simply make the following

<u>Assumption</u>:

(v) For any number $c < \infty$, $E[T_c] < \infty$.

This condition is usually easy to check. For example, if $\mathcal{C}$ is given by (1) and $\{N(t) : t \geq 0\}$ is a Poisson process with parameter λ, then $\mathcal{C}'(t) = hN(t)$. Now $T_c = T \vee D$, where

$$T = \min\{t \geq 0 : hN(t) \geq c\} ,$$

so it suffices to check $E[T] < \infty$. It is straightforward to compute

$$E[T] = \int_0^\infty P(T > t)dt = \int_0^\infty P\Big(hN(t) < c\Big)dt$$

$$= \int_0^\infty P\Big(N(t) < [c/h]\Big)dt = [c/h]1/\lambda ,$$

where $\lceil c/h \rceil$ equals the smallest integer greater than or equal to c/h.

We now present two lemmas.

Lemma 1. For any $T \in \mathcal{T}$, $g(T_{max}) \le g(T)$, where $T_{max} \equiv T \vee T_{g(T)}$.

Proof. Let A denote the event $\{T < T_{g(T)}\}$, A^c denote the complement of A, and 1(A) denote the indicator function for the event A. Then

$$g(T_{max}) = \frac{K + E\int_0^{T_{max}} \mathcal{C}'(t)dt}{E[T_{max}]}$$

$$= \frac{K + E\left[1(A)\int_0^{T_{max}} \mathcal{C}'(t)dt\right] + E\left[1(A^c)\int_0^{T_{max}} \mathcal{C}'(t)dt\right]}{E[T_{max}]}$$

$$= \frac{K + E\left[1(A)\int_0^{T} \mathcal{C}'(t)dt\right] + E\left[1(A)\int_T^{T_{g(T)}} \mathcal{C}'(t)dt\right] + E\left[1(A^c)\int_0^{T} \mathcal{C}'(t)dt\right]}{E[T_{max}]}$$

$$= \frac{K + E\left[\int_0^{T} \mathcal{C}'(t)dt\right] + E\left[1(A)\int_T^{T_{g(T)}} \mathcal{C}'(t)dt\right]}{E[T_{max}]}$$

$$\le \frac{K + E\left[\int_0^{T} \mathcal{C}'(t)dt\right] + g(T)E[1(A)\{T_{g(T)} - T\}]}{E[T_{max}]}$$

$$= \frac{E[T]g(T) + g(T)\{E[T_{max}] - E[T]\}}{E[T_{max}]} = g(T) ,$$

where the inequality is due to the fact that $\mathcal{C}'(t) \le g(T)$ for all $T \le t < T_{g(T)}$, and the succeeding equality follows from the definition of g(T) and the identity

$$E[T_{max}] = E[1(A)T_{g(T)}] + E[1(A^c)T]$$
$$= E[1(A)T_{g(T)}] - E[1(A)T] + E[T] .$$

Lemma 2. For any $T \in \mathcal{T}$, $g(T_{g(T)}) \leq g(T)$.

Proof. In view of Lemma 1, it suffices to show $g(T_{g(T)}) \leq g(T_{max})$. Let A denote the event $\{T_{g(T)} < T_{max}\}$, and notice that $A^c = \{T_{g(T)} = T_{max}\}$, since $T_{max} = T \vee T_{g(T)} \geq T_{g(T)}$. Proceeding in a manner similar to the proof of Lemma 1, we obtain

$$g(T_{max}) = \frac{K + E\left[1(A^c)\int_0^{T_{g(T)}} \mathcal{C}'(t)dt\right] + E\left[1(A)\int_0^{T_{g(T)}} \mathcal{C}'(t)dt\right] + E\left[1(A)\int_{T_{g(T)}}^{T_{max}} \mathcal{C}'(t)dt\right]}{E[T_{max}]}$$

$$\geq \frac{E[T_{g(T)}]g(T_{g(T)}) + g(T_{max})E[1(A)\{T_{max} - T_{g(T)}\}]}{E[T_{max}]}$$

$$= \frac{E[T_{g(T)}]}{E[T_{max}]} [g(T_{g(T)}) - g(T_{max})] + g(T_{max}) ,$$

where the inequality is due to the fact that $\mathcal{C}'(t) \geq g(T) \geq g(T_{max})$ for all $T_{g(T)} \leq t < T_{max}$, and the last equality follows from the identity

$$E[T_{g(T)}] = E[1(A)\{T_{g(T)} - T_{max}\}] + E[T_{max}] .$$

Hence $g(T_{g(T)}) - g(T_{max}) \leq 0$, and this proof is completed.

Theorem 3. The derivative policy with $c = c^*$ is the optimal policy, that is, $g(T_{c^*}) = c^*$.

Proof. There exists a sequence $\{T_n\}$ of stopping rules with $g(T_n) \downarrow c^*$ as $n \to \infty$. If we set $c_n = g(T_n)$, then by Lemma 2 the sequence $\{T_{c_n}\}$ of stopping rules satisfies $g(T_{c_n}) \downarrow c^*$ as $n \to \infty$.

Now $c_n \downarrow c^*$, so $T_{c^*} \le T_{c_n}$ and

$$g(T_{c_n}) = \frac{K + E\left[\int_0^{T_{c^*}} \mathcal{C}'(t)dt\right] + E\left[\int_{T_{c^*}}^{T_{c_n}} \mathcal{C}'(t)dt\right]}{E[T_{c_n}]}$$

$$\ge \frac{E[T_{c^*}]}{E[T_{c_n}]} g(T_{c^*}) + \frac{E[T_{c_n} - T_{c^*}]}{E[T_{c_n}]} c^* .$$

For any $\epsilon > 0$ and all large enough n, it follows that

$$\epsilon > g(T_{c_n}) - c^* \ge \frac{E[T_{c^*}]}{E[T_{c_n}]} [g(T_{c^*}) - c^*]$$

$$\ge \frac{E[T_{c^*}]}{E[T_{c_1}]} [g(T_{c^*}) - c^*] .$$

Since ϵ is arbitrary, we must have $g(T_{c^*}) = c^*$, and this proof is completed.

Corollary 4. If $\mathcal{C}(\cdot)$ is strictly convex on the set $\{t \ge 0 : \mathcal{C}(t) > 0\}$, then the optimal policy is unique.

Proof. Of course, T_{c^*} is one optimal policy; suppose $T \in \mathcal{T}$ is another. Let $T_{max} = T \vee T_{g(T)} = T \vee T_{c^*}$ and A denote the event $\{T < T_{c^*}\}$. Following the proof of Lemma 1, we see that the strict convexity implies $\mathcal{C}'(t) < g(T)$

for all $T \le t < T_{g(T)} = T_{c^*}$ which, in turn, implies the inequality there is strict unless $P(A) = 0$. But $g(T_{max}) < g(T) = c^*$ would be a contradiction, so we must have $P(A) = 0$ and $T \ge T_{c^*}$.

Now $T \ge T_{c^*}$ implies $T_{max} = T$, so we can let A denote the event $\{T_{c^*} < T\}$ and follow the proof of Lemma 2. Again, the strict convexity would imply $g(T_{c^*}) < g(T_{max}) = g(T)$, a contradiction, unless $P(A) = 0$. Hence $P(T_{c^*} = T) = 1$, and this proof is completed.

Remark. The hypothesis of Corollary 4 is satisfied if the cost functional contains a term like (2) with the function h there strictly convex.

2. POLICY COMPUTATION AND IMPLEMENTATION

According to Theorem 3, knowing the minimum expected cost per unit time c^* is equivalent to knowing the optimal policy. In Theorem 6 below we shall provide an explicit algorithm for computing c^*. This algorithm is similar in spirit to the method of successive approximations in Markov decision theory.

Lemma 5. If $c > c^*$, then $c - g(T_c) \ge \frac{E[T_{c^*}]}{E[T_c]} [c - c^*]$.

Proof. This lemma follows immediately from the following inequality:

$$g(T_c) = \frac{K + E\left[\int_0^{T_{c^*}} \mathcal{C}'(t)dt\right] + E\left[\int_{T_{c^*}}^{T_c} \mathcal{C}'(t)dt\right]}{E[T_c]}$$

$$= \frac{E[T_{c^*}]}{E[T_c]} c^* + \frac{E\left[\int_{T_{c^*}}^{T_c} \mathcal{C}'(t)dt\right]}{E[T_c]}$$

$$\le \frac{E[T_{c^*}]}{E[T_c]} c^* + \frac{E[T_c - T_{c^*}]}{E[T_c]} c .$$

Theorem 6. Let $T \in \mathcal{T}$ be arbitrary, and set $c_1 = g(T)$. Define recursively $c_{n+1} = g(T_{c_n})$, $n = 1, 2, \ldots$ Then $\lim_{n\to\infty} c_n = c^*$.

Proof. If at any stage $c_{n+1} = c_n$, then by Lemma 5 we must have $c_n = c^*$. Alternatively, for some number $c \geq c^*$, we must have $c_n \downarrow c$. It remains to show $c = c^*$.

Suppose $c > c^*$. Then, by Lemma 5,

$$c_n - c_{n+1} \geq \frac{E[T_{c^*}]}{E[T_{c_n}]} [c_n - c^*] \geq \frac{E[T_{c^*}]}{E[T_{c_1}]} [c - c^*] > 0 ,$$

which contradicts the fact that $\{c_n\}$ is Cauchy.

This algorithm is generally an infinite one, so it would be nice to be able to estimate a lower bound for c^*. One fairly practical procedure for doing this is as follows: Based upon the initial iterations of the algorithm, make an initial estimate, say c_b, of a lower bound for c^*. Next, calculate $g(T_{c_b})$. If $c_b > g(T_{c_b})$, then you know $c_b > c^*$ and your estimate was too high. Alternatively, if $c_b < g(T_{c_b})$, then $c_b < c^*$ (by Lemma 5), so you can either revise upward your estimate c_b, continue with the algorithm, or terminate.

Although a derivative policy is the optimal continuous review control policy, the nature of the process $\{N(t) : t \geq 0\}$ is such that it suffices to examine the system at just the customer arrival epochs and the service completion epoch. At each customer arrival epoch, the controller revises his calculation of the time $T = \min\{t \geq 0 : \mathcal{C}'(t) \geq c^*\}$, given the assumption that no more customers will arrive. The controller then waits until either $T \vee D$ or the next customer arrival epoch, whichever occurs first.

3. SOME GENERALIZATIONS AND AN APPLICATION TO POPULATION CONTROL

Although our control model was presented in the context of a queueing system, it has two other applications that we know of. Here we shall briefly describe some generalizations of our model and then mention a population control problem. In section 4 we shall analyse an inventory control problem.

The primary feature of our model is a population of people, customers, or things that can be described by a stationary, nondecreasing Markov process $\{N(t) : t \geq 0\}$. We can generalize the state space for this process to the set of all integers greater than or equal to N, a nonnegative integer. Then the cost functional $\mathcal{C}(t)$ satisfies the same Assumptions (i-v) except that, in (iv), $N(S_n) = 0$ is replaced by $N(S_n) = N$. All of the arguments we made for the queueing problem then follow. In particular, it is optimal to use a derivative policy where the population is always reduced to state N at each intervention.

Two other generalizations are possible if we can restrict our attentions to policies that always return the process to state N at each service epoch. In the queueing context of section 1, we tacitly assumed the random variable D for the minimum time between service epochs was independent of the process $\{N(t) : t \geq 0\}$. However, for Theorem 3, we only needed the fact that $\{D_n\}$ comprises a sequence of independent and identically distributed random variables, so we can generalize D to be any stopping time for the process $\{N(t) : t \geq 0\}$ with $E[D] < \infty$. If the process always resumes from state N, then the control cannot affect D and thus cannot affect the set $\mathcal{T}$ of admissible controls.

We can also assume the service cost is a random variable which is a certain function of the process. Let the random variable K_n equal the cost of the nth service, assume $E[K_n] < \infty$, and assume K_n is a function of $\{N(t) : 0 \leq t \leq D_n\}$, where D_n is the stopping time previously mentioned. If the process always resumes from state N, then the control cannot affect K_n, and $\{K_n\}$ is a sequence of independent and identically distributed random variables. The average cost for any $T \in \mathcal{T}$ is as before, only with K replaced by $E[K_n]$.

A word of warning: it is possible to formulate a control problem with a stopping time D and random service cost K_n such that it is never optimal to return the system to state N. Hence, to make the preceding two generalizations, it is usually necessary to either verify that it is indeed optimal for the process to resume from state N or, failing that, to arbitrarily restrict one's attentions to such policies.

The motivation for the preceding two generalizations will become apparent in the following section. To illustrate the generality that is possible for the underlying process $\{N(t) : t \geq 0\}$, consider the problem of controlling a population of, for example, wild game, pests, or infected individuals (as in the case of an epidemic model). Populations such as these have been modeled as branching processes. Periodically, a controller intervenes and reduces the population to some state $N > 0$. For our purposes, we need only add the requirement that the process be nondecreasing.

4. A CONTINUOUS REVIEW INVENTORY MODEL

In this section we consider a control problem quite similar to the continuous review inventory model presented by Sivazlian [5]. The cumulative demands for an item form a Poisson process. There is a nonnegative holding cost $h(R_i)$ per item, where R_i is the length of time the ith unit is held in the inventory. In addition, there is a nonnegative shortage cost $p(W_j)$ per unit of excess demand, where W_j is the length of time that the jth demand is unsatisfied, and p is a nonnegative, increasing, convex function on $[0, \infty)$.

The system is continuously reviewed, and orders are periodically made. For some fixed numbers F and S, each time an order is made, a fixed cost F is incurred and the inventory level is instantaneously replenished up to level S. The problem is to determine the optimal policy for the ordering epochs so as to minimize the average cost. If we can restrict our attentions to ordering policies that wait until the inventory is depleted before reordering, then it is optimal to follow what we call an (S, derivative) policy. By an (S, derivative) policy, we mean that the timing of the ordering epochs is dictated by a derivative policy.

To show that an (S, derivative) policy is indeed optimal, and to compute an optimal policy, it suffices to formulate our inventory problem in terms of the model of section 3. For the underlying Markov process $\{N(t) : t \geq 0\}$, we shall take the Poisson process demand model. Set $N = 0$. State 0 corresponds to inventory level S, state S corresponds to inventory level 0, and $\left(N(t) - S\right) \vee 0$ equals the excess demand. Set $D = \inf\{t : N(t) = S\}$, so that the set $\mathcal{T}$ consists of all policies that wait until the inventory is depleted before reordering. Notice that $\mathcal{T}$ contains all ordinary (S, s) policies.

It remains to specify the costs. Set K equal to F plus the expected holding cost per ordering period. Notice the holding cost may be a function of $\{N(t) : 0 \leq t \leq D\}$. Finally, for the cost functional, set

$$\mathcal{C}(t) = \begin{cases} \sum_{j=N(D)+1}^{N(t)} p(W_j) & , \quad t > D \\ 0 & , \quad t \leq D . \end{cases}$$

It is easy to verify that Assumptions (i) - (v) are satisfied, so we can summarize our ideas in the following

Theorem 7. The inventory system described above, where all admissible ordering policies wait for the inventory to become depleted before re-ordering, is equivalent to the controlled Markov process formulated above, and an (S, derivative) policy is the optimal ordering policy.

Remark. It should be emphasized that the inventory replenishment level S is a fixed constant, not a variable in the optimization problem. Naturally, if the optimal (S, derivative) policy has been determined for different values of S, then the inventory can select that which corresponds to the minimum average cost.

We only considered policies that order at nonpositive inventory levels. For certain holding cost functions, in particular, if h is nonnegative, convex, and increasing, then such policies are indeed optimal, and Theorem 7 applies.

In the case of linear shortage costs the derivative policy is given by a control-limit policy. Hence, an ordinary (s, S) policy is optimal.

REFERENCES

1. Deb, R. K. and R. F. Serfozo, "Optimal Control of Batch Service Queues," Advances in Applied Probability, 5 (1973), 340-361.

2. Ignall, E. and P. Kolesar, "Optimal Dispatching of an Infinite-Capacity Shuttle: Control at a Single Terminal," Operations Research, vol. 22, no. 5 (1974), 1008-1024.

3. Kosten, L., Stochastic Theory of Service Systems, New York: Pergamon Press, 1973.

4. Ross, S. M., Applied Probability Models with Optimization Applications, San Francisco: Holden Day, 1970.

5. Sivazlian, B. D., "A Continuous-Review (s, S) Inventory System with Arbitrary Interarrival Distribution between Unit Demand," Operations Research, vol. 22, no. 1 (1974), 65-71.

6. Weiss, H. J., Vehicle Management Systems for High Performance Personal Rapid Transit Systems, Ph.D. Dissertation, Northwestern University, August, 1975.