
Monotonic and Insensitive Optimal Policies for Control of Queues with Undiscounted Costs

Author(s): Shaler Stidham Jr. and Richard R. Weber

Source: *Operations Research*, Vol. 37, No. 4 (Jul. - Aug., 1989), pp. 611-625

Published by: [INFORMS](#)

Stable URL: <http://www.jstor.org/stable/171261>

Accessed: 15/09/2011 09:35

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at
<http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



INFORMS is collaborating with JSTOR to digitize, preserve and extend access to *Operations Research*.

MONOTONIC AND INSENSITIVE OPTIMAL POLICIES FOR CONTROL OF QUEUES WITH UNDISCOUNTED COSTS

SHALER STIDHAM, JR.

University of North Carolina, Chapel Hill, North Carolina

RICHARD R. WEBER

Queens' College, Cambridge, England

(Received November 1983; revisions received July 1984; September 1986; November 1987; April 1988; accepted May 1988)

We consider the problem of controlling the service and/or arrival rates in queues, with the objective of minimizing the total expected cost to reach state zero. We present a unified, simple method for proving that an optimal policy is monotonic in the number of customers in the system. Applications to average-cost minimization over an infinite horizon are given. Both exponential and nonexponential models are considered; the essential characteristic is a left-skip-free transition structure and a nondecreasing (not necessarily convex) holding-cost function. Some of our results are insensitive to service-time distributions.

The literature on optimal control of queues contains many proofs that an optimal control is monotonic in some naturally selected variable. (For surveys see Sobel 1974; Stidham and Prabhu 1974; Crabill, Gross, and Magazine 1977; Stidham 1984, 1985, 1986.) For example, often it is optimal to serve at a faster rate or admit fewer arrivals when more customers are present. The proofs of these *obvious* properties, however, are often disconcertingly long, tedious, and without apparent grounding in intuition. Undeniably there are problems (notably, some involving networks of queues: Weber and Stidham 1987) where careful attention to mathematical rigor—at the occasional expense of intuitive transparency—is not only necessary, but brings ancillary benefits. These benefits include the discovery of additional, unanticipated properties of optimal policies or counterexamples to the *obvious* monotonicity of optimal policies, when not so obvious economic or probabilistic conditions are not satisfied. Still, some of the most basic monotonicity results cry out for proofs that are rigorous, and that exploit common sense and discard superfluous assumptions.

As an example, consider control of the service rate in an M/M/1 queue (Crabill 1974; Sobel 1974, 1982; Lippman 1975; Serfozo 1981; Bengtsson 1982; Jo 1983, among others). Assume, as most authors do, that more customers in the system is worse than fewer, all other things being equal. (That is, the holding cost function is nondecreasing.) Then it seems self-evident

that one would wish to serve at a faster rate when more customers are present, at least when there is no time preference regarding monetary expenditures (that is, no discounting of future benefits and costs). The intuition is simple: use of a given service rate produces larger potential benefits (savings in holding costs) when more customers are present, and it costs the same in both cases. So higher service rates should be reserved for times when the system is more congested. (Discounting can void this argument. We may then wish to postpone the expense of a higher service rate until the future, so that the present value of its cost will be less.) In the literature, however, most of the proofs of this type of result appear to make no use of this intuition. In addition, many proofs depend on what turn out to be superfluous assumptions, such as the convexity of the holding-cost function. Convexity is needed when there is discounting or more than one facility: see Weber and Stidham. Many monotonicity proofs approach the undiscounted problem via a sequence of discounted problems. Exceptions among the papers cited above are Sobel (1982) and Bengtsson.

The purpose of this paper is to provide a unified set of simple, rigorous arguments for the monotonicity of optimal policies, for service-rate control problems among others, in a setting where there is no discounting. The driving engine of our proofs is the intuition expressed above, transformed into a rigorous argument. We are able to provide proofs for many of the

Subject classifications: Dynamic programming, semi-Markov; undiscounted costs, infinite horizon. Queues, optimization: control of queues with undiscounted costs.

service-rate control problems considered in the literature, plus some that have apparently never been treated before. One of our (apparently) new results is the discovery that the optimal service rates are *insensitive* to the service-time distribution (they depend only on the mean) in systems where either there are no arrivals or a last-come, first-served, preemptive-resume queue discipline is in effect. Insensitivity is a phenomenon that has been explored thoroughly in the context of descriptive models of queues, but not, to our knowledge, in control models.

Our approach to each of the control problems studied takes the same general form. We first consider the problem of minimizing the expected total cost until state 0 is reached, from each possible starting state. (Problems in which long-run average cost is to be minimized are converted to this format.) Our formulation relies on ideas from dynamic programming, but with a nonstandard embedding of decision points, based on the left-skip-free transition structure of the models studied: to get from state i to state 0, the system must first pass through state $i - 1$. This property is crucial to our method of analysis. Arrivals are treated as temporary interruptions in the process of moving the system from state i to state $i - 1$ —a treatment that is analogous to the idea of a *superprocess* in the theory of dynamic allocation indices (Gittins 1979, Whittle 1980, 1981).

In its simplest form (for control of the service rate in an M/M/1 queue: cf. Theorem 2), our proof is based on the following intuitive ideas. By the left-skip-free property, the optimal service rate in state i minimizes the expected cost to go from i to $i - 1$. But this problem is probabilistically and economically identical to the problem of going from $i + 1$ to i , except that the holding costs are uniformly larger in the latter problem. Hence, the optimal service rate should be no smaller for state $i + 1$ than for state i . For systems with a state-dependent arrival rate, we need additional conditions; the proofs are less intuitive, but nonconvex holding costs are still allowed.

Section 1 studies exponential models with control of the service rate. First a model with no arrivals is discussed, primarily to introduce our methodology in a simple setting. Then we treat models with arrivals, first with the objective of minimizing total expected cost until the system is empty, then with long-run average cost as the criterion.

The remaining sections of the paper demonstrate how the ideas introduced in Section 1 can be applied to related problems. In most instances we just sketch the proofs because they are straightforward modifica-

tions of those in Section 1. Section 2 discusses exponential models for combined control of service and arrival rates and for control of arrival rates alone. In Section 3, we present some control models for non-exponential systems. The first is an M/G/1 model with a controllable service rate and a last-come, first-served, preemptive-resume discipline. We show that a continuous-time version of the analysis of Section 1 applies to this problem. In fact, the optimal policy is insensitive to the service-time distribution in this model. The other models considered are M/G/1 with selection of the service-time distribution at the start of each service and service-rate control in systems with phase-type service.

1. SERVICE-RATE CONTROL IN EXPONENTIAL QUEUES

1.1. The Model With No Arrivals

We begin with the problem of optimally controlling the service rate in a system with a single exponential server and no arrivals. That is, all jobs to be processed already are present at time zero. The problem may be formally stated as follows.

At time $t = 0$, there are i jobs in the system, which are to be processed one at a time. The service mechanism is memoryless. That is, the probability of a service completion in the interval $(t, t + dt)$, given that the service rate at time t is μ , equals $\mu dt + o(dt)$, independent of the current state or past history of the system. Our goal is to choose the service rate μ at the start of each service in order to minimize the total expected cost to process all jobs (equivalently, the total expected cost until the system reaches state 0). We assume that μ must be chosen from a compact set $A \subset [0, \infty)$. There is a cost of providing service, which is incurred at rate $c(\mu)$ per unit time while the service rate in effect is μ . We assume that $c(\mu)$ is nonnegative, nondecreasing, and continuous on A . (These assumptions about $c(\mu)$ and its domain are not essential. They are introduced here only for ease of exposition. See Remark 1 after Theorem 1.) There is also a holding cost that is incurred at rate $h(j)$ while there are j jobs in the system, $j = 0, 1, \dots$. We assume that $h(0) = 0$ and $h(j) > 0$ for all $j \geq 1$.

The problem can be formulated as a finite-state semi-Markov decision process (Ross 1970, Whittle 1983), with an infinite planning horizon and 0 as an absorbing state. Let $v(i)$ denote the minimum expected total cost incurred until the first visit to state 0, given that the system starts in state $i \geq 0$. The function $v(\cdot)$ is uniquely determined by $v(0) = 0$ and

the recursive optimality equations ($i \geq 1$)

$$v(i) = \min_{\mu \in A} \left\{ \frac{c(\mu) + h(i)}{\mu} + v(i-1) \right\}. \quad (1)$$

The assumptions about $c(\mu)$ and $h(i)$ guarantee that for each $i \geq 1$ the minimum is finite and is attained by some $\mu \in A$, which we shall denote $\mu(i)$. By convention, we shall resolve ties by choosing the largest minimizer.

Let $z(i, j)$ denote the minimum expected cost until the system first enters state j , starting from state i ($j < i$). Since services occur one at a time, starting in state i , the system must visit state $i-1$ on its way to state 0. Thus

$$v(i) = z(i, i-1) + v(i-1). \quad (2)$$

In fact, this result holds for any Markovian decision process that is *skip-free to the left* (Keilson 1965, Wijngaard and Stidham 1986). That is, an instantaneous transition from state i to state j is not possible if $j < i-1$.

It follows from (1) and (2) that

$$z(i, i-1) = \min_{\mu \in A} \frac{c(\mu) + h(i)}{\mu}. \quad (3)$$

Thus $\mu(i)$, the optimal service rate in state i , may be found by minimizing the expected cost to go from i to $i-1$:

$$g(i, \mu) := [c(\mu) + h(i)]/\mu.$$

Theorem 1. *If $h(i)$ is nondecreasing in i , then $\mu(i)$ is nondecreasing in i , $i \geq 1$. In other words, it is optimal to serve at a faster rate when more jobs are in the system.*

Proof. Suppose $h(i)$ is nondecreasing. Let $\mu_2 > \mu_1$. Then

$$\begin{aligned} g(i, \mu_2) - g(i, \mu_1) &= \frac{c(\mu_2)}{\mu_2} - \frac{c(\mu_1)}{\mu_1} - h(i) \left[\frac{1}{\mu_1} - \frac{1}{\mu_2} \right] \end{aligned}$$

which is nonincreasing in i . Letting $\mu_1 = \mu(i)$ and using the fact that $\mu(i)$ is the largest minimizer of $g(i, \mu)$, it follows that

$$\begin{aligned} g(i-1, \mu_2) - g(i-1, \mu(i)) &\geq g(i, \mu_2) - g(i, \mu(i)) > 0 \end{aligned}$$

for all $\mu_2 > \mu(i)$. Hence $\mu(i-1) \leq \mu(i)$.

Remarks

1. The continuity/compactness assumptions about $c(\mu)$ and A were used in the proof of Theorem 1 only to ensure the existence of a largest minimizer of $g(i, \mu)$ for each i . Clearly, the conclusion of Theorem 1 applies as long as this latter condition is satisfied. In particular, the feasible region for μ could be unbounded.

2. **Submodularity.** Our proof of Theorem 1 was based on showing that $g(i, \mu_2) - g(i, \mu_1)$ is nonincreasing in i for $\mu_2 > \mu_1$. A function with this property is said to be *submodular*. Submodular functions play a central role in the theory of lattice programming (Topkis 1978), where they provide a general framework for proving monotonicity of an optimal policy in a variety of control models.

3. **Insensitivity.** Our results clearly do not apply just to memoryless service mechanisms. Suppose that at the beginning of the processing of a job we must choose a service-time distribution from a class of distributions, the members of which are indexed by the reciprocals of their means: $\mu := (E[\text{service time}])^{-1}$. Once a choice of distribution has been made for a particular job, it must remain in effect until the job finishes service. Whatever the choices of service-time distributions, the service times of successive jobs are assumed to be independent. Obviously the minimum total expected cost until all jobs have been served, starting with j jobs present at time t and the service of a job about to begin, does not depend on t nor on the history of states and actions up to time t . Denoting this minimum expected total cost by $v(j)$, $j = 1, 2, \dots$, we see that the optimal value function $v(\cdot)$ again is determined uniquely by the recursive equations (1), with $v(0) = 0$. The total expected cost depends only on the *means* of the various service-time distributions chosen. Moreover, the optimal service rates for this model are *insensitive* to the distributions of service time. The phenomenon of insensitivity has been studied extensively in the context of descriptive models for queues but it has been rarely, if ever, encountered in control models.

4. Viewed in the context of the original problem—starting with i jobs present at time 0, choose service rates for each job so as to minimize the expected total cost incurred until all jobs have been processed—our result says that we should process at a faster rate at first, while more jobs are present, and at a slower rate later on, when fewer jobs are in the system. This property is qualitatively similar to the optimality of the *shortest-processing-time* (SPT) discipline for a sequence of jobs on a single processor (Conway,

Maxwell and Miller 1967). Here, instead of rearranging the jobs so that the shortest one comes first, we are able to *select* the processing times (or at least their distributions) in such a way that the shortest one (in mean) comes first.

5. A *classical* proof of the optimality of SPT (Conway et al.) proceeds by looking at an arbitrary sequence that violates SPT and showing that the total cost can be reduced by making a pairwise switch of any two jobs whose sequence violates SPT. In a parallel way, in our model, one can construct an alternate proof that the optimal service rates are monotonic, based on pairwise switches of any two service rates for adjacent states, say $i + 1$ and i , that violate monotonicity. (See Section 3.2 for an example of an argument based on pairwise switches.) More complex pairwise-switch arguments play an important role in the method of *dynamic allocation indices* (often called *Gittins indices*) for solving a variety of dynamic control problems, including multiarmed-bandit problems and priority selection in queueing systems (Gittins; Whittle 1980, 1981).

1.2. The Model With Arrivals

Now we add arrivals to the previous model. Assume that new jobs arrive to the system according to a state-dependent Poisson process with mean rate λ_i , $i \geq 0$. In all other respects the model is assumed to be the same as already described in the previous subsection.

We formulate the problem of minimizing the total expected cost to go from state i to state 0 as an infinite horizon, semi-Markov decision process, with state 0 as an absorbing state. We observe the system and choose a service rate at each change of state (arrival or service-completion epoch). Let $v(i)$ denote the minimal expected total cost to move the system from state i to state 0, $i = 1, 2, \dots$ ($v(0) = 0$). Since the costs between observation points are nonnegative, the problem and its optimal cost function v are well defined (Strauch 1966, Schäl 1975) and the $v(i)$ satisfy the optimality equation ($i \geq 1$)

$$v(i) = \min_{\mu \in \mathcal{I}} \{(\lambda_i + \mu)^{-1} [c(\mu) + h(i) + \lambda_i v(i + 1) + \mu v(i - 1)]\} \quad (4)$$

with $v(0) = 0$. Moreover, a stationary policy that uses a service rate μ that minimizes the right-hand side of (4) whenever the system is in state i is optimal. Once again, we resolve ties by choosing the largest minimizer, denoted $\mu(i)$.

We shall find it more convenient to work with a transformed optimality equation that has the same form as (1). Observe that $v(i)$ satisfies (4) for $i \geq 1$ if

and only if for all μ

$$v(i)(\lambda_i + \mu) \leq c(\mu) + h(i) + \lambda_i v(i + 1) + \mu v(i - 1)$$

or equivalently

$$v(i) \leq [c(\mu) + h(i) + \lambda_i(v(i + 1) - v(i))]/\mu + v(i - 1)$$

with equality for a value of μ that achieves the minimum in (4). Thus, (4) is equivalent to the optimality equation ($i \geq 1$)

$$v(i) = \min_{\mu \in \mathcal{I}} \{[c(\mu) + h(i) + \lambda_i(v(i + 1) - v(i))]/\mu + v(i - 1)\} \quad (5)$$

The left-skip-free property of state transitions again ensures that

$$v(i) = z(i, i - 1) + v(i - 1) \quad (6)$$

where, as before, $z(i, j)$ is defined as the minimum total expected cost to go from state i to state $j < i$. Combining (5) and (6) we see that $z(i, i - 1)$ satisfies

$$z(i, i - 1) = \min_{\mu \in \mathcal{I}} \left\{ \frac{c(\mu) + h(i) + \lambda_i z(i + 1, i)}{\mu} \right\} \quad (7)$$

The following intuitive interpretation of (7) may be instructive. Consider the problem of minimizing the total expected cost to go from state i to state $i - 1$. Because of the memoryless property of the exponential service-time distribution and the state-dependent Poisson arrival process, an optimal choice of the service rate μ clearly depends only on the current state i and not on the current time t nor on the past history of the system. Whatever the choice of μ , an arrival may occur before a transition from i to $i - 1$. When this happens, the system moves to state $i + 1$ and spends a certain amount of time in states $j \geq i + 1$ before returning to state i , at which point the service mechanism begins to serve again at rate μ . Each excursion into states $j \geq i + 1$ may be regarded as a temporary interruption to the process of moving the system from state i to state $i - 1$.

While the state is i and the service rate is μ , the system incurs cost at rate $c(\mu) + h(i)$. The expected cost while in state i before a transition to $i - 1$ is thus $[c(\mu) + h(i)]/\mu$. Excursions into states $j \geq i + 1$ occur at rate λ_i . The expected number of excursions before a transition to $i - 1$ is, therefore, λ_i/μ . The expected total cost during each excursion is $z(i + 1, i)$. Thus, the expected total cost of going from state i to $i - 1$, with service rate μ in effect, is $[c(\mu) + h(i) + \lambda_i z(i + 1, i)]/\mu$, the minimand in (7).

It follows from (6) and (7) that the problem with external arrivals takes exactly the same form as the problem with no arrivals as embodied by (2) and (3), but with a modified holding cost function

$$\tilde{h}(i) := h(i) + \lambda_i z(i + 1, i). \tag{8}$$

Thus, as an immediate corollary to Theorem 1, we have the following lemma.

Lemma 1. *For the problem with external arrivals, if $\tilde{h}(i)$ is finite and nondecreasing in i , where $\tilde{h}(i)$ is defined by (8), then the optimal service rate $\mu(i)$ is nondecreasing in $i \geq 0$.*

Poisson Arrival Process

Our first application of this lemma yields a simple proof of the monotonicity of the optimal service rate in the special case, $\lambda_i \equiv \lambda$ (an ordinary Poisson arrival process).

Theorem 2. *Suppose that the arrivals come from a Poisson process with a mean rate λ , and that $h(i)$ is nondecreasing in i . Then the optimal service rate $\mu(i)$ is nondecreasing in i .*

Proof. By Lemma 1 it suffices to show that $z(i + 1, i)$ is nondecreasing in i . Since the arrival rate is constant, $z(i + 1, i)$ and $z(i, i - 1)$ may both be regarded as solutions to the same problem, namely that of minimizing the total expected cost to go from state 1 to state 0, but with different holding cost functions: in the former case, $h(i + \cdot)$, and in the latter case, $h(i - 1 + \cdot)$. Since $h(\cdot)$ is nondecreasing, $h(i + \cdot)$ is uniformly at least as large as $h(i - 1 + \cdot)$. Hence, the minimum expected total cost to go from 1 to 0 is at least as large in the former case as in the latter, from which it follows that $z(i + 1, i) \geq z(i, i - 1)$.

The key idea in the proof is a generalization of the familiar insight that an M/M/1 queue exhibits probabilistically identical behavior in going from state $i + 1$ to state i as in going from state i to state $i - 1$, and that in both cases the behavior is probabilistically identical to that over an ordinary busy period (going from state 1 to state 0). The generalization permits state-dependent service rates, provided that, while going from i to $i - 1$, we use in each state $j \geq i$, the rate we would use in state $j + 1$ while going from $i + 1$ to i . It is clear that this idea breaks down when the arrival rates are allowed to be state dependent, in which case, a more intricate argument and additional conditions will be needed.

State-Dependent Arrival Rates

We give a set of sufficient conditions for the monotonicity of the optimal service rate in the general case of state-dependent arrival rates. The conditions are in the same spirit as those of Crabill for a special case of our problem. As in Section 1.1, we assume that A is a compact set, with $\bar{\mu} = \max\{\mu: \mu \in A\} < \infty$.

$$(C1) \quad \bar{\lambda} := \sup_{i \geq 0} \lambda_i < \infty \quad \text{and} \quad \sum_{i=1}^{\infty} h(i) \left(\frac{\bar{\lambda}}{\bar{\mu}}\right)^i < \infty.$$

$$(C2) \quad \gamma_0 := \sup \left\{ \frac{c(\bar{\mu}) - c(\mu)}{\bar{\mu} - \mu} : \mu \in A, \mu < \bar{\mu} \right\} < \infty.$$

$$(C2') \quad \gamma_1 := \sup \left\{ \frac{c(\mu_2) - c(\mu_1)}{\mu_2 - \mu_1} : \mu_1, \mu_2 \in A, \mu_1 \neq \mu_2 \right\} < \infty.$$

$$(C3) \quad h(i) \rightarrow \infty \quad \text{as} \quad i \rightarrow \infty.$$

$$(C3') \quad h(i) \text{ converges monotonically to } \infty \quad \text{as} \quad i \rightarrow \infty.$$

$$(C4) \quad \text{for all } i \geq 1, \quad h(i) - h(i - 1) \geq (\lambda_{i-1} - \lambda_i)\gamma_1.$$

Conditions C2 and C2' are trivially satisfied when A consists of a finite set of service rates. If A consists of a closed interval, it suffices for C2 that $c(\mu)$ be continuous on A and that a left derivative exists at $\bar{\mu}$. Likewise, if A is a closed interval, it suffices for C2' that $c(\mu)$ be Lipschitz continuous on A , e.g., when $c(\mu)$ is continuously differentiable on A . (None of these conditions is, in fact, very restrictive. See Remark 6 below.)

Lemma 2. *Assume Condition C1. Then $z(i, i - 1) < \infty$, and hence, $v(i) < \infty$, for all $i \geq 1$.*

Proof. It suffices to show that $z_f(i, i - 1) < \infty$, where $z_f(i, i - 1)$ is the total expected cost to go from i to $i - 1$ under the full-service policy: the policy that uses service rate $\bar{\mu}$ in all states $i \geq 1$. Now

$$z_f(i, i - 1) = t_f(i, i - 1)c(\bar{\mu}) + h_f(i, i - 1)$$

where $t_f(i, i - 1)$ and $h_f(i, i - 1)$ are, respectively, the total expected time and the total expected holding cost incurred during the passage from state i to state $i - 1$, under the full-service policy. It is easy to see by a stochastic-dominance argument that $t_f(i, i - 1)$ is bounded above by the expected length of a busy period in an M/M/1 queue with an arrival rate $\bar{\lambda}$ and a service rate $\bar{\mu}$, and therefore is finite. Similarly, using the facts that the expected time spent in state j during a busy period in this M/M/1 queue equals

$(1/\bar{\lambda})(\bar{\lambda}/\bar{\mu})^j$ and that $h(\cdot)$ is nondecreasing, we conclude that

$$\begin{aligned} \bar{\lambda}h_j(i, i-1) &\leq \sum_{j=1}^{\infty} h(i-1+j) \left(\frac{\bar{\lambda}}{\bar{\mu}}\right)^j \\ &= \left(\frac{\bar{\mu}}{\bar{\lambda}}\right)^{i-1} \sum_{j=i}^{\infty} h(j) \left(\frac{\bar{\lambda}}{\bar{\mu}}\right)^j < \infty \end{aligned}$$

by Condition C1.

Lemma 3. *Assume Condition C1. Then, for $\mu, \mu' \in A, \mu < \mu(i) < \mu'$*

$$\frac{c(\mu(i)) - c(\mu)}{\mu(i) - \mu} \leq z(i, i-1) \leq \frac{c(\mu') - c(\mu(i))}{\mu' - \mu(i)}. \quad (9)$$

Proof. It follows from Lemma 2 that $z(i, i-1) < \infty$. Then, from the definition of $\mu(i)$, for all $\mu \in A$,

$$\begin{aligned} z(i, i-1) &= [c(\mu(i)) + \tilde{h}(i)]/\mu(i) \leq [c(\mu) + \tilde{h}(i)]/\mu. \quad (10) \end{aligned}$$

A little algebra shows that (10) holds for $0 < \mu < \mu(i)$ if and only if

$$\frac{c(\mu(i)) - c(\mu)}{\mu(i) - \mu} \leq \frac{c(\mu(i)) + \tilde{h}(i)}{\mu(i)} = z(i, i-1).$$

The proof of the right-hand inequality of (9) is analogous.

Lemma 4. *Assume Conditions C1, C2, and C3. Then for all sufficiently large $i, \mu(i) = \bar{\mu}$.*

Proof. In view of C2 and C3, we have $h(i) > \bar{\mu}\gamma_0$ for all i sufficiently large. For all such states $i, \mu(i) = \bar{\mu}$, since, otherwise, $\gamma_0 < [c(\mu(i)) + \tilde{h}(i)]/\mu(i) \leq [c(\bar{\mu}) - c(\mu(i))]/(\bar{\mu} - \mu(i)) \leq \gamma_0$, in view of Lemma 3.

We are now ready to prove monotonicity of $\mu(i)$.

Theorem 3. *Assume Conditions C1, C2', C3', and C4. Then $\mu(i)$ is nondecreasing in $i \geq 1$.*

Proof. Let $k := \max\{i: \mu(i) < \bar{\mu}\}$. (Lemma 4 implies $k < \infty$.) It suffices to show that $\mu(i) \geq \mu(i-1)$, or in view of Lemma 1, that $\tilde{h}(i) \geq \tilde{h}(i-1)$, for all $i \leq k$. The proof proceeds by downward induction on i .

We first observe that

$$h(i) - h(i-1) \geq (\lambda_{i-1} - \lambda_i) \frac{c(\mu_2) - c(\mu_1)}{\mu_2 - \mu_1} \quad (11)$$

for all $\mu_1, \mu_2 \in A, \mu_2 > \mu_1$. If $\lambda_{i-1} > \lambda_i$, then (11) follows from Conditions C2' and C4. Otherwise, it follows from C3'.

Let $i < k$ and suppose that $\tilde{h}(i+1) \geq \tilde{h}(i)$. Then $\mu(i+1) \geq \mu(i)$. If $\mu(i+1) > \mu(i)$, Lemma 3 and (11) imply

$$\begin{aligned} \tilde{h}(i) &= h(i) + \lambda_i z(i+1, i) \\ &\geq h(i) + \frac{\lambda_i [c(\mu(i+1)) - c(\mu(i))]}{\mu(i+1) - \mu(i)} \\ &\geq h(i-1) + \frac{\lambda_{i-1} [c(\mu(i+1)) - c(\mu(i))]}{\mu(i+1) - \mu(i)} \\ &\geq h(i-1) + \lambda_{i-1} z(i, i-1) = \tilde{h}(i-1). \quad (12) \end{aligned}$$

On the other hand, if $\mu(i+1) = \mu(i) = \mu$ (say), then

$$\begin{aligned} z(i+1, i) &= \frac{c(\mu) + \tilde{h}(i+1)}{\mu} \\ &\geq \frac{c(\mu) + \tilde{h}(i)}{\mu} = z(i, i-1). \quad (13) \end{aligned}$$

Since $\mu < \bar{\mu}$ (by the definition of k and the assumption that $i < k$), it follows from (11) that

$$h(i) - h(i-1) \geq (\lambda_{i-1} - \lambda_i) \frac{c(\bar{\mu}) - c(\mu)}{\bar{\mu} - \mu}$$

so that

$$h(i) - h(i-1) \geq (\lambda_{i-1} - \lambda_i) z(i+1, i) \quad (14)$$

if $\lambda_{i-1} > \lambda_i$. Otherwise, (14) follows from Condition C3'. Together (14) and (13) imply that

$$\begin{aligned} \tilde{h}(i) &= h(i) + \lambda_i z(i+1, i) \\ &\geq h(i-1) + \lambda_{i-1} z(i+1, i) \\ &\geq h(i-1) + \lambda_{i-1} z(i, i-1) = \tilde{h}(i-1). \end{aligned}$$

This completes the inductive step. To start the induction, we must verify that $\tilde{h}(k) \geq \tilde{h}(k-1)$. Since $\bar{\mu} = \mu(k+1) > \mu(k)$ by the definition of k , the argument is formally the same as used in the proof of (12).

Remarks

6. Consider the case where $A = [0, \bar{\mu}]$ and the service-cost function $c(\mu)$ is convex on A . Then Conditions C2 and C2' coincide and hold if and only if $(d^-/d\mu)c(\bar{\mu}) < \infty$. In fact, any nonnegative, nondecreasing service-cost function $c(\mu)$ defined on $[0, \bar{\mu}]$ can be replaced without loss of optimality by its lower convex envelope, $\underline{c}(\mu)$. Moreover, cases where not all service rates $\mu \in [0, \bar{\mu}]$ are feasible also can be accommodated by our model by replacing $c(\mu)$ by its lower convex envelope, $\underline{c}(\mu)$, defined on the set \underline{A} , the convex

hull of A . To see why, let

$$\alpha := [c(\mu(i)) + \tilde{h}(i)]/\mu(i) = \min_{\mu \in A} \{[c(\mu) + \tilde{h}(i)]/\mu\}.$$

Lemma 3 implies that $c(\mu) \geq a(\mu) := c(\mu(i)) + \alpha(\mu - \mu(i))$, for all $\mu \in A$. Hence $\underline{c}(\mu) \geq a(\mu)$ for all $\mu \in \underline{A}$, since $\underline{c}(\mu)$ is by definition the supremum at μ over all affine functions that bound $c(\mu)$ from below on the set A . Therefore

$$\frac{\underline{c}(\mu) + \tilde{h}(i)}{\mu} \geq \frac{a(\mu) + \tilde{h}(i)}{\mu} = \frac{c(\mu(i)) + \tilde{h}(i)}{\mu(i)}$$

for all $\mu \in \underline{A}$, from which it follows that

$$\min_{\mu \in \underline{A}} \left\{ \frac{c(\mu) + \tilde{h}(i)}{\mu} \right\} = \min_{\mu \in \underline{A}} \left\{ \frac{\underline{c}(\mu) + \tilde{h}(i)}{\mu} \right\}$$

and $\mu(i)$ attains both minima.

As an example, consider the case where there are only a finite number of feasible rates, $0 < \mu_1 < \mu_2 < \dots < \mu_n \leq \bar{\mu}$, so that $\underline{c}(\mu)$ is piecewise linear. It follows that a service rate μ_j will never be used in an optimal policy if there exist rates μ_i and μ_k such that $\mu_i < \mu_j < \mu_k$ and $[c(\mu_j) - c(\mu_i)]/[\mu_j - \mu_i] > [c(\mu_k) - c(\mu_j)]/[\mu_k - \mu_j]$, since in this case, $c(\mu_j) > \underline{c}(\mu_j)$ (cf. Crabill).

It follows from these observations that Condition C2' is actually stronger than necessary; it suffices that C2' holds for the lower convex envelope, $\underline{c}(\mu)$, which is true if and only if $(d^-/d\mu)\underline{c}(\bar{\mu}) < \infty$.

7. Without a condition like C4, the optimal service rate with state-dependent arrivals will not be monotonic, in general. For example, if the arrival rate decreases when more customers are present, then there may be an incentive to move to a higher state and lose some arrivals, as this may lead to lower congestion costs in the future, even with a nondecreasing holding-cost function. Whether or not this is the case depends on the relative values of $h(\cdot)$, $c(\cdot)$, and $\{\lambda_i\}$. A simple counterexample in which C4 does not hold is as follows. Suppose there are two feasible service rates, $\mu = 1$ and $\mu = 2$, with $c(1) = 1$, $c(2) = 4$; $\lambda_1 = 1$, $\lambda_2 = 0$; and $h(1) = h(2) = 1$. In this example, we have $\tilde{h}(2) = h(2) = 1$, and $z(2, 1) = \min[c(\mu) + 1]/\mu = [c(1) + 1]/1 = 1 + 1 = 2$, so that $\mu(2) = 1$. But this implies that $\tilde{h}(1) = h(1) + z(2, 1) = 1 + 2 = 3$, so that $z(1, 0) = \min[c(\mu) + 3]/\mu = [c(2) + 3]/2 = 7/2$, so that $\mu(1) = 2 > 1 = \mu(2)$.

Note that Condition C4 is trivially satisfied with any nondecreasing $h(\cdot)$ when the arrival rates λ_i are nondecreasing in i .

1.3. Average-Cost Criterion

We will consider a service system operating over an infinite horizon. The model is the same as in the

previous subsection, but now the objective is to minimize the long-run average cost per unit time from each starting state i , $i \geq 0$. Throughout this subsection we shall assume that $\lambda_i > 0$ for all $i \geq 0$. (This assumption can be relaxed. See Remark 8 below.)

We shall establish monotonicity of an optimal policy by showing that the average-cost problem is equivalent to a problem in which the objective is to minimize the total expected g -revised cost until the next visit to state 0, in the sense that an optimal policy for the latter problem is average-cost optimal as well, when $g = g^* :=$ the minimal long-run average cost. Here g -revised cost means the cost incurred by a system in which a constant g is subtracted from the cost rate at each time point. This is equivalent to replacing the holding-cost rate $h(i)$ by $h(i) - g$ in each state $i \geq 0$. We then can apply the results of Section 1.2 to conclude that a monotonic policy is optimal for the average-cost problem.

To prove the equivalence of the average-cost problem and the total g^* -revised cost problem, we shall use an argument based on the renewal-reward theorem, which makes it possible to express the average cost of a stationary policy as the ratio of the expected total cost to the expected total time elapsed between two successive visits to state 0. But first, we must establish that we can restrict attention to stationary policies without loss of average-cost optimality.

Lemma 5. *Assume Conditions C1 and C3. Then there exists a stationary policy that minimizes the long-run average expected cost per unit time from each starting state $i \geq 0$. Its long-run average expected cost g^* is independent of the starting state and $g^* < \infty$.*

Proof. The desired result will follow if we can show that conditions a-g of Weber and Stidham hold. Conditions a-d and f are obviously satisfied. Condition C1 implies e: it is possible to go from any state to any other state with finite expected cost (cf. proof of Lemma 2). Finally, Condition C3 implies g: there are only a finite number of states in which the one-stage cost $[c(\mu) + h(i)]/(\lambda_i + \mu)$ does not exceed the average cost from a fixed policy, say the full-service policy. (Using the notation of Lemma 2, one concludes from the renewal-reward theorem that the average cost of the full-service policy is given by $\{c(\mu)t_f(1, 0) + h_f(1, 0)\}/[t_f(1, 0) + \lambda_0^{-1}] < \infty$.)

Remark

8. An alternate proof of this theorem uses the conditions of Sennott (1987), which are similar to those of Weber and Stidham.

It remains to show that a monotonic stationary optimal policy exists for the average-cost problem. As indicated before, we shall do this by showing that the average-cost problem can be solved by minimizing the expected total g^* -revised cost until the next visit to state 0, and then applying the results of Section 1.2. Two technical issues complicate this approach. First, we do not know a priori that state 0 is positive recurrent under a stationary average-cost optimal policy. Second, the one-stage costs in the g^* -revised cost problem are not necessarily nonnegative, as assumed in Section 1.2.

For these reasons, we shall begin by formulating a slightly more general version of the problem studied in Section 1.2. To be specific, consider the service-rate control problem of Section 1.2 with the following alterations: 1) a constant g is subtracted from the holding cost $h(i)$ in each state $i \geq 0$; 2) the objective is to minimize the expected total (g -revised) cost until the next visit to a given state $m \geq 0$, starting from each state $i \geq 0$; and 3) the set of feasible service rates $A(i)$ in state i is given by $A(0) = \{0\}$, $A(i) = A$ for $0 \leq i \leq m$, and $A(i) = A \cap [\underline{\mu}, \infty)$ for $i > m$, where $\underline{\mu} \geq 0$. (In other words, we require the server to operate at least at the rate $\underline{\mu}$ in states $i > m$.) We call this Problem $\mathbf{P}(g, m, \underline{\mu})$. Note that the problem in Section 1.2 is a special case of Problem $\mathbf{P}(g, m, \underline{\mu})$, in which $g = 0$, $m = 0$, and $\underline{\mu} = 0$.

Problem $\mathbf{P}(g, m, \underline{\mu})$ can be formulated as a SMDP if we replace the transitions into state m by transitions into an absorbing, costless state. The optimal value function for this SMDP is $z^*(\cdot, m)$, where $z^*(i, m)$ denotes the minimal expected total g -revised cost until the next visit to state m , starting from state i , $i \geq 0$.

Lemma 6. *Assume Condition C3. Consider Problem $\mathbf{P}(g, m, \underline{\mu})$. Let $k := \max\{i: h(i) < g\}$ and assume that $\underline{\mu} > 0$ if $m < k$. Then the SMDP formulation of $\mathbf{P}(g, m, \underline{\mu})$ and its associated optimal value function $z^*(\cdot, m)$ are well defined. Moreover, $z^*(\cdot, m)$ satisfies the optimality equations ($i \geq 0$)*

$$\begin{aligned} z^*(i, m) &= \min_{\mu \in A(i)} \{(\lambda_i + \mu)^{-1} [c(\mu) + h(i) \\ &\quad - g + \lambda_i z^*(i + 1, m) + \mu z^*(i - 1, m)]\} \quad (15) \end{aligned}$$

with $z^*(i + 1, m)$ ($z^*(i - 1, m)$) replaced by 0 when $i = m - 1$ ($i = m + 1$). A stationary policy is optimal for Problem $\mathbf{P}(g, m, \underline{\mu})$ if, in each state $i \geq 0$, it chooses a service rate $\mu(i)$ that minimizes the bracketed expression on the right-hand side of (15).

Proof. Since $h(i) \rightarrow \infty$ as $i \rightarrow \infty$ (Condition C3), for any fixed g we have $k < \infty$. Then for all $i > k$ the g -revised cost per unit time while in state i is nonnegative, while for states $i \leq k$, the g -revised cost per unit time is bounded below by $-g$. Under any policy, starting in a state $0 \leq i < m$, the expected time spent in states $[0, m)$ before the next visit to state m is uniformly bounded above by the expected time until the next visit to state m under the full-service policy f , which is obviously finite. Similarly, if $m < k$, then, starting in a state $i > m$, the expected time spent in states $(m, k]$ before the next visit to state m is uniformly bounded above by the expected time spent in states $(m, k]$ under the policy that uses service rate $\underline{\mu} > 0$ in all states (m, ∞) , which is also finite. A similar argument shows that the expected time spent in states $[0, k]$, starting in state m , is uniformly bounded above. It follows that General Assumption (G) of Schäl is satisfied. Hence, Problem $\mathbf{P}(g, m, \underline{\mu})$ and its associated optimal value functions are well defined. The rest of the theorem follows from standard SMDP theory (cf. Schäl).

Lemma 7. *Assume Conditions C1, C2', C3' and C4. Consider Problem $\mathbf{P}(g, m, \underline{\mu})$, with $\underline{\mu} > 0$ if $m < k$. For each $i > 0$, let $\mu(i)$ be the largest minimizer of the right-hand side of (15). Then $\mu(i)$ is nondecreasing in $i > m$, that is, a monotonic policy is optimal for Problem $\mathbf{P}(g, m, \underline{\mu})$ for states $i > m$.*

Proof. As in the problem in Section 1.2, a stationary optimal policy for states $i > m$ in Problem $\mathbf{P}(g, m, \underline{\mu})$ can be found by choosing the largest minimizer in the following optimality equation, which is equivalent to (15) for $i > m$:

$$\begin{aligned} z^*(i, i - 1) &= \min_{\mu \in A(i)} \left\{ \frac{[c(\mu) + h(i) - g + \lambda_i z^*(i + 1, i)]}{\mu} \right\}. \quad (16) \end{aligned}$$

Essentially the same arguments as used in Section 1.2 imply that the optimal service rates for (16) are monotonic: $\mu(i + 1) \geq \mu(i)$, $i > m$. In particular, neither the presence of a constant lower bound on μ nor subtraction of the constant g from $h(i)$ affect the inductive proof of Theorem 3.

We are now ready to prove the main result of this section.

Theorem 4. *Assume Conditions C1, C2', C3' and C4. There exists a monotonic stationary policy that is average-cost optimal. That is, $\mu^*(i)$ is nondecreasing*

in $i \geq 1$, where $\mu^*(i)$ is an average-cost optimal service rate to be used whenever the system is in state i . Moreover, $\mu^*(i) > 0$ for all $i > 0$ and $\mu^*(i) = \bar{\mu}$ for all sufficiently large i .

Proof. Lemma 5 implies that we can restrict our attention to stationary policies without loss of average-cost optimality. Moreover, a policy that has no positive recurrent states cannot be average-cost optimal, as the following argument shows.

Consider a fixed stationary policy under which all states $i \geq 0$ are transient. It follows from the theory of birth-death processes (Karlin and Taylor 1975) that $p_i = 0$ for every state $i \geq 0$, where p_i is the long-run fraction of time spent in state i . Since $h(i)$ is nondecreasing (Condition C3'), the average cost over the interval $[0, t]$ is bounded below by

$$h(i) \cdot (\text{fraction of } [0, t] \text{ spent in states } j \geq i) \\ \rightarrow h(i) \cdot \left[1 - \sum_{j=0}^{i-1} p_j \right] = h(i), \quad \text{as } t \rightarrow \infty$$

for every i . It follows that the long-run average cost under such a policy is infinite. But Condition C1 implies that there exists a policy (namely, the full-service policy) with finite long-run average cost. Hence, a stationary average-cost optimal policy must have at least one positive recurrent state m .

Let π be a stationary policy and suppose that state $i \geq m$ is positive recurrent under π , where $m \geq 0$. Let $z_\pi(i, j)$ and $t_\pi(i, j)$ denote the total expected cost and time, respectively, until the next visit to state j , starting from state i and following policy π . Since $\lambda_i > 0$ for all $i \geq 0$, it follows that all states $i \geq m$ are positive recurrent and reach m , so that $t_\pi(i, m) < \infty$, for all $i \geq 0$. The renewal-reward theorem (Ross), therefore, implies that the long-run average cost under policy π is independent of the starting state i and is given by

$$g_\pi := z_\pi(m, m)/t_\pi(m, m).$$

Let π^* be an average-cost optimal policy and let m be positive recurrent under π^* , with $m \geq k = \max\{i: h(i) < g^*\}$. For $i = 0, 1, \dots$, and $\mu \geq 0$, let $\Pi(i, \mu)$ denote the class of all stationary policies under which state i is positive recurrent and which use a service rate no smaller than μ in all states $j > i$. Then, without loss of average-cost optimality, we can restrict attention to policies in $\Pi(m, \mu)$, with $\mu = 0$. Moreover

$$z_\pi(m, m) - g^*t_\pi(m, m) \geq 0$$

for all policies $\pi \in \Pi(m, 0)$ with equality for the average-cost optimal policy π^* . Therefore, an average-

cost optimal policy can be found by minimizing $z_\pi(m, m) - g^*t_\pi(m, m)$ over all policies in $\Pi(m, 0)$.

For any g , the quantity $z_\pi(i, m) - gt_\pi(i, m)$ equals the total expected g -revised cost incurred until the next visit to state m , starting in state i and following policy π . The problem of minimizing this quantity over all policies in $\Pi(m, 0)$ is just Problem $\mathbf{P}(g, m, 0)$. Thus, an average-cost optimal policy can be found by solving Problem $\mathbf{P}(g, m, \underline{\mu})$ with $\underline{\mu} = 0$ and $g = g^*$. It follows from Lemma 7 that this policy has monotonic service rates in all states $i > m$.

Our next step is to show that an average-cost optimal policy uses a positive service rate in all states $i > 0$, so that state 0 is positive recurrent. For this result we need Condition C4, in addition to C1, C2', and C3'. First we show that $\mu(m) > 0$, where $\mu(m)$ is the optimal service rate in state m for Problem $\mathbf{P}(g, m, \underline{\mu})$ with $\underline{\mu} = 0$ and $g = g^*$. Since $\mu(m)$ is also average-cost optimal in state m (see above), it follows that state $m - 1$ is positive recurrent under an average-cost optimal policy. Continuing in this fashion, one concludes that state 0 is positive recurrent under an average-cost optimal policy.

Consider the optimality equation (15) for Problem $\mathbf{P}(g, m, \underline{\mu})$ with $\underline{\mu} = 0$ and $g = g^*$. Our previous discussion implies that

$$z^{s*}(m, m) = z_{\pi^*}(m, m) - g^*t_{\pi^*}(m, m) = 0.$$

Suppose $\mu(m) = 0$. Then (15) for $i = m$ and $g = g^*$ implies

$$0 = [h(m) - g^*]/\lambda_m + z^{s*}(m + 1, m) \\ < [c(\mu) + h(m) - g^* + \lambda_m z^{s*}(m + 1, m) \\ + \mu z^{s*}(m - 1, m)]/(\lambda_m + \mu) \quad (17)$$

for all $\mu > 0$. (The strict inequality follows from our convention of always selecting the largest minimizer of the right-hand side of the optimality equation.) We can use the equality in (17) to eliminate g^* from the right-hand side of the inequality. Letting $\mu = \mu(m + 1) > 0$, rearranging terms, and using the fact that $z^{s*}(m - 1, m) \leq [h(m - 1) - g^*]/\lambda_{m-1}$, we obtain

$$\frac{\lambda_{m-1}c(\mu(m + 1))}{\mu(m + 1)} \\ > -h(m - 1) + h(m) + \lambda_m z^{s*}(m + 1, m). \quad (18)$$

But Lemma 3 implies $z^{s*}(m + 1, m) = [c(\mu(m + 1)) + h(m + 1) - g^*]/\mu(m + 1) \geq c(\mu(m + 1))/\mu(m + 1)$, which, combined with (18), yields

$$h(m) - h(m - 1) \\ < (\lambda_{m-1} - \lambda_m)c(\mu(m + 1))/\mu(m + 1)$$

a contradiction of Condition C4. Thus $\mu(m) > 0$, and we have an average-cost optimal policy in which $\mu^*(m) > 0$, and $\mu^*(i)$ is positive and nondecreasing in $i > m$. It follows that $\underline{\mu} := \min\{\mu^*(i)\} = \min\{\mu^*(m), \mu^*(m+1)\} > 0$, so that all states $i \geq m-1$ are positive recurrent. Another application of the renewal-reward argument shows that an average-cost optimal policy can be found by solving Problem $\mathbf{P}(g^*, m-1, \underline{\mu})$. But for this problem, Condition C4 will imply that $\mu(m-1) > 0$, by a repetition of the above argument by contradiction, with m replaced by $m-1$. It follows by downward induction that there exists an average-cost optimal policy in which $\mu^*(i) > 0$ for all $i > 0$, and $\mu^*(i)$ is nondecreasing in i . Finally, it follows from Lemma 4 that $\mu^*(i) = \bar{\mu}$ for sufficiently large i . This completes the proof of Theorem 4.

Remark

9. The Assumption that $\lambda_i > 0$ for all $i \geq 0$ was made for simplicity of exposition. If $\lambda_n = 0$ for some $n > 0$, then the problem reduces to one with a finite number of states $i = 0, 1, \dots, n$. The proofs of this section all go through with minor modifications; in some cases, they are simplified considerably.

2. OTHER EXPONENTIAL CONTROL MODELS

In this section, we use the methods of Section 1, with minor modifications, to establish the monotonicity of optimal policies for combined control of arrival and service rates and control of arrival rates alone in exponential systems. We shall only study the problem of minimizing the expected cost until the system becomes empty. The extension to average-cost minimization proceeds along the same lines as in Section 1.

2.1. Combined Control of Arrival and Service Rates

In this model we are free to choose both λ and μ at each change of state. The feasible pairs (λ, μ) are indexed by a real number a which must belong to a compact set A . The corresponding arrival and service rates are denoted $\lambda(a)$ and $\mu(a)$, respectively. We assume that $\mu(a)$ is continuous and nondecreasing in a and that $\lambda(a)/\mu(a)$ is continuous and nonincreasing in a . There is a cost per unit time $c(a)$ which we assume is nonnegative, nondecreasing, and continuous in a . There is a positive holding cost that is incurred at rate $h(i)$ per unit time while there are i customers in the system, $i \geq 1$. We assume that $h(\cdot)$ is nondecreasing.

With $v(i)$ and $z(i, i-1)$ defined as in Section 1.2, we have $v(0) = 0$ and, for $i \geq 1$, $v(i) = z(i, i-1) + v(i-1)$ and

$$z(i, i-1) = \min_{a \in A} \left\{ \frac{[c(a) + h(i) + \lambda(a)z(i+1, i)]}{\mu(a)} \right\}. \quad (19)$$

Once again, in order to show that an optimal policy is monotonic (in this case, that an optimal action $a(i)$ is nondecreasing in i), we need to show that $g(i, a)$, the right-hand side of (19), is submodular in i and a . For $a_2 > a_1$, we have

$$\begin{aligned} g(i, a_2) - g(i, a_1) &= \frac{c(a_2)}{\mu(a_2)} - \frac{c(a_1)}{\mu(a_1)} + h(i) \left[\frac{1}{\mu(a_2)} - \frac{1}{\mu(a_1)} \right] \\ &\quad + \left[\frac{\lambda(a_2)}{\mu(a_2)} - \frac{\lambda(a_1)}{\mu(a_1)} \right] z(i+1, i) \end{aligned}$$

which is nonincreasing in i if $z(i+1, i)$ is nondecreasing in i . But the argument for this is essentially the same as in the proof of Theorem 2.

Note that the requirement that $c(a)$ be nonnegative may be relaxed if $\mu(a) \geq \underline{\mu} > 0$ for all $a \in A$, since this assumption by itself guarantees that the optimization problem (19) is well defined and that the minimum is finite.

A variant of this model is considered by Serfozo.

2.2. Control of the Arrival Rate

The model of the previous subsection includes control of the arrival rate alone as a special case. Suppose the service rate is fixed at $\mu > 0$, and we are free to select the arrival rate λ at each change of state from a feasible set $L \subset [0, \infty)$. Whenever the arrival rate is λ , we incur a cost at rate $c(\lambda)$, which is nonincreasing. As usual, there is a nondecreasing holding cost rate $h(i)$. If we set $a := -\lambda$, $c(a) := c(-a)$, $\mu(a) := \mu$, then the assumptions of the previous subsection are satisfied and we conclude that the optimal arrival rate is nonincreasing in $i \geq 1$.

Note that because $\mu > 0$, we do not have to require that $c(\lambda)$ be nonnegative. Thus, the model includes the case where a reward is earned at rate $r(\lambda)$, where $r(\lambda)$ is a nondecreasing function. Most of the arrival-control models in the literature can be put into this format, sometimes after an appropriate change in the decision variable. See Stidham (1985) for a further discussion.

3. SERVICE-RATE CONTROL WITH NONEXPONENTIAL SERVICE-TIME DISTRIBUTIONS

In this section, we consider some models for the control of service in systems with general service-time distributions. In each case, we shall only study the problem of minimizing the total expected cost until the system becomes empty. The extension to average-cost minimization proceeds along the same lines as in Section 1.

3.1. M/G/1 Queue With LCFS-PR Queue Discipline; Insensitivity

In Section 1.1, we observed that the optimal policy for control of service rates in a system with no arrivals is insensitive to the service-time distribution. We shall show that the insensitivity phenomenon extends to systems with arrivals, provided that the queue discipline is restricted to be last-come, first-served, preemptive-resume (LCFS-PR).

Under this discipline, a job that arrives when the system is in state $i - 1$ goes immediately into service and continues in service as long as the system is in state i . If an arrival occurs before the job in question finishes service, it is preempted and resumes service, where it left off, at the next time point when none of the jobs that arrived after it are still in the system, that is, when the system next enters state i (necessarily from state $i + 1$). It follows that the service time of such a job coincides exactly with the total time spent by the system in state i between a transition from $i - 1$ to i and the next transition from i to $i - 1$. Thus, in the system with exponential service-time distribution, selecting the service rate to be in effect while the system is in state i is equivalent, in the case of the LCFS-PR discipline, to selecting the service rate to be used throughout the service of each job whose arrival initiates a time interval spent in state i .

Consider a system with a state-dependent Poisson arrival process operating under the LCFS-PR discipline, but suppose that the service-time distribution is no longer restricted to be exponential. Specifically, assume that the work requirements of successive arriving jobs are i.i.d. as a random variable X with mean 1. The server works at deterministic rate μ , where we can choose μ from a compact set $A \subset [0, \infty)$, with $\bar{\mu} = \max\{\mu: \mu \in A\} < \infty$. In all other respects the system is as described in Section 1. The objective, as in Section 1.2, is to move the system to state 0 at minimum total expected cost, from each possible starting state.

Since the service-time distributions are no longer

exponential, the optimal service rate conceivably could depend on the remaining work required by each of the jobs in the system, as well as the number i of such jobs. Let x_1, \dots, x_i denote the remaining work of the jobs in the system, numbered in order of arrival, and let $\underline{x}^i := (x_1, \dots, x_i)$ be the new state variable. Because of the Poisson property of the arrival process, the system clearly has the Markov property with respect to the state variable \underline{x}^i . The problem can be formulated as a continuous-time Markov decision process (Doshi 1978).

Let $v(\underline{x}^i)$ denote the minimum total expected cost until the system first becomes empty, starting in state \underline{x}^i . Because the queue discipline is LCFS-PR none of the $i - 1$ customers, with remaining work x_1, \dots, x_{i-1} , will resume service until after the next time point at which the system has $i - 1$ jobs (or fewer) present. It follows that the system has a *modified* left-skip-free property: starting in state $\underline{x}^i = (\underline{x}^{i-1}, x_i)$, the system must pass through state \underline{x}^{i-1} on its way to becoming empty. Therefore, we can write

$$v(\underline{x}^i) = z(\underline{x}^i, \underline{x}^{i-1}) + v(\underline{x}^{i-1})$$

where $z(\underline{x}^i, \underline{x}^{i-1}) :=$ the minimum total expected cost to move the system from state \underline{x}^i to state \underline{x}^{i-1} . (Note that this definition only makes sense when \underline{x}^i and \underline{x}^{i-1} are *compatible*, in the sense that $\underline{x}^i = (\underline{x}^{i-1}, x_i)$.) Thus, to determine the properties of an optimal policy for moving the system from state \underline{x}^i to state 0, it suffices to study the problem of moving the system from state \underline{x}^i to state \underline{x}^{i-1} .

Lemma 8. For all $i \geq 1$ and all $\underline{x}^i = (\underline{x}^{i-1}, x_i)$, $z(\underline{x}^i, \underline{x}^{i-1}) = x_i z(i, i - 1)$, where the function $z(i, i - 1)$ satisfies the optimality equation

$$z(i, i - 1) = \min_{\mu \in A} \left\{ \frac{[c(\mu) + h(i) + \lambda_i z(i + 1, i)]}{\mu} \right\}. \quad (20)$$

The optimal service rate for each state \underline{x}^i depends only on i , the number of customers in the system, and not on x_1, \dots, x_i , the remaining work requirements of the customers. It may be found by choosing the (largest) minimizer, $\mu(i)$, of the right-hand side of (20).

Proof. Consider the problem of moving the system from state \underline{x}^i to state \underline{x}^{i-1} at minimum total expected cost. It follows from the modified left-skip-free property that this problem may be formulated as the following deterministic control problem \mathbf{D}_i . The system starts in state $\underline{x}^i = (x_1, \dots, x_i)$ at time $t = 0$. Let

$x^i(t)$ denote the state of the system and $\mu(t)$ the service rate in effect at time $t \geq 0$.

Problem D_i

$$z(x^i, x^{i-1}) = \min \int_0^\infty \{c(\mu(t)) + h(i) + \lambda_i E[z((x^i(t), X), x^i(t))]\} \cdot 1(x_i(t) > 0) dt$$

subject to

$$x^i(0) = x^i$$

$$(d/dt)x_1(t) = \dots = (d/dt)x_{i-1}(t) = 0$$

$$(d/dt)x_i(t) = -\mu(t)$$

$$\mu(t) \in A, \quad t \in [0, \infty).$$

Since the queue discipline is LCFS-PR, the cost incurred by any policy going from a state (x^i, x) to state x^i is independent of x^i . Hence $z((x^i, x), x^i)$ is independent of x^i . Let $z(i+1, i) := E[z((x^i(t), X), x^i(t))]$ and $\tilde{h}(i) := h(i) + \lambda_i z(i+1, i)$.

It follows that Problem D_i is a special case, with $x = x_i$ and $K = \tilde{h}(i)$, of the following generic deterministic control problem: choose $T, 0 \leq T < \infty$, and $\{\mu(t), t \in [0, T]\}$ to

Problem D

$$\text{minimize } \int_0^T [c(\mu(t)) + K] dt$$

$$\text{subject to } \int_0^T \mu(t) dt = x$$

$$\mu(t) \in A, \quad t \in [0, T]$$

where $K \geq 0, x \geq 0$. Problem D has the following interpretation: There are x units of work to be processed. We are free to choose the time interval $[0, T]$ during which the work will be processed and processing rate $\mu(t)$ for each $t \in [0, T]$. We incur a service cost at rate $c(\mu)$ while the service rate is μ and a penalty at constant rate K throughout the interval $[0, T]$. Since the cost rate in Problem D does not depend on the quantity of work remaining, there is no incentive to serve one unit of work at a different rate than another. Therefore, we expect that the optimal service rate $\mu(t)$ will be constant and chosen to minimize the cost per unit of work: $[c(\mu) + K]/\mu$. A rigorous proof follows.

Let μ^* attain the minimum in

$$\min_{\mu \in A} [c(\mu) + K]/\mu.$$

Then Problem D has optimal solution $\mu(t) = \mu^*$ for all $t \leq x/\mu^*, \mu(t) = 0$ for $t > x/\mu^*$. Suppose that $\mu(t), t \in [0, T]$, is a feasible solution to D. Then

$$\begin{aligned} \int_0^T [c(\mu(t)) + K] dt &= \int_0^T \mu(t) \left\{ \frac{c(\mu(t)) + K}{\mu(t)} \right\} dt \\ &\geq \left\{ \frac{c(\mu^*) + K}{\mu^*} \right\} \int_0^T \mu(t) dt \\ &= [c(\mu^*) + K](x/\mu^*) \end{aligned}$$

which is the total cost of processing work at constant rate μ^* .

Returning to Problem D_i, we conclude that the optimal service rate $\mu(i)$ to be used when i customers are present does not depend on x^i and can be found by choosing the (largest) minimizer of the right-hand side of (20). It follows that $z(x^i, x^{i-1}) = x_i z(i, i-1)$, where $z(i, i-1)$ satisfies (20).

Since (20) is identical to (7), Lemmas 8 and 4, and Theorem 3 imply the following.

Theorem 5. *The optimal service rates for the LCFS-PR queue with state-dependent Poisson arrival process depend only on the number of jobs in the system, i , and are the same as those for the system with exponential service-time distribution. That is, the optimal service rates are insensitive to the service-time distribution. If Conditions C1, C2, and C3 hold, then the optimal service rate $\mu(i) = \bar{\mu}$ for all sufficiently large i . If Conditions C1, C2', C3', and C4 hold, then the optimal service rate $\mu(i)$ is nondecreasing in $i \geq 1$.*

3.2. M/G/1 Queue With Batch Arrivals and Nonpreemptive Discipline

Consider a single-server queue with a compound Poisson arrival process. That is, batches of jobs arrive according to an ordinary Poisson process; the sizes of successive batches are i.i.d. random variables. At each service completion, if the queue is not empty, a job is removed from the queue and placed into service. We are free to choose the service-time distribution for this job from a class of distributions indexed by the service rate μ chosen from a compact set $A \subset [0, \infty)$. The choice of service-time distribution remains fixed until service is completed; services cannot be interrupted nor preempted. Service times of successive jobs, conditional on their indices, are independent. Special cases of this model, with individual arrivals, have been considered by Schassberger (1975) and Gallisch (1979). Let $S(\mu)$ denote a generic service-time random variable from the distribution indexed by μ . Then $E[S(\mu)] = 1/\mu$. We assume that $\mu < \mu'$ implies that

$S(\mu')$ is stochastically smaller than $S(\mu)$; that is, $\Pr\{S(\mu') > t\} \leq \Pr\{S(\mu) > t\}$ for all $t \geq 0$. As in our previous models, we assume that there is a cost rate $c(\mu)$ incurred while service rate $\mu \in A$ is in effect, and a holding cost incurred at rate $h(i)$ while the system is in state i , where $c(\cdot)$ is nonnegative, nondecreasing, and continuous on A , and $h(\cdot)$ is positive and nondecreasing in $i \geq 1$.

The system clearly has the Markov property with respect to the state variable i when observed at the beginnings of successive services. If we let $v(i)$ denote the minimum expected total cost until the system first reaches state 0, starting from state i at the beginning of a service, then using the left-skip-free property of state transitions as in the previous models, we have (with $v(0) = 0$)

$$v(i) = z(i, i-1) + v(i-1), \quad i \geq 1$$

where $z(i, j) :=$ minimum total expected cost until the system first reaches state j , starting from state i at the beginning of a service. Moreover

$$\begin{aligned} z(i, i-1) &= \min_{\mu \in A} \left\{ \frac{c(\mu)}{\mu} + f(i; \mu) \right. \\ &\quad \left. + E[z(i + K(\mu) - 1, i - 1)] \right\}, \end{aligned}$$

where $f(i; \mu)$ is the expected total holding cost incurred during the service time $S(\mu)$ that begins with i jobs in the system, and $K(\mu)$ is the number of arrivals during the service time $S(\mu)$. The assumptions that arrivals are from a compound Poisson process and that service times are independent guarantee that $K(\mu)$ is independent of i and the history of the system before the beginning of the service.

We shall exploit these properties to prove that the optimal service rate is monotonic in i . In contrast to our previous proofs, we shall give a proof based on pairwise switches.

Consider a policy that is not monotonic. Specifically, suppose that the service rate in state $i + 1$ is μ_1 and the service rate in state i is μ_2 , where $\mu_2 > \mu_1$. Let us start the system in state $i + 1$ with a service about to begin and consider the total expected cost until the system next enters state $i - 1$. (The minimum of this total expected cost is $z(i + 1, i - 1) = z(i + 1, i) + z(i, i - 1)$.) Now consider the effect of a pairwise switch of the service rates so that μ_1 is used in state i and μ_2 is used in state $i + 1$. Since $S(\mu)$ is stochastically decreasing in μ , $f(i + 1; \mu_2) \leq f(i + 1; \mu_1)$. Moreover,

$K(\mu_2)$ is stochastically smaller than $K(\mu_1)$. The total expected savings during the passage from $i + 1$ to i is

$$\begin{aligned} &f((i + 1; \mu_1) - f(i + 1; \mu_2) \\ &+ E[z(i + K(\mu_1), i)] - E[z(i + K(\mu_2), i)]). \end{aligned} \quad (21)$$

Similarly, during the passage from i to $i - 1$, we incur additional expected costs because of the switch, which are equal to

$$\begin{aligned} &f(i; \mu_1) - f(i; \mu_2) \\ &+ E[z(i + K(\mu_1) - 1, i - 1)] \\ &- E[z(i + K(\mu_2) - 1, i - 1)]. \end{aligned} \quad (22)$$

We shall show that the net savings, namely (21) minus (22), is nonnegative, so that the expected total cost from state $i + 1$ to state $i - 1$ is not increased by the pairwise switch, thus establishing the optimality of a monotonic policy.

It suffices to show that

$$\begin{aligned} &f(i + 1; \mu_1) - f(i; \mu_1) \\ &\geq f(i + 1; \mu_2) - f(i; \mu_2) \end{aligned} \quad (23)$$

and

$$\begin{aligned} &E[z(i + K(\mu_1), i)] - E[z(i + K(\mu_1) - 1, i - 1)] \\ &\geq E[z(i + K(\mu_2), i)] \\ &- E[z(i + K(\mu_2) - 1, i - 1)]. \end{aligned} \quad (24)$$

Note that $f(i + 1; \mu)$ may be viewed as the total holding cost during a service time $S(\mu)$ that begins with i jobs in the system, but with holding-cost function $\hat{h}(j) := h(j + 1)$. Since $h(\cdot)$ is nondecreasing, it follows that $f(i + 1; \mu) - f(i; \mu) = E[\varphi(S(\mu))]$, where $\varphi(t)$ is nondecreasing in t . Then (23) follows from the fact that $S(\mu_1)$ is stochastically larger than $S(\mu_2)$. To prove (24), note that

$$\begin{aligned} &E[z(i + K(\mu), i)] - E[z(i + K(\mu) - 1, i - 1)] \\ &= E \left[\sum_{k=0}^{K(\mu)-1} (z(i + k + 1, i + k) \right. \\ &\quad \left. - z(i + k, i + k - 1)) \right] \end{aligned}$$

It can easily be shown that $z(i + k + 1, i + k) \geq z(i + k, i + k - 1)$, for all $i \geq 1, k \geq 0$. (See the proof of Theorem 2; the proof is essentially the same for the system at hand.) Thus (24) follows from the fact that $K(\mu_1)$ is stochastically larger than $K(\mu_2)$.

3.3. An M/PH/1 Queue With Nonpreemptive Discipline

Customers arrive to a single-server queue according to a Poisson arrival process with a mean arrival rate λ . The work requirement of each job consists of m phases, numbered $k = 1, 2, \dots, m$. Phase k contains an exponentially distributed amount of work, with mean $1/r_k$. The units in which work is measured are normalized so that

$$1/r_1 + \dots + 1/r_m = 1. \quad (25)$$

The server completes work at a deterministic rate μ , where we are free to choose μ from the compact set $A \subset [0, \infty)$. Thus, given that phase k is in progress and service rate μ is in effect, the probability that phase k is completed in the interval $(t, t + dt)$ is $\mu r_k dt + o(dt)$. Moreover, (25) implies that the expected duration of a customer's service is $1/\mu$ if the service rate μ is in effect throughout the service. There is a service-cost rate $c_k(\mu)$ incurred while service rate μ is in effect during phase k . Holding cost is incurred at a rate $h(i)$ while i customers are in the system. As usual, we assume that $c(\cdot)$ is nonnegative, nondecreasing, and continuous on A , and $h(\cdot)$ is positive and nondecreasing in $i \geq 1$.

The system has the Markov property with respect to the state variable (i, k) , where i is the number of jobs in the system and k is the index of the service phase currently in progress. Let $v(i, k)$ denote the minimum expected total cost until the system reaches state 0 (becomes empty), starting from state (i, k) . Let $z(i, k; j, l)$ denote the minimum expected total cost until the system reaches state (j, l) , starting from state (i, k) , where $j < i$. The system must pass through state $(i - 1, k)$ on its way from state (i, k) to 0, from which it follows that

$$v(i, k) = z(i, k; i - 1, k) + v(i - 1, k), \\ i \geq 1, \quad 1 \leq k \leq m$$

where state $(0, k)$ is identified with state 0. Turning our attention to $z(i, k; i - 1, k)$ and reasoning in much the same way as in Section 1, we conclude that

$$z(i, k; i - 1, k) \\ = \min_{\mu \in A} \left\{ \frac{[c_k(\mu) + h(i) + \lambda z(i + 1, k; i, k)]}{r_k \mu} \right\} \quad (26)$$

for all $i \geq 1, 1 \leq k \leq m$. Let $\mu(i, k)$ denote the optimal service rate in state (i, k) , specifically, the largest minimizer of the right-hand side of (26). It follows from essentially the same argument as used in the

proof of Theorem 2 that $\mu(i, k)$ is nondecreasing in i , for each fixed k .

In general, we cannot say anything about how $\mu(i, k)$ varies with k (cf. Jo and Stidham 1983). However, in the case where both $c_k(\mu)$ and r_k do not depend on k , it can be shown that $\mu(i, k)$ is nonincreasing in k for each $i \geq 1$, with $\mu(i, m) \geq \mu(i - 1, 1)$. Perhaps the easiest way to show this is to use the state transformation $j = (i - 1)m + m - k + 1$, so that the new state j measures the total number of work phases in the system, and then use a batch-arrival extension of the results in Section 1. (The new holding-cost function $h'(j)$ defined by $h'(im - k + 1) := h(i)$, for all $i \geq 1$ and $1 \leq k \leq m$, is nondecreasing in j if $h(i)$ is nondecreasing in i .)

Jo and Stidham studied a generalization of this model, in which the number of phases required by a job is a random variable. They assumed that $h(i)$ was convex as well as nondecreasing and proved the monotonicity of the optimal service rate first for the discounted problem, then for the average-cost problem by the usual Tauberian arguments. A random number of phases is not allowed in our approach because one loses the modified left-skip-free property that the system must pass through state $(i - 1, k)$ on its way from state (i, k) to state 0.

Phase-type service distributions, such as the ones in this model and in Jo and Stidham, can be used to approximate general service-time distributions. When the total number of phases is a random variable, one has a generalized Erlang distribution; it is possible to approximate any distribution of a nonnegative random variable arbitrarily closely with a distribution from this family (Schassberger 1973). With a deterministic number of phases, however, it is possible only to fit distributions with a coefficient of variation less than or equal to one.

ACKNOWLEDGMENT

The research of the first author was partially supported by the U.S. Army Research Office, contract DAAG29-82-K-0151, at North Carolina State University, Raleigh, and by a grant from the Science and Engineering Research Council at the University of Cambridge.

REFERENCES

- BENGTSSON, B. 1982. On Some Control Problems for Queues. Linköping Studies in Science and Technology, Dissertation No. 87, Department of Electrical

- Engineering, Linköping University, Linköping, Sweden.
- CONWAY, R., W. MAXWELL AND L. W. MILLER. 1967. *Theory of Scheduling*. Addison-Wesley, Reading, Mass.
- CRABILL, T. B. 1974. Optimal Control of a Maintenance System With Variable Service Rates. *Opns. Res.* **22**, 736–745.
- CRABILL, T. B., D. GROSS AND M. MAGAZINE. 1977. A Classified Bibliography of Research on Optimal Design and Control of Queues. *Opns. Res.* **25**, 219–232.
- DOSHI, B. T. 1978. Optimal Control of the Service Rate in an M/G/1 Queueing System. *Adv. Appl. Prob.* **10**, 682–701.
- GALLISCH, R. 1979. On Monotone Optimal Policies in a Queueing Model of M/G/1 Type With Controllable Service-Time Distribution. *Adv. Appl. Prob.* **11**, 870–887.
- GITTINS, J. C. 1979. Bandit Processes and Dynamic Allocation Indices. *J. Roy. Stat. Soc. B.* **41**, 148–164.
- JO, K. Y. 1983. Optimal Service-Rate Control of Exponential Queueing Systems. *J. Opns. Res. Soc. Jpn.* **26**, 147–165.
- JO, K. Y., AND S. STIDHAM. 1983. Optimal Service-Rate Control of M/G/1 Queueing Systems Using Phase Methods. *Adv. Appl. Prob.* **15**, 616–637.
- KARLIN, S., AND H. M. TAYLOR. 1975. *A First Course in Stochastic Processes*. Academic Press, New York.
- KEILSON, J. 1965. *Green's Function Methods in Probability Theory*. Griffin, London.
- LIPPMAN, S. 1975. Applying a New Device in the Optimization of Exponential Queueing Systems. *Opns. Res.* **23**, 687–710.
- MITCHELL, W. 1973. Optimal Service-Rate Selection in an M/G/1 Queue. *SIAM J. Appl. Math.* **24**, 19–35.
- ROSS, S. 1970. *Applied Probability Models With Optimization Applications*. Holden-Day, San Francisco.
- SCHÄL, M. 1975. Conditions for Optimality in Dynamic Programming and the Limit of n -Stage Optimal Policies to be Optimal. *Z. Wahrscheinlichkeitstheorie Verw. Geb.* **32**, 179–196.
- SCHASSBERGER, R. 1973. *Warteschlangen*. Springer-Verlag, Berlin.
- SCHASSBERGER, R. 1975. A Note on Optimal Service Selection in a Single-Server Queue. *Mgmt. Sci.* **21**, 1326–1331.
- SENNOTT, L. 1987. Average-Cost Optimal Stationary Policies in Infinite-State Markov Decision Processes—Existence and an Algorithm. Department of Mathematics, Illinois State University. Submitted for publication.
- SERFOZO, R. F. 1981. Optimal Control of Random Walks, Birth and Death Processes, and Queues. *Adv. Appl. Prob.* **13**, 61–83.
- SOBEL, M. 1974. Optimal Operation of Queues. In *Mathematical Methods in Queueing Theory*, pp. 231–261, A. B. Clarke (ed.). Springer-Verlag, Berlin.
- SOBEL, M. 1982. The Optimality of Full-Service Policies. *Opns. Res.* **30**, 636–649.
- STIDHAM, S. 1984. Optimal Control of Admission, Routing and Service in Queues and Networks of Queues: A Tutorial Review. In *Proceedings of the ARO Workshop, Analytical and Computational Issues in Logistics R and D*, pp. 330–337. George Washington University, Washington, D.C.
- STIDHAM, S. 1985. Optimal Control of Admission to a Queueing System. *IEEE Trans. Auto. Control* **AC-30**, 705–713.
- STIDHAM, S. 1988. Scheduling, Routing, and Flow Control in Stochastic Networks. In *Stochastic Differential Systems, Stochastic Control Theory and Applications*, IMA Vol. 10, pp. 529–561, W. Fleming and P. L. Lions (eds.). Springer-Verlag, Berlin.
- STIDHAM, S., AND N. U. PRABHU. 1974. Optimal Control of Queueing Systems. In *Mathematical Methods in Queueing Theory*, pp. 263–294, A. B. Clarke (ed.). Springer-Verlag, Berlin.
- STRAUCH, R. 1966. Negative Dynamic Programming. *Ann. Math. Statist.* **37**, 871–890.
- TOPKIS, D. 1978. Minimizing a Submodular Function on a Lattice. *Opns. Res.* **26**, 305–321.
- WEBER, R., AND S. STIDHAM. 1987. Optimal Control of Service Rates in Networks of Queues. *Adv. Appl. Prob.* **19**, 202–218.
- WHITTLE, P. 1980. Multi-Armed Bandits and the Gittins Index. *J. Roy. Stat. Soc. B.* **42**, 143–149.
- WHITTLE, P. 1981. Arm-Acquiring Bandits. *Ann. Prob.* **9**, 284–292.
- WHITTLE, P. 1983. *Optimization Over Time, Vol. II*. John Wiley, Chichester.
- WIJNGAARD, J., AND S. STIDHAM. 1986. Forward Recursion for Markov Decision Processes With Skip-Free-to-the-Right Transitions, Part I: Theory and Algorithms. *Math. Opns. Res.* **11**, 295–308.