Mathematical Methods in Operations Research 2005
Special issue in honor of *Arie Hordijk*

On the optimality of a full-service policy for a queueing system with discounted costs

Shaler Stidham Jr.

Department of Statistics and Operations Research
CB 3180, Smith Building
University of North Carolina
Chapel Hill, NC 27599-3180
USA

February 1, 2005 (revised, February 15, 2005)

Abstract We provide weak sufficient conditions for a full-service policy to be optimal in a queueing control problem in which the service rate is a dynamic decision variable. In our model there are service costs and holding costs and the objective is to minimize the expected total discounted cost over an infinite horizon. We begin with a semi-Markov decision model for a single-server queue with exponentially distributed inter-arrival and service times. Then we present a general model with weak probabilistic assumptions and demonstrate that the full-service policy minimizes both finite-horizon and infinite-horizon total discounted cost on each sample path.

1 Introduction

We consider a single-server queueing system in which the service rate is a dynamic decision variable. There is a service cost per unit time (a non-decreasing function of the service rate) and a holding cost per unit time (a non-decreasing function of the number of customers in the system). The objective is to minimize the expected discounted cost over a finite or infinite horizon.

For the problem without discounting, under rather general conditions, it is well known that an optimal policy always serves at the fastest possible rate (the *full-service policy*). The conditions include a requirement that the fastest rate minimize the expected service cost incurred during the service time of a single customer. Since the full-service policy clearly minimizes the holding cost incurred up to any time point, it is intuitively obvious in this case that it minimizes the total expected cost until the n-th departure for any n, and hence the long-run average expected cost as well. This intuition was the basis for a rigorous proof of the optimality of the full-service policy for a very general single-server queueing system in a seminal paper by Sobel [4].

For the problem with discounting, this intuition is no longer valid. It may be preferable to use a slower, cheaper rate to save on the service-cost rate now, since the resulting higher future cost is discounted and thus has a diminished impact on the present value of the total cost. Nonetheless, it is clear that the full-service policy will still often be optimal in the problem with discounting, provided that the discount rate is sufficiently small. There are scattered results in the literature that confirm this conjecture, but often under strong economic and probabilistic conditions. In particular, most papers assume exponentially distributed inter-arrival and service times, in order to formulate the optimization problem as a semi-Markov decision process (*SMDP*).

It is the objective of this paper to provide a unified treatment of the optimality of a full-service policy for the discounted-cost problem. In particular, we seek the weakest possible economic and probabilistic conditions under which a full-service policy is optimal. To fix ideas and provide a brief review of the literature, we begin (Section 2) with a general *SMDP* model for the exponential case, which leads naturally to a general set of sufficient conditions for the optimality of a full-service policy in that setting, for the case of a linear holding-cost function. The key to this analysis is a transformation, in which we measure the cost of each policy relative to that of a fixed policy, namely the policy that always uses the zero service rate (the *no-service policy*). The analysis of this transformed *SMDP* model suggests a set of sufficient conditions for a much more general model, which is presented in Section 3. In this setting we are able to use sample-path analysis to prove that the full-service policy is optimal in a more general sense. Namely, it minimizes the total discounted cost until the n-th departure, for any realization of the sequence of arrival times and service requirements of the customers. The analysis is similar to that used by Sobel [4] for the problem without discounting. Finally, in Section 4 we return to the *SMDP* model for the exponential case and show how the optimality of a full-service policy extends to non-linear holding-cost functions.

2 An *M/M/1* Queue with Discounted Costs.

We consider an *M/M/1* queue with fixed arrival rate λ and service rate μ, which can be dynamically varied between 0 and $\bar{\mu} < \infty$. There is a service cost which is incurred at rate $c(\mu) \geq 0$ while service rate μ is in effect, where $c(\cdot)$ is continuous and non-decreasing in $\mu \in [0, \bar{\mu}]$ with $c(0) = 0$. There is a nonnegative holding cost rate $h(i)$ per unit time while there are i customers in system, where $h(i)$ is non-decreasing in $i \geq 0$. Future costs are continuously discounted at rate $\alpha > 0$. The objective is to minimize the expected α-discounted total cost over an infinite horizon, from each starting state $i \in S = \{0, 1, ...\}$, where the state is the number of customers in the system.

We observe the system at each arrival and service completion instant (i.e., at each change of state) in order to formulate the problem as a semi-Markov decision process (*SMDP*). Let $v(i)$ denote the minimum expected total discounted cost over an infinite horizon, starting from state $i \geq 0$. At each change of state, a service rate $\mu \in [0, \bar{\mu}]$ is selected, which then remains in effect until the next change of state. Without loss of generality we assume that $\mu = 0$ is always selected when $i = 0$. Using standard arguments, one can show that v satisfies the optimality equations,

$$v(i) = \min_{\mu \in [0,\bar{\mu}]} \left\{ \frac{h(i) + c(\mu)}{\alpha + \lambda + \mu} + \left(\frac{\lambda}{\alpha + \lambda + \mu} \right) v(i+1) + \left(\frac{\mu}{\alpha + \lambda + \mu} \right) v(i-1) \right\} , \; i \geq 1 , \tag{1}$$

$$v(0) = \frac{h(0)}{\alpha + \lambda} + \left(\frac{\lambda}{\alpha + \lambda} \right) v(1) . \tag{2}$$

For $i \geq 1$, note that (1) holds if and only if

$$(\alpha + \lambda + \mu)v(i) \leq h(i) + c(\mu) + \lambda v(i+1) + \mu v(i-1) ,$$

or, equivalently,

$$(\alpha + \lambda + \bar{\mu})v(i) \leq h(i) + c(\mu) + \lambda v(i+1) + \mu v(i-1) + (\bar{\mu} - \mu)v(i) ,$$

for all $\mu \in [0, \bar{\mu}]$, with equality for at least one μ. But the last assertion is true if and only if

$$(\alpha+\lambda+\bar{\mu})v(i) = \min_{\mu \in [0,\bar{\mu}]} \{h(i)+c(\mu)+\lambda v(i+1)+\mu v(i-1)+(\bar{\mu}-\mu)v(i)\} . \tag{3}$$

On the other hand, equation (2) for $i = 0$ holds if and only if

$$\begin{aligned}(\alpha + \lambda + \bar{\mu})v(0) &= h(0) + c(0) + \lambda v(1) + \bar{\mu} v(0) \\ &\leq h(0) + c(\mu) + \lambda v(1) + (\bar{\mu} - \mu)v(0) + \mu v(-1) ,\end{aligned}$$

for all $\mu \in [0, \bar{\mu}]$, where the latter inequality holds since $c(\mu) \geq c(0) = 0$ for all $\mu \in [0, \bar{\mu}]$, and we take $v(-1) = v(0)$. Hence (2) is equivalent to

$$(\alpha + \lambda + \bar{\mu})v(0) = \min_{\mu \in [0,\bar{\mu}]} \{h(0) + c(\mu) + \lambda v(1) + \mu v(-1) + (\bar{\mu} - \mu)v(0)\} ,$$

which in turn is equivalent to (3) for $i = 0$.

It follows that $v(\cdot)$ satisfies the uniformized version (3) of the optimality equation for all $i \geq 0$. By appropriate choice of time units we assume that $\alpha + \lambda + \bar{\mu} = 1$, without loss of generality. Thus $v(\cdot)$ satisfies

$$v(i) = h(i) + \lambda v(i+1) + \min_{\mu \in [0,\bar{\mu}]} \{c(\mu) + \mu v(i-1) + (\bar{\mu} - \mu)v(i)\} , \; i \geq 0 ,$$

or, equivalently,

$$v(i) = h(i) + \lambda v(i+1) + \bar{\mu} v(i) + \min_{\mu \in [0,\bar{\mu}]} \{c(\mu) - \mu[v(i) - v(i-1)]\} \ , \ i \geq 0 \ . \quad (4)$$

Note that $v(i) - v(i-1)$ is the benefit of a service completion in state i, i.e., the savings in expected future discounted cost from being in state $i-1$ rather than i. While in state i and using service rate μ, we incur a direct service cost at rate $c(\mu)$, while service completions occur at rate μ, each bringing a benefit $v(i) - v(i-1)$. Therefore, the *net* variable cost rate (to be minimized) is $g(i,\mu) := c(\mu) - \mu[v(i) - v(i-1)]$. Let $\mu(i)$ denote the (largest) minimizer of $g(i,\mu)$ in $[0,\bar{\mu}]$. The following lemma is immediate.

Lemma 1 *Suppose $v(i) - v(i-1)$ is nondecreasing in $i \geq 1$. That is, $v(\cdot)$ is convex in the integer variable $i \geq 0$. Then $\mu(i)$ is nondecreasing in $i \geq 0$.*

2.1 Inductive Proof of Convexity of $v(\cdot)$

Thus, to prove that an optimal policy is monotonic, it suffices to show that $v(i)$ is convex in i. Henceforth we shall make the following assumption.

Assumption. The holding cost rate, $h(i)$, is convex in $i \geq 0$.

The standard approach for proving that the optimal value function, $v(\cdot)$, is convex is induction on a sequence of successive approximations, $v_n(i)$, to $v(i)$. Let $\{v_n, n \geq 0\}$ be defined recursively by $v_0 \equiv 0$ and, for $n \geq 1$, $i \geq 0$, by

$$v_n(i) = h(i) + \lambda v_{n-1}(i+1) + \min_{\mu \in [0,\bar{\mu}]} \{c(\mu) + \mu v_{n-1}(i-1) + (\bar{\mu} - \mu) v_{n-1}(i)\} \ .$$

It follows from the theory of Markovian decision processes (see Puterman [3]) that $v_n(i) \to v(i)$ as $n \to \infty$, for all $i \geq 0$. Convexity of v will then follow if we can show that v_n is convex for each n. This can easily be done by induction on n (see, e.g., Lippman [2]). Thus we have the following theorem.

Theorem 1 *For all $n \geq 0$, $v_n(i)$ is convex in $i \geq 0$. Hence $v(i)$ is convex in $i \geq 0$ and an optimal service-rate control policy for the infinite-horizon, discounted problem is monotonic: $\mu(i)$ is non-decreasing in $i \geq 0$.*

2.2 Extensions and comments.

1) More general service-cost function and feasible action set.

Nothing in the proof of monotonicity specifically required that $c(\mu)$ be continuous or that the feasible action set A be an interval, $[0,\bar{\mu}]$. We simply need for maximum of the r.h.s. of the optimality equation to be attained for each i and for A to contain both $\mu = 0$ (see above) and a finite maximal element $\bar{\mu}$ (in order to be able to uniformize the process). For example, we could have a finite set of possible actions, $A = \{\mu_0, \mu_1, \ldots, \mu_m\}$, where $0 = \mu_0 < \mu_1 < \ldots < \mu_m = \bar{\mu}$. More generally, it suffices for A to be a compact (that is, closed and bounded) set on which $c(\mu)$ is continuous.

2) Elimination of non-optimal actions.

For each state $i \geq 1$, our problem has the form $\min_{\mu \in A}\{c(\mu) - \mu K\}$, where K is a given constant. Suppose $\mu_1 < \mu_2 < \mu_3$ ($\mu_1, \mu_2, \mu_3 \in A$) and

$$\frac{c(\mu_2) - c(\mu_1)}{\mu_2 - \mu_1} \geq \frac{c(\mu_3) - c(\mu_2)}{\mu_3 - \mu_2} . \tag{5}$$

Then it follows that either $c(\mu_2) - \mu_2 K \geq c(\mu_1) - \mu_1 K$ or $c(\mu_2) - \mu_2 K \geq c(\mu_3) - \mu_3 K$ and hence action μ_2 need never be used in an optimal policy. This implies that, without loss of optimality, we may replace A by its convex hull $[0, \bar{\mu}]$ and $c(\mu)$ by its lower convex envelope on $[0, \bar{\mu}]$. (Here it is assumed that 0 and $\bar{\mu}$ are, respectively, the minimal and maximal elements of A.)

3) Optimality of extremal (bang-bang) policies.

Suppose A has the property that $c(\bar{\mu})/\bar{\mu} \leq c(\mu)/\mu$, for all $\mu \in A$. Then (5) holds with $\mu_1 = 0$, $\mu_2 = \mu$, and $\mu_3 = \bar{\mu}$, and it follows that an *extremal* or *bang-bang* policy is optimal: in each state $i \geq 1$, use either action 0 or action $\bar{\mu}$. When combined with the monotonicity of an optimal policy, this implies that there exists an integer i_0 such that $\mu(i) = 0$, $0 \leq i \leq i_0$, $\mu(i) = \bar{\mu}$, $i > i_0$.

2.3 Optimality of a Full-Service Policy.

A *full-service* policy is a special case of an extremal policy, in which $i_0 = 0$: that is, the maximal service rate, $\bar{\mu}$, is used in all states $i \geq 1$. In many cases, it is intuitively plausible that a full-service policy should be optimal, at least when there is no discounting. To see why, note that $c(\mu)/\mu = c(\mu) \times (1/\mu) =$ (service cost rate per unit time) $\times$ (expected service time) $=$ expected service cost per completed service. Thus, the extremal condition, $c(\bar{\mu})/\bar{\mu} \leq c(\mu)/\mu$, for all $\mu \in A$, implies that the full-service policy minimizes the expected cost per completed service, and hence the expected service cost incurred up to the n-th departure, for any n. But it is obvious that the full-service policy also minimizes the expected holding cost up to any time t, and therefore (in particular) up to the n^{th} departure, for any n. Thus the full-service policy minimizes the expected *total* cost and the expected *average* cost per departure until the n^{th} departure. In any stable system, the long-run average departure rate equals the long-run average arrival rate, which is independent of policy (and equals λ in the present problem). Thus, the full-service policy minimizes the *long-run average total cost per unit time.*

A rigorous proof, based on this argument, is in Sobel [4]. As the above argument suggests, the proof works for very general interarrival-time and service-time distributions. Essentially, the interarrival-time sequence can be arbitrary, possibly dependent and/or non-identically distributed, as long as the long-run average interarrival time exists (i.e., as long as the sequence satisfies an ergodic theorem). The service times are drawn from a family of distributions, indexed by the rate μ, that is, the reciprocal of the mean service time, with the property that a service time performed at rate μ is stochastically larger than one performed at rate $\mu' > \mu$.

The argument does not work when there is discounting, however, since then we may wish to use a slow service rate and incur a lower cost rate $c(\mu)$ now, postponing the service completion and incurring higher holding costs in the future, when their present value is lower.

Let us now return to the discounted problem, to see if we can use the *SMDP* formulation to determine conditions under which a full-service policy is optimal. First, we shall assume that

$$c(\bar{\mu})/\bar{\mu} \leq c(\mu)/\mu \text{ , for all } \mu \in A \text{ ,} \tag{6}$$

which implies that an extremal, monotonic policy is optimal (see comment (3) above). Thus, there exists an integer i_0 such that $\mu(i) = 0$, $0 \leq i \leq i_0$, $\mu(i) = \bar{\mu}$, $i > i_0$. Our goal is to find a condition which, together with (6), implies that $i_0 = 0$. To keep the exposition simple, we shall assume that the holding cost is linear:

$$h(i) = h \cdot i \text{ , } i \geq 0 \text{ .} \tag{7}$$

(We sketch the extension to non-linear holding costs in Section 4.) In the linear-holding-cost case the optimality equation (4) takes the form

$$v(i) = h \cdot i + \lambda v(i+1) + \bar{\mu} v(i) + \min_{\mu \in A} \{c(\mu) - \mu[v(i) - v(i-1)]\} \text{ , } i \geq 0 \text{ .} \tag{8}$$

It turns out to be convenient to work with a transformed model. We do this by expressing all value functions relative to the value function for a fixed policy: the policy that chooses service rate $\mu = 0$ in all states i (the *no-service* policy). Let v^o denote the value function associated with this policy. That is, $v^o(i)$ is the expected total α-discounted cost over an infinite horizon, starting from state i at time 0, assuming that no customers are ever served. It follows from the usual conditioning arguments that v^o satisfies the following (uniformized) functional equation,

$$v^o(i) = h \cdot i + \lambda v^o(i+1) + \bar{\mu} v^o(i) \text{ , } i \geq 0 \text{ ,} \tag{9}$$

where as before we have assumed that $\alpha + \lambda + \bar{\mu} = 1$.

Since no service cost is ever incurred, $v^o(i)$ equals the α-discounted total holding cost incurred over the infinite horizon by the i customers in the system at time 0 and by all future arrivals, assuming that no customers are ever served. Thus

$$v^o(i) = \mathrm{E}\left[\int_0^\infty e^{-\alpha t} h \cdot (i + N(t)) dt\right] \text{ ,} \tag{10}$$

where $\{N(t), t \geq 0\}$ is a Poisson process with parameter λ. It is a straightforward exercise to use this formula to derive an explicit algebraic expression for $v^o(i)$, in terms of i and the parameters h, α and λ. But we shall only need (10) and the fact that v^o satisfies the functional equation (9).

First observe (from (10) that

$$v^o(i) - v^o(i-1) = \mathrm{E}\left[\int_0^\infty e^{-\alpha t} h dt\right] = h/\alpha \ , \ i \geq 1 \ . \tag{11}$$

The difference $v^o(i) - v^o(i-1)$ is precisely the expected α-discounted holding cost incurred from time 0 to time ∞ by the additional customer who is in the system from time 0 to time ∞.

Now let us define $\tilde{v}(i) := v(i) - v^o(i)$, $i \geq 0$. Subtracting (9) from the optimality equation (8) for v, we obtain

$$\tilde{v}(i) = \lambda \tilde{v}(i+1) + \bar{\mu}\tilde{v}(i) + \min_{\mu \in A}\{c(\mu) - \mu[v(i) - v(i-1)]\} \ , \ i \geq 0 \ ,$$

so that $\tilde{v}$ satisfies the optimality equations

$$\begin{aligned} \tilde{v}(i) &= \lambda \tilde{v}(i+1) + \bar{\mu}\tilde{v}(i) \\ &\qquad + \min_{\mu \in A}\{c(\mu) - \mu[\tilde{v}(i) - \tilde{v}(i-1) + h/\alpha]\} \ , \ i \geq 1 \ , \quad (12) \\ \tilde{v}(0) &= \lambda \tilde{v}(1) + \bar{\mu}\tilde{v}(0) \ , \quad (13) \end{aligned}$$

where we have used (11).

Remark. We see from (12) and (13) that the effect of subtracting v^o from the optimal value functions is to replace the original problem, in which holding costs were charged continuously through time, with an equivalent problem with no holding cost *per se*, in which a reward h/α is earned at every service completion.

The intuitive idea behind this transformation is the following. Charging holding costs continuously through time is equivalent to charging the system, at each arrival point, the discounted total holding cost, h/α, which the arriving customer would incur over the infinite horizon if never served, and then, when the customer departs, rewarding the system with the savings in discounted holding cost (once again, h/α) caused by the customer not being in the system over the infinite horizon remaining from that time point. Since the amounts charged at the arrival points are independent of the service-rate control policy, we can remove them from the value functions without affecting the optimization. This is precisely what we are doing when we subtract v^o from the optimal value functions. This idea, which is due to Bell [1], is discussed more rigorously in a general setting in Stidham [5].

From (12) we see that the full-service policy will be optimal if (and only if)

$$\tilde{v}(i) - \tilde{v}(i-1) \geq \frac{c(\bar{\mu})}{\bar{\mu}} - h/\alpha \ , \ i \geq 1 \ . \tag{14}$$

Theorem 2 *Suppose (6) and (7) hold and*

$$\frac{c(\bar{\mu})}{\bar{\mu}} \leq h/\alpha \ . \tag{15}$$

Then the full-service policy is optimal.

Proof . We shall use induction on a sequence of successive approximations, $\tilde{v}_n$, to the transformed optimal value function, $\tilde{v}$, to show that (14) holds. Let $\tilde{v}_n, n \geq 0$, be defined recursively by $\tilde{v}_0 \equiv 0$ and, for $n \geq 1$,

$$\begin{aligned} \tilde{v}_n(i) &= \lambda \tilde{v}_{n-1}(i+1) + \bar{\mu}\tilde{v}_{n-1}(i) \\ &\quad + \min_{\mu \in A}\{c(\mu) - \mu[\tilde{v}_{n-1}(i) - \tilde{v}_{n-1}(i-1) + h/\alpha]\} \ , \ i \geq 1 \ , \quad (16) \\ \tilde{v}_n(0) &= \lambda \tilde{v}_{n-1}(1) + \bar{\mu}\tilde{v}_{n-1}(0) \ , \quad (17) \end{aligned}$$

Suppose $\tilde{v}_{n-1}$ satisfies (14) (the induction hypothesis). Then the full-service policy is optimal for the n-stage problem, so that

$$\tilde{v}_n(i) = \lambda \tilde{v}_{n-1}(i+1) + c(\bar{\mu}) + \bar{\mu}[\tilde{v}_{n-1}(i-1) - h/\alpha] \ , \ i \geq 1 \ ,$$

and $\tilde{v}_n(0)$ satisfies (17). It follows that $(i \geq 2)$

$$\begin{aligned} \tilde{v}_n(i) - \tilde{v}_n(i-1) &= \lambda[\tilde{v}_{n-1}(i+1) - \tilde{v}_{n-1}(i)] + \bar{\mu}[\tilde{v}_{n-1}(i-1) - \tilde{v}_{n-1}(i-2)] \\ &\geq (\lambda + \bar{\mu}) \left(\frac{c(\bar{\mu})}{\bar{\mu}} - h/\alpha \right) \\ &\geq \frac{c(\bar{\mu})}{\bar{\mu}} - h/\alpha \ , \end{aligned}$$

where the last inequality follows from (15) and the assumption that $\alpha + \lambda + \bar{\mu} = 1$. Similarly, for $i = 1$ we have

$$\begin{aligned} \tilde{v}_n(1) - \tilde{v}_n(0) &= \lambda[\tilde{v}_{n-1}(2) - \tilde{v}_{n-1}(1)] + c(\bar{\mu}) - \bar{\mu}(h/\alpha) \\ &\geq (\lambda + \bar{\mu}) \left(\frac{c(\bar{\mu})}{\bar{\mu}} - h/\alpha \right) \\ &\geq \frac{c(\bar{\mu})}{\bar{\mu}} - h/\alpha \ . \end{aligned}$$

Thus we have shown that $\tilde{v}_n$ satisfies (14), which completes the inductive step. To start the induction, note that

$$\tilde{v}_0(i) - \tilde{v}_0(i-1) = 0 \geq \frac{c(\bar{\mu})}{\bar{\mu}} - h/\alpha \ ,$$

since $\tilde{v}_0 \equiv 0$. Since $\tilde{v}_n \to \tilde{v}$ as $n \to \infty$, it follows that $\tilde{v}$ satisfies (14) and hence a full-service policy is optimal. ∎

Referring again to the transformed optimality equation (12) and observing that

$$c(\mu) - \mu \cdot h/\alpha = \mu \left(\frac{c(\mu)}{\mu} - h/\alpha \right) \ ,$$

we see that the problem is equivalent to one in which the system earns a lump-sum benefit, $b(\mu) := h/\alpha - c(\mu)/\mu$, at the instant of each service completion in state i when the service rate is μ, and no other rewards or costs are earned or incurred. With this interpretation, it is intuitively obvious that, if $b(\bar{\mu}) \geq b(\mu)$ for all $\mu \in A$ and $b(\bar{\mu}) \geq 0$ – that is, if (6) and (15) hold, respectively – then the full-service policy will be optimal, since it both maximizes the benefit earned at each service completion and minimizes the time until each service completion. In the presence of discounting, a given benefit has a larger present value if it is earned earlier, provided the benefit is non-negative – hence the need for condition (15). Compare this argument to the intuitive argument for optimality of the full-service policy in the undiscounted case, given at the beginning of this section. In the undiscounted case we needed only the first of the two conditions, namely (6).

This intuitive argument for the optimality of the full-service policy in the discounted case can be made the basis of a rigorous proof in a more general setting, in which the interarrival times and service times need not be exponential nor mutually independent and in which optimality holds on a sample-path basis. This is the subject of the next section.

3 The *G/G/1* Queue with Discounting

In this section we consider a single-server queue with arbitrarily distributed interarrival and service times in the presence of discounting. The notation and assumptions follow (roughly) those of Sobel [4], who considered a general single-server queue without discounting. Like Sobel, we obtain strong (stochastic-order and/or sample-path) finite- and infinite-horizon versions of the optimality of the full-service policy.

Customers are numbered $j = 1, 2, \ldots$, in order of arrival. Let $A_0 = 0$ and let $A_j \geq 0$ denote the instant of arrival of customer j, where $A_j \geq A_{j-1}$, $j \geq 1$. (Note that batch arrivals are permitted.) We make no distributional nor independence assumptions about the sequence of random variables, $\{A_j, j \geq 1\}$. Let X_j denote the work required by customer j, $j \geq 1$. Assume that X_j, $j \geq 1$, are i.i.d. non-negative random variables distributed as the generic random variable X with distribution function

$$G(x) = \mathrm{P}\{X \leq x\} \ , \ x \geq 0 \ .$$

The stochastic sequences $\{X_j, j \geq 1\}$ and $\{A_j, j \geq 1\}$ are independent of each other.

The server works at a deterministic rate μ, $\mu \in A$, where μ is a decision (control) variable and A is the set of feasible service rates, which contains a maximal element, $\bar{\mu}$. Thus, if the server works at rate μ throughout the service of customer j, then customer j's service time is $S_j = X_j/\mu$, with $\mathrm{E}[S_j] = 1/\mu$.

As an example, suppose the customers are messages in a communication system and the server is a communication channel. Then X_j could be the number of packets in message j and the service rate μ could be the number of packets per second transmitted by the channel, so that the time required to transmit message j is

$$\frac{X_j \text{ packets}}{\mu \text{ packets/second}} = X_j/\mu \text{ seconds} \ .$$

Let F_μ denote the distribution function of the service time, when service rate μ is chosen, where $\mu \in A$. Then under our assumptions we have, for any $\mu, \mu' \in B$, $\mu < \mu'$,

$$F_\mu(z) = \mathrm{P}\{X/\mu \leq z\} = G(\mu z) \leq G(\mu' z) = \mathrm{P}\{X/\mu' \leq z\} = F_{\mu'}(z) \ , \ z \geq 0 \ .$$

Thus the service time is stochastically decreasing in μ. In particular, for all $\mu \in A$, we have

$$F_\mu(z) \leq F_{\bar{\mu}}(z) \ , \ z \geq 0 \ ,$$

so that the service time is stochastically minimized by choosing the maximal service rate, $\bar{\mu}$.

Without loss of generality, we shall assume henceforth that $E[X] = \int_0^\infty x dG(x) = 1$. That is, we shall measure work in units of the mean work required by a customer. In the communication example, if the average message length is, say, 10 mb, this would mean that the transmission rate μ is measured in the number of multiples of 10 mb that can be transmitted per second. It follows that the mean service time, when using service rate μ, is given by

$$\int_0^\infty z dF_\mu(z) = \mathrm{E}[X/\mu] = 1/\mu \ .$$

Thus, μ corresponds to what we often mean by service rate: the reciprocal of the mean service time.

A service rate μ is chosen at the beginning of the service of each customer and remains in effect until that service is completed. Motivated by the results for the *M/M/1* queue, we assume that the system earns a deterministic, lump-sum benefit, $b(\mu)$, at the instant of each service completion when service rate μ is in effect, and that no other rewards or costs are earned or incurred. We make the following assumptions about the benefit function, $b(\mu)$.

Assumption 1. $b(\bar{\mu}) \geq b(\mu)$ for all $\mu \in A$.

Assumption 2. $b(\bar{\mu}) \geq 0$.

Future benefits are continuously discounted at discount rate α.

Let μ_j denote the service rate chosen for the service of customer j. We claim that setting $\mu_j = \bar{\mu}$ for all $j \geq$ maximizes the total discounted benefit earned up to and including the k^{th} departure, for all $k \geq 1$, for any realization of the customer arrival times, A_j, and work requirements, X_j,

$j \geq 1$. This is intuitively obvious, following the argument at the end of the previous section, but it can be formally proved as follows.

Consider an arbitrary policy for choosing the service rates, μ_j, associated with customers $j = 1, 2, \ldots$. Consider a fixed and arbitrary realization of the arrival times, A_j, $j \geq 1$, and work requirements, X_j, $j \geq 1$. Let D_j denote the departure time of customer j, $j \geq 1$. First note that, since the queue discipline is first-come, first served,

$$\begin{aligned} D_j &= \max\{A_j, D_{j-1}\} + S_j \\ &= \max\{A_j, D_{j-1}\} + X_j/\mu_j \ . \end{aligned}$$

for all $j \geq 1$ (with $D_0 := 0$). Since X_j/μ_j is minimized by setting $\mu_j = \bar{\mu}$, it follows by induction on $j = 1, 2, \ldots,$ that the departure time, D_j, of each customer j is minimized by the full-service policy. Now, for any $k \geq 1$, the total discounted benefit earned in $[0, D_k]$ is

$$\sum_{j=1}^{k} e^{-\alpha D_j} b(\mu_j)$$

Hence the full-service policy maximizes this quantity.

Thus, we have the following theorem, which is a sample-path counterpart of the result for the *M/M/1* queue.

Theorem 3 *Under Assumptions 1 and 2, for any realization of the arrival times, A_j, $j \geq 1$, and work requirements, X_j, $j \geq 1$, setting $\mu_j = \bar{\mu}$ for all $j \geq 1$ maximizes the total discounted benefit earned during $[0, D_k]$,*

$$\sum_{j=1}^{k} e^{-\alpha D_j} b(\mu_j) \ ,$$

for all $k \geq 1$.

The following corollaries are immediate.

Corollary 1 *The full-service policy maximizes the expected discounted benefit earned until the k^{th} departure,*

$$\mathrm{E}\left[\sum_{j=1}^{k} e^{-\alpha D_j} b(\mu_j)\right] \ ,$$

for all $k \geq 1$.

Corollary 2 *The full-service policy maximizes the expected discounted benefit earned over an infinite horizon,*

$$\mathrm{E}\left[\sum_{j=1}^{\infty} e^{-\alpha D_j} b(\mu_j)\right] \ .$$

4 Extension to Non-Linear Holding Cost in the *SMDP* Model

In this section we return to the *SMDP* model of Section 2 for the system with exponentially distributed inter-arrival and service times and show how the optimality of a full-service policy extends to non-linear holding costs.

Let $\mathbf{h} = \{h(i), i \in S\}$ denote the holding-cost function and assume (as before) that $h(i)$ is a convex, non-decreasing function of $i \in S$. To indicate the dependence of the optimal value function on the holding-cost function, we use the notation, $v(i; \mathbf{h})$, for the minimal total discounted cost over an infinite horizon, starting in state i, given that the holding-cost function is $\mathbf{h}$. Then $v(i; \mathbf{h})$ satisfies the (uniformized) optimality equation (cf. (4))

$$\begin{aligned} v(i;\mathbf{h}) &= h(i) + \lambda v(i+1;\mathbf{h}) + \bar{\mu} v(i;\mathbf{h}) \\ &\qquad + \min_{\mu \in A}\{c(\mu) - \mu[v(i;\mathbf{h}) - v(i-1;\mathbf{h})]\}\ ,\ i \ge 1\ , \\ v(0;\mathbf{h}) &= h(0) + \lambda v(1;\mathbf{h}) + \bar{\mu} v(0;\mathbf{h}) \end{aligned}$$

Similarly, let $v^f(i; \mathbf{h})$ denote the value function for the full-service policy, which satisfies the following functional equation:

$$\begin{aligned} v^f(i;\mathbf{h}) &= c(\bar{\mu}) + h(i) + \lambda v^f(i+1;\mathbf{h}) + \bar{\mu} v^f(i-1;\mathbf{h})\ ,\ i \ge 1\ , \\ v^f(0;\mathbf{h}) &= h(0) + \lambda v^f(1;\mathbf{h}) + \bar{\mu} v^f(0;\mathbf{h}) \end{aligned}$$

We know that the full-service policy is optimal for the holding-cost function $\mathbf{h}$ if and only if

$$c(\bar{\mu}) - \bar{\mu}[v^f(i;\mathbf{h}) - v^f(i-1;\mathbf{h})] \le c(\mu) - \mu[v^f(i;\mathbf{h}) - v^f(i-1;\mathbf{h})]\ ,\ \mu \in A\ . \tag{18}$$

for all $i \ge 1$. It is intuitively plausible that, if the full-service policy is optimal for a particular holding-cost function $\mathbf{h}$, then it is also optimal for all "larger" holding-cost functions. The exact meaning of "larger", however, must be specified with care. The following lemma gives the result that we shall need.

Lemma 2 *Consider two holding-cost functions,* $\mathbf{h}^1 = \{h^1(i), i \in S\}$ *and* $\mathbf{h}^2 = \{h^2(i), i \in S\}$, *and suppose that* $h^1(i) - h^1(i-1) \le h^2(i) - h^2(i-1)$ *for all* $i \ge 1$. *Then, if the full-service policy is optimal for* $\mathbf{h}^1$, *it is also optimal for* $\mathbf{h}^2$.

Proof Suppose the full-service policy is optimal for $\mathbf{h}^1$, so that (18) holds for all $i \ge 1$ with $\mathbf{h} = \mathbf{h}^1$. Then to show that the full-service policy is optimal for $\mathbf{h}^2$, it suffices to show that

$$v^f(i;\mathbf{h}^1) - v^f(i-1;\mathbf{h}^1) \le v^f(i;\mathbf{h}^2) - v^f(i-1;\mathbf{h}^2)\ ,\ i \ge 1\ . \tag{19}$$

We verify this inequality by induction on a sequence of approximations to the infinite-horizon value functions. For any holding-cost function $\mathbf{h}$, let the (finite-horizon) value functions $v^f_n(i; \mathbf{h})$ be defined recursively for $n \ge 1$ by

$$\begin{aligned} v^f_n(i;\mathbf{h}) &= c(\bar{\mu}) + h(i) + \lambda v^f_{n-1}(i+1;\mathbf{h}) + \bar{\mu} v^f_{n-1}(i-1;\mathbf{h})\ ,\ i \ge 1\ , \\ v^f_n(0;\mathbf{h}) &= h(0) + \lambda v^f_{n-1}(1;\mathbf{h}) + \bar{\mu} v^f_{n-1}(0;\mathbf{h})\ , \end{aligned}$$

with $v_0^f(\cdot;\mathbf{h}) \equiv 0$. Let $n \geq 1$ be given and suppose

$$v_{n-1}^f(i;\mathbf{h}^1) - v_{n-1}^f(i-1;\mathbf{h}^1) \leq v_{n-1}^f(i;\mathbf{h}^2) - v_{n-1}^f(i-1;\mathbf{h}^2)\ ,\ i \geq 1\ .$$

Then for $i \geq 2$ we have

$$\begin{aligned}
&v_n^f(i;\mathbf{h}^1) - v_n^f(i-1;\mathbf{h}^1)\\
&= h^1(i) - h^1(i-1) + \lambda(v_{n-1}^f(i+1;\mathbf{h}^1) - v_{n-1}^f(i;\mathbf{h}^1))\\
&\qquad +\bar{\mu}(v_{n-1}^f(i-1;\mathbf{h}^1) - v_{n-1}^f(i-2;\mathbf{h}^1))\\
&\leq h^2(i) - h^2(i-1) + \lambda(v_{n-1}^f(i+1;\mathbf{h}^2) - v_{n-1}^f(i;\mathbf{h}^2))\\
&\qquad +\bar{\mu}(v_{n-1}^f(i-1;\mathbf{h}^2) - v_{n-1}^f(i-2;\mathbf{h}^2))\\
&= v_n^f(i;\mathbf{h}^2) - v_n^f(i-1;\mathbf{h}^2)\ ,
\end{aligned}$$

whereas for $i = 1$ we have

$$\begin{aligned}
&v_n^f(1;\mathbf{h}^1) - v_n^f(0;\mathbf{h}^1)\\
&= c(\bar{\mu}) + h^1(1) - h^1(0) + \lambda(v_{n-1}^f(2;\mathbf{h}^1) - v_{n-1}^f(1;\mathbf{h}^1))\\
&\qquad +\bar{\mu}(v_{n-1}^f(0;\mathbf{h}^1) - v_{n-1}^f(0;\mathbf{h}^1))\\
&\leq c(\bar{\mu}) + h^2(1) - h^2(0) + \lambda(v_{n-1}^f(2;\mathbf{h}^2) - v_{n-1}^f(1;\mathbf{h}^2))\\
&\qquad +\bar{\mu}(v_{n-1}^f(0;\mathbf{h}^2) - v_{n-1}^f(0;\mathbf{h}^2))\\
&= v_n^f(1;\mathbf{h}^2) - v_n^f(0;\mathbf{h}^2)\ .
\end{aligned}$$

It follows by induction on n that (19) holds and hence the full-service policy is optimal for the holding-cost function $\mathbf{h}^2$. ∎

We shall presently use this lemma to extend our results on the optimality of the full-service policy from linear to non-linear holding-cost functions. The lemma has some interest in its own right, however. Note first that its conditions require an ordering between the first differences of $\mathbf{h}^1$ and those of $\mathbf{h}^2$ – not the values themselves. In particular, the conditions put no restrictions on $h^1(0)$ and $h^2(0)$; it could be the case that $h^1(0) > h^2(0)$. Indeed, we could have $h^1(i) > h^2(i)$ for all $i \geq 0$ and still satisfy the conditions of the lemma. Conversely, it could be the case that $h^1(i) \leq h^2(i)$ for all $i \geq 0$ but the conditions of the lemma are not satisfied.

Now consider an arbitrary convex, non-decreasing holding-cost function, $\mathbf{h}^2 = \{h^2(i), i \in S\}$. Define the linear holding-cost function $\mathbf{h}^1 = \{h^1(i), i \in S\}$ as follows:

$$h^1(i) = (h^2(1) - h^2(0)) \cdot i\ ,\ i \geq 0\ .$$

Then it follows from the convexity of $\mathbf{h}^2$ that $h^1(i) - h^1(i-1) \leq h^2(i) - h^2(i-1)$ for all $i \geq 1$, so that the conditions of Lemma 2 are satisfied. The following theorem is then a direct consequence of Lemma 2 and Theorem 2.

Theorem 4 *Suppose (6) holds and*

$$\frac{c(\bar{\mu})}{\bar{\mu}} \leq (h^2(1) - h^2(0))/\alpha\ . \tag{20}$$

Then the full-service policy is optimal for the problem with holding-cost function $\mathbf{h}^2$.

References

1. Bell C (1971) Characterization and computation of optimal policies for operating an *M/G/1* queuing system with removable server. Oper. Res. **19** 208–218.
2. Lippman S (1975) Applying a new device in the optimization of exponential queuing systems. Oper. Res. **23** 687–710.
3. Puterman M *Markov Decision Processes: Discrete Stochastic Dynamic Programming.* (Wiley, New York 1994)
4. Sobel M (1982) The optimality of full service policies. Oper. Res. **30** 636–649.
5. Stidham S (1972) $L = \lambda W$: a discounted analogue and a new proof. Oper. Res. **20** 1115–1126.