Probability in the Engineering and Informational Sciences, **13**, 1999, 309–327. Printed in the U.S.A.

BIAS OPTIMALITY IN A QUEUE WITH ADMISSION CONTROL*

MARK E. LEWIS
Faculty of Commerce and Business Administration
University of British Columbia
2053 Main Mall
Vancouver, British Columbia, Canada V6T 1Z2

HAYRIYE AYHAN AND ROBERT D. FOLEY
School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332-0205

We consider a finite capacity queueing system in which each arriving customer offers a reward. A gatekeeper decides based on the reward offered and the space remaining whether each arriving customer should be accepted or rejected. The gatekeeper only receives the offered reward if the customer is accepted. A traditional objective function is to maximize the gain, that is, the long-run average reward. It is quite possible, however, to have several different gain optimal policies that behave quite differently. Bias and Blackwell optimality are more refined objective functions that can distinguish among multiple stationary, deterministic gain optimal policies. This paper focuses on describing the structure of stationary, deterministic, optimal policies and extending this optimality to distinguish between multiple gain optimal policies. We show that these policies are of trunk reservation form and must occur consecutively. We then prove that we can distinguish among these gain optimal policies using the bias or transient reward and extend to Blackwell optimality.

1. INTRODUCTION

We consider a slight generalization of the following. Suppose we add a gatekeeper to an $M/M/n/m$ queueing system, and the gatekeeper decides whether each arriving customer is admitted or rejected. Different customers bring different rewards, and

*This work was partially supported by the NSF-NATO Postdoctoral Fellowship.

© 1999 Cambridge University Press 0269-9648/99 $12.50

https://doi.org/10.1017/S0269964899133047 Published online by Cambridge University Press

the gatekeeper receives the reward only if the customer is accepted. The gatekeeper knows the reward offered by the customer and the number of customers currently in the system. The gatekeeper's objective is either to maximize the long-run average reward received or to achieve a more stringent form of optimality such as bias or Blackwell optimality. Clearly, as the number of customers in the system nears the system capacity $m < \infty$, the gatekeeper would be inclined to reject customers offering smaller rewards, as accepting these customers might prevent the admission of a customer offering a larger reward.

More precisely, we assume that the arriving customers form a Poisson process with arrival rate $\lambda > 0$. The amount of work required by each customer is exponentially distributed with mean one. Each arriving customer independently offers a reward r_k with probability $p_k > 0$ for $k = 1,2,\ldots,\ell$. Thus, the arrival process can be decomposed into ℓ independent Poisson processes with rates $\lambda_k \equiv \lambda p_k$ for $k = 1,\ldots,\ell$, corresponding to the ℓ different rewards offered. In our generalization, we can equivalently consider a single server which works at rate μ_i when the number of customers in the system is i, and we assume that $0 = \mu_0 < \mu_1 \leq \cdots \leq \mu_m < \infty$. Thus, to model a system with n identical servers in which the customer's service times are exponentially distributed with mean $1/\mu$, we simply set $\mu_i = \min\{n, i\}\mu$. Our system has a maximum capacity of $m < \infty$. Hence, customers arriving to a full system are rejected. Customers offering a reward r_k will be called class k customers for $k = 1,\ldots,\ell$, and we assume that $r_1 > r_2 > \cdots > r_\ell > 0$. The gatekeeper, fully cognizant of the system parameters $m, \ell, (r_1,\ldots,r_\ell)$, $(\lambda_1,\ldots,\lambda_\ell)$, and $(\mu_0,\ldots,\mu_m)$, but without any additional knowledge concerning actual service times or on the future of the arrival process, decides for each arriving customer based on the class of the arriving customer and the current number of customers in the system whether to accept or reject the customer. All rejected customers are lost.

In many previous models, the objective was to maximize the long-run average reward. As is standard in the literature, we call a policy that maximizes the long-run average reward gain optimal. However, there is the possibility of multiple gain optimal solutions. In order to decide which policy of the group of gain optimal solutions should be chosen, we turn to a more sensitive optimality criterion in the form of *bias optimality*, which in turn leads us to the even more sensitive criterion of Blackwell optimality. Suppose there are two gain optimal policies that receive exactly the same sequence of rewards except that one policy receives an extra reward at time zero. With respect to gain optimality, the one time reward would be ignored, but with respect to bias optimality, the policy yielding the extra reward would be preferred. We illustrate this with an example.

Example 1: Suppose $m = 4$, $n = 4$, and $l = 2$, so that there is no buffer and there are two customer classes. Let the arrival rate be $\lambda = 0.75$. The probability of arrivals being of class 1 is $p_1 = \frac{2}{3}$ while the probability of class 2 is $p_2 = \frac{1}{3}$. Let the reward of class 1 and class 2 be $r_1 = 1$ and $r_2 = 0.8$, respectively. Finally, let the service rate of each server be 0.0625, so that $\mu_i = 0.0625i$. Suppose we are only concerned with policies that accept both customer classes until the state reaches a level k and then

https://doi.org/10.1017/S0269964899133047 Published online by Cambridge University Press

only accepts class one (since it has the higher reward). Such a policy is called a *trunk reservation policy* with control level k and will be defined precisely later. For this model the gain optimal control level is $k = 3$. That is, we accept both classes as long as the number of customers in the system is two or fewer, and accept only class one if there are three or more busy servers. It is also the case that three is both bias and Blackwell optimal.

Now suppose we keep the same model except let $r_2 = 0.74439$. In this case, we have two gain optimal trunk reservation policies, with control levels two and three. However, it remains that the control level three is both bias and Blackwell optimal, but level two is neither.

One should notice that there were either one or two gain optimal policies and if there were two, the larger was bias and Blackwell optimal. It is no coincidence that the number of gain optimal policies is less than or equal to $\min\{2^{l-1}, 2^{m-1}\}$ while the number of bias optimal policies is one. In fact, this turns out to be true in general for our problem. Hence, this paper is devoted to the proof of the following theorem:

THEOREM 1: *There exists a trunk reservation policy that is gain, bias, and Blackwell optimal. Furthermore, suppose* $(d^*)^\infty$ *is the gain optimal trunk reservation policy with the largest control level for each class. If* $(L^*)_j$ *is the control level of class j under* d^*, $j = 1, \ldots, \ell$, *then among the class of Markovian, stationary, deterministic policies,*

1. *The only other possible gain optimal control level for each class j is* $(L^*)_j - 1$, $j = 1, \ldots, \ell$.
2. $(d^*)^\infty$ *is the only bias optimal policy.*
3. $(d^*)^\infty$ *is the only Blackwell optimal policy.*

Remark 1: In proving Theorem 1 we arrive at some secondary results which are significant as well. Lemma 1 shows that under the assumption of irreducibility, the set of policies which satisfy a set of equations we call the gain optimality equations is independent of the vectors which satisfy the equations. Hence, any policy which satisfies these equations with one set of parameters will satisfy the equations when the parameters of another solution are substituted.

Proposition 3 asserts that there are no gain optimal Markovian, stationary, deterministic policies outside the set of policies which satisfies the gain optimality equations. Hence, we are assured that finding the structure of policies which satisfy these equations characterizes optimal policies.

Also, notice that Theorem 1 implies that, while there are at most $\min\{2^{l-1}, 2^{m-1}\}$ Markovian, stationary, deterministic gain optimal policies, only **one** of these is bias optimal.

The remainder of the paper is organized as follows. Section 2 is a review of related research. In Section 3, we define the action space and the related transition matrix and reward vector. Section 4 is devoted to reviewing Markov decision process theory that will be used throughout the paper. We formulate the problem in

https://doi.org/10.1017/S0269964899133047 Published online by Cambridge University Press

terms of a Markov Decision Process in Section 5. Results pertaining to the gain optimality criterion can be found in Section 6. We conclude with Section 7, the bias optimality results, which are immediately extended to Blackwell optimality.

2. LITERATURE SURVEY

We rely heavily on the paper of Haviv and Puterman [3]. In addition, we will follow closely the notation used in Puterman [7]. Previously, Blackwell [1] has shown that for the discounted case with a finite state and action space, there exists an optimal stationary policy. Upon applying the standard result from Markov decision processes (see Puterman [7, Corollary 8.2.5]),

$$g = \lim_{\alpha \uparrow 1} (1-\alpha) v_\alpha,$$

where g is the gain and v_α is the total discounted reward with discount factor α; Blackwell's result extends to the gain optimality case. In the same paper, Blackwell introduced what is now known as *Blackwell* optimality and showed the existence of a stationary Blackwell optimal policy.

In a classic paper, Miller [6] showed the existence of an optimal *trunk reservation* policy for a finite customer class system with multiple servers and no buffer. This was extended to an uncountable number of customer classes but with one server by Lippman and Ross [5]. Although, the model that we propose does not have an uncountable number of customer classes, we add the generality of a finite buffer and variable, state-dependent service rates. Recently, Feinberg and Reiman [2] added the generality of a constraint on the blocking probability of the highest-paying customers. We also find useful the method of uniformization discussed in Lippman [4]. These problems and others like them were solved using many of the methods of Semi-Markov decision processes. A good source for a discussion of problems similar to these is Stidham [9].

3. SETUP

In order to clarify the model, we make the following assumptions. There are no holding costs associated with customers. As both interarrival times and service times are exponential and thus memoryless, the state of the system can be described by the number of customers in the system and the class of the arriving customer. Since m is the system capacity, the state space $S = \{0,1,\ldots,m\} \times \{0,1,\ldots,\ell\}$, where class zero customers offer zero reward and will correspond to service completions and fictitious transitions in uniformization. For each state $s \in S$ there are associated allowable actions. Thus, we define A_s to be the allowable actions in state $s = (i,c)$ and have,

$$A_{(i,c)} = \begin{cases} \{\text{accept, reject}\} & i \leq m-1,\ c > 0, \\ \{\text{reject}\} & i = m,\ c > 0. \end{cases}$$

https://doi.org/10.1017/S0269964899133047 Published online by Cambridge University Press

Now, define D^∞ to be the class of Markovian, stationary, deterministic policies and let d^∞ denote an arbitrary element in this set. We apply uniformization techniques originally developed by Lippman [4] to discretize the model and use discrete-time methods. To this end, let $\lambda_i = \lambda p_i$, $i = 1, \ldots, \ell$, and let $\Lambda = \mu_m + \sum_{i=1}^{l} \lambda_i$ be the uniformization constant. Without loss of generality, assume that $\Lambda = 1$. Notice, that at each decision epoch, we have an $\{(m+1) \times \ell + 1\}$-dimensional vector d, whose elements describe, when in state s, what decision rule (in A_s) should be applied. That is to say, $d(s) \in A_s$. We order the states lexicographically, so that the first $\ell + 1$ elements are $(0,0), \ldots, (0,\ell)$, the second $\ell + 1$ are $(1,0), \ldots, (1,\ell)$, and so on. Let D denote the set of deterministic decision rules available at each decision epoch. Thus, $d \in D$. Although the model that we propose is a continuous-time Markov Decision Process, we will discretize the model. Hence, we will define the optimality criteria in discrete time.

DEFINITION 1: *The long-run average reward or gain of a policy π given that the system started in state s, denoted $g_\pi(s)$ is (provided the limit exists) given by*

$$g_\pi(s) = \lim_{N \to \infty} E_s \frac{1}{N} \sum_{n=0}^{N} r(X_n, \pi_n(X_n)),$$

where r is the reward function. Further, a policy, π^*, is called gain optimal if

$$g_{\pi^*}(s) \geq g_\pi(s) \quad \text{for all } s \in S$$

for all π.

DEFINITION 2: *The bias of a stationary policy d^∞, given that the system started in state s, denoted $h_d(s)$, is defined to be*

$$h_d(s) = \sum_{n=0}^{\infty} E_s[r(X_n, d(X_n)) - g_d(X_n)].$$

We say that a policy, $(d^*)^\infty$ is bias optimal if it is gain optimal, and

$$h_{d^*}(s) \geq h_d(s) \quad \text{for all } s \in S$$

for all d.

If the Markov chain generated by d^∞ is aperiodic (true for our model) this limit exists. Because any gain optimal policy has the same average reward, we see that the bias may distinguish the best policy by examining the transient reward. Hence, it is often referred to as the transient reward.

4. MARKOV DECISION PROCESS THEORY

We now state a few facts from Markov decision process theory for the average reward objective function. For each of the following results, we assume that both the state and action space are finite and that the rewards are bounded. Both assumptions hold for our model. Recall that the gain of a policy is defined to be the long-run

https://doi.org/10.1017/S0269964899133047 Published online by Cambridge University Press

average reward of that policy. It is well-known that an optimal gain vector can be determined by solving the following set of equations:

$$g = \max_{d \in D} P_d g \tag{1}$$

and

$$h = \max_{d \in G(g)} \{r_d - g + P_d h\} \tag{2}$$

for g and h where $G(g) \subseteq D$ is the set of decision rules that attain the maximum in Eq. (1). We will refer to the system of Eqs. (1) and (2) as *gain optimality equations*. Let $H(g,h) \subseteq G(g)$ be the set of decision rules that attain the maximum in (2). The next result from Puterman [7, Prop. 8.6.1(b)] will be useful later.

Let g_{d^∞} and h_{d^∞} be the gain and bias vectors, respectively, of policy $d^\infty \in D^\infty$.

LEMMA 1: *Suppose* $P_{d'} g_{d^\infty} = P_d g_{d^\infty}$ *for two decision rules* d *and* d' *and that*

$$r_{d'}(s) + P_{d'} h_{d^\infty}(s) > r_d(s) + P_d h_{d^\infty}(s)$$

for a recurrent state s *in the Markov chain generated by* d', *then* $g_{d'^\infty} \geq g_{d^\infty}$ *with strict inequality in at least one component.*

We now note that under the assumption that all states communicate for each policy, there is only one set $H(g,h)$. That is, the decision rules that attain the maximum in (2) do not change as we change g and h.

LEMMA 2: *Suppose all states in* S *communicate for each stationary, deterministic policy, and a decision rule* $d_1 \in D$ *satisfies* (1) *and* (2) (*obtains the argmax*) *with associated vectors* g_1 *and* h_1. *If there exists* $d_2 \in D$ *that satisfies* (1) *and* (2), *then* $d_2 \in H(g_1, h_1)$.

PROOF: We prove this by contradiction using Lemma 1. Suppose there exists a $d_2 \in D$ that satisfies the gain optimality equations, but is not in $H(g_1, h_1)$. Thus, d_2^∞ is gain optimal and, since the chain is irreducible, must have the same gain vector g_1. In fact, the gain vectors are constant vectors. Hence, we know that $P_{d_2} g_1 = P_{d_1} g_1$. Since $d_2 \notin H(g_1, h_1)$, and, since d_1 attains the maximum in (2), we have for some state s

$$r_{d_1}(s) - g_1(s) + P_{d_1} h_1(s) > r_{d_2}(s) - g_1(s) + P_{d_2} h_1(s)$$

or equivalently,

$$r_{d_1}(s) + P_{d_1} h_1(s) > r_{d_2}(s) + P_{d_2} h_1(s).$$

Hence, by Lemma 1 we have that $g_1 > g_2$ which contradicts the optimality of d_2^∞. Thus, $d_2 \in H(g_1, h_1)$, as desired. ■

The following proposition, which appears in Haviv and Puterman [3], will be used to compute the gain and the bias of a policy d^∞.

https://doi.org/10.1017/S0269964899133047 Published online by Cambridge University Press

PROPOSITION 1: *The gain and bias of a policy may be computed by solving the following set of linear equations*

$$\begin{aligned} g &= P_d g, \\ h &= r_d - g + P_d h, \\ w &= -h + P_d w, \end{aligned} \tag{3}$$

the gain, g_d, and bias, h_d, satisfy the first two of the equations and there exists some w which along with h_d satisfies the third. In fact, g_d and h_d are unique.

The next result that appears in Puterman [7, Thm. 10.1.5], essentially gives us the first part of Theorem 1.

PROPOSITION 2:

1. *There exists a stationary, deterministic n-discount optimal policy for $n = -1,0,1,\ldots$.*
2. *Suppose $(d^*)^\infty$ is a Blackwell optimal policy. It is n-discount optimal for $n = -1,0,1,\ldots$.*

Note that (-1)-optimality is equivalent to gain optimality and 0-optimality is equivalent to bias-optimality. Hence, to get the first assertion in Theorem 1, we need only show that these policies can be of trunk reservation form.

The final result from Markov decision process theory that we will discuss is the definition of trunk reservation and how it pertains to our problem. We borrow from many sources the following definition of trunk reservation policies (in particular, see Reiman [8]).

DEFINITION 3: *A trunk reservation decision rule, with reservation levels $m - k_c$, where $0 \le k_c \le m$, $c = 1,\ldots,\ell$, is one that*:

1. *always accepts Class* 1 *customers* (*if the number of customers in the system is less than m*)
2. *accepts Class c customers if and only if there are strictly less than k_c customers in the system.*

We will refer to the level, k_c, at which we begin to reject customers of class c as the control level for class c customers. A policy made up of trunk reservation decision rules is called a trunk reservation policy.

Let D_T be the subset of D that contains only the trunk reservation decision rules. Feinberg and Reiman [2] have examined the model that we propose with the added generality of a constraint on the blocking probability of class one. In their analysis, they reformulate the problem of maximizing the gain as a nonlinear programming problem. By neglecting the constraint, they are able to arrive at the fact that among stationary, deterministic policies, only policies of trunk reservation form can be optimal. We will prove this result for our model using Markov Decision Processes

https://doi.org/10.1017/S0269964899133047 Published online by Cambridge University Press

methods and extend them to bias and Blackwell optimality. Hence, we will be able to restrict our attention to policies in D_T.

5. PROBLEM FORMULATION

In order to apply the above theory, we note that since we have a finite state space and all states in S communicate for each stationary, deterministic policy, the gain vector of any policy in D^∞ is constant. Hence, Eq. (1) is satisfied by all Markovian deterministic decision rules. That is, $G(g) = D$. We will sometimes employ the standard convention of using g as both the constant element of the gain vector as well as the gain vector itself. Moreover, since Lemma 2 implies that $H(g,h)$ does not depend on g and h for the process in question we suppress this dependence and write H instead of $H(g,h)$. In order to simplify the notation, we note that for decision rules $d \in D_T$ it suffices to know the control levels for each customer class. Therefore, we will denote the trunk reservation decision rule (an ℓ-dimensional vector) with control levels L_c for class c customers by an L.

If we write Eq. (2) in component notation, we have that H is the set of rules in D that satisfies for all $0 \le i < m$ (there is but one action if $i = m$),

$$h(i,c) + g = \max\{r_c + U(i+1), U(i)\}, \tag{4}$$

where

$$U(i) = \sum_{j=1}^{\ell} \lambda_j h(i,j) + \mu_i h(i-1,0) + \left[1 - \left(\sum_{j=1}^{\ell} \lambda_j + \mu_i\right)\right] h(i,0). \tag{5}$$

Remark 2: $U(i)$ can be interpreted as the "remaining" bias of the infinite horizon problem. Hence, Eq. (4) has the interpretation that we can receive the immediate reward that the customer offers or forego this reward in hopes of better offers presenting themselves in the future. Using this interpretation, it is easy to see that $h(i,0) = U(i) - g$.

6. GAIN OPTIMALITY RESULTS

We begin this section by showing that in the class of stationary, deterministic policies only trunk reservation policies can be gain optimal. We will need the following lemmas.

LEMMA 3: *Let* $\Delta f(i) \equiv f(i+1) - f(i)$. *Suppose* $d^\infty \in D^\infty$ *is optimal, then we have* $\Delta U_d(i) < 0$ *for* $i = 0, \ldots, m-1$.

PROOF: We prove the assertion by induction. Let $i = 0$. It is clear that if $d(0,c) = \{\text{reject}\}$, for some $c > 0$ we get from the optimality equations $r_c + \Delta U_d(0) \le 0$, from which $\Delta U_d(0) < 0$ follows immediately. Hence, assume that $d(0,c) = \{\text{accept}\}$ for all c. Note that

https://doi.org/10.1017/S0269964899133047 Published online by Cambridge University Press

$$U_d(0) = \sum_{j=1}^{\ell} \lambda_j h_d(0,j) + \left[1 - \sum_{j=1}^{\ell} \lambda_j\right] h_d(0,0)$$
$$= \sum_{j=1}^{\ell} \lambda_j h_d(0,j) + \left[1 - \sum_{j=1}^{\ell} \lambda_j\right](U_d(0) - g_d).$$

Applying the assumption that {accept} is optimal for all classes, we get

$$U_d(0) = \sum_{j=1}^{\ell} \lambda_j (r_j + U_d(1) - g_d) + \left[1 - \sum_{j=1}^{l} \lambda_j\right](U_d(0) - g_d).$$

Hence,

$$\sum_{j=1}^{\ell} \lambda_j \Delta U_d(0) = g_d - \sum_{j=1}^{\ell} \lambda_j r_j.$$

Let α_d denote the stationary distribution of the Markov process generated by the decision rule d. Further, let R_c be the set of states in which class c customers are rejected. Note that

$$g_d = \sum_{j=1}^{l} \sum_{k \in R_j'} \alpha_d(k)(\lambda_j r_j).$$

Thus, $g_d < \sum_{j=1}^{\ell} \lambda_j r_j$ and $\Delta U_d(0) < 0$ as desired.

Now we assume the assertion holds for $i-1$ and show that it is true for i. Recall that

$$U_d(i) = \sum_{j=1}^{\ell} \lambda_j h_d(i,j) + \mu_i h_d((i-1)^+,0) + \left[1 - \left(\sum_{j=1}^{\ell} \lambda_j + \mu_i\right)\right] h_d(i,0).$$

It is again sufficient to consider the case when $d(s) = \{\text{accept}\}$ for all $s \in S$. Hence,

$$\sum_{j=1}^{\ell} \lambda_j \Delta U_d(i) = g_d - \sum_{j=1}^{\ell} \lambda_j r_j + \mu_i \Delta U_d(i-1)$$
$$< 0. \qquad \blacksquare$$

LEMMA 4: *Let* $\Delta^2 f(i) = \Delta(\Delta f)(i) = f(i+2) - f(i+1) - (f(i+1) - f(i))$. *If* $d^\infty \in D^\infty$ *is optimal, then* $\Delta^2 U_d(i) < 0$ *for* $i = 0, \ldots, m-2$.

PROOF: We will also prove this by induction. Note that for $k = 0,1$ we have,

$$g_d = \sum_{j=1}^{\ell} \lambda_j h_d(0,j) - \sum_{j=1}^{\ell} \lambda_j [U_d(0) - g_d]$$
$$g_d = \sum_{j=1}^{\ell} \lambda_j h(1,j) + \mu_1 h(0,0) - \left(\sum_{j=1}^{\ell} \lambda_j + \mu_1\right)[U_d(1) - g_d],$$

https://doi.org/10.1017/S0269964899133047 Published online by Cambridge University Press

which leads to

$$\sum_{j=1}^{\ell} \lambda_j h_d(0,j) - \sum_{j=1}^{\ell} \lambda_j [U_d(0) - g] = \sum_{j=1}^{\ell} \lambda_j h(1,j) + \mu_1 h(0,0) - \left(\sum_{j=1}^{\ell} \lambda_j + \mu_1 \right) [U_d(1) - g_d]$$

or

$$\sum_{j=1}^{\ell} \lambda_j (\Delta h_d(0,j) - \Delta U_d(0)) = \mu_1 \Delta U_d(0). \tag{6}$$

Recall that class 1 customers are always accepted, so that $\Delta h_d(0,1) = \Delta U_d(1)$. We now make some definitions. Let

1. AA = the set of customer classes that are accepted when there are 0 or 1 customers in the system,
2. RR = the set of customer classes that are rejected when there are 0 or 1 customers in the system,
3. AR = The set of customer classes that are accepted when there are 0 customers and rejected when there is 1 customer in the system,
4. RA = The set of customer classes that are rejected when there is one customer and accepted when there are zero customers in the system.

We now get from Equation (6)

$$\begin{aligned}
\mu_1 \Delta U_d(0) &= \sum_{j \in AA} \lambda_j (\Delta h_d(0,j) - \Delta U_d(0)) + \sum_{j \in RR} \lambda_j (\Delta h_d(0,j) - \Delta U_d(0)) \\
&\quad + \sum_{j \in AR} \lambda_j (\Delta h_d(0,j) - \Delta U_d(0)) + \sum_{j \in RA} \lambda_j (\Delta h_d(0,j) - \Delta U_d(0)) \\
&= \sum_{j \in AA} \lambda_j (\Delta^2 U_d(0)) + 0 + \sum_{j \in AR} \lambda_j (-r_j - \Delta U_d(0)) + \sum_{j \in RA} \lambda_j (r_j + \Delta U_d(1)) \\
&= \sum_{j \in AA} \lambda_j (\Delta^2 U_d(0)) + \sum_{j \in AR} \lambda_j (-r_j - \Delta U_d(1) + \Delta^2 U_d(0)) \\
&\quad + \sum_{j \in RA} \lambda_j (r_j + \Delta U_d(1)).
\end{aligned}$$

Hence,

$$\sum_{j \in AA \cup AR} \lambda_j (\Delta^2 U_d(0)) = \mu_1 \Delta U_d(0) + \sum_{j \in AR} \lambda_j (r_j + \Delta U_d(1)) - \sum_{j \in RA} \lambda_j (r_j + \Delta U_d(1)).$$

The second term is nonpositive because it is optimal to reject these classes when there is 1 customer in the system. Similarly for the last term, it is optimal to accept these classes ($r_j + \Delta U_d(1) \geq 0$) when there is 1 customer in the system. Hence, applying Lemma 3 we get $\Delta^2 U_d(0) < 0$.

https://doi.org/10.1017/S0269964899133047 Published online by Cambridge University Press

Now assume the result is true for $k = i - 1$ and consider the state $k = i$. From the definition of $U_d(i)$ we get

$$g_d = \sum_{j=1}^{\ell} \lambda_j h(i,j) + \mu_i h(i-1,0) - \left(\sum_{j=1}^{l} \lambda_j + \mu_i \right)[U_d(i) - g_d],$$

$$g_d = \sum_{j=1}^{\ell} \lambda_j h(i+1,j) + \mu_{i+1} h(i,0) - \left(\sum_{j=1}^{\ell} \lambda_j + \mu_{i+1} \right)[U_d(i+1) - g_d].$$

Hence, we have

$$\sum_{j=1}^{\ell} \lambda_j h(i+1,j) + \mu_{i+1} h(i,0) - \left(\sum_{j=1}^{\ell} \lambda_j + \mu_{i+1} \right)[U_d(i+1) - g_d]$$

$$= \sum_{j=1}^{\ell} \lambda_j h(i,j) + \mu_i h(i-1,0) - \left(\sum_{j=1}^{\ell} \lambda_j + \mu_i \right)[U_d(i) - g_d]$$

or, equivalently,

$$\sum_{j=1}^{\ell} \lambda_j [\Delta h(i,j) - \Delta U_d(i)] = [\mu_{i+1} - \mu_i] \Delta U_d(i) + \mu_i \Delta^2 U_d(i-1).$$

Let $(AA)_i$, $(RR)_i$, $(AR)_i$, and $(RA)_i$ be sets analogous to those in the case when $k = 0$ replacing 0 and 1 with i and $i + 1$, respectively. We get

$$\begin{aligned}
&[\mu_{i+1} - \mu_i] \Delta U_d(i) + \mu_i \Delta^2 U_d(i-1) \\
&\quad = \sum_{j \in (AA)_i} \lambda_j [\Delta h(i,j) - \Delta U_d(i)] + \sum_{j \in (RR)_i} \lambda_j [\Delta h(i,j) - \Delta U_d(i)] \\
&\qquad + \sum_{j \in (AR)_i} \lambda_j [\Delta h(i,j) - \Delta U_d(i)] + \sum_{j \in (RA)_i} \lambda_j [\Delta h(i,j) - \Delta U_d(i)] \\
&\quad = \sum_{j \in (AA)_i} \lambda_j [\Delta^2 U_d(i)] + \sum_{j \in (AR)_i} \lambda_j [-r_j - \Delta U_d(i+1) + \Delta^2 U_d(i)] \\
&\qquad + \sum_{j \in (RA)_i} \lambda_j [r_j + \Delta U_d(i+1)].
\end{aligned}$$

Hence,

$$\begin{aligned}
\sum_{j \in (AA)_i \cup (AR)_i} \lambda_j [\Delta^2 U_d(i)] = &\sum_{j \in (AR)_i} \lambda_j [r_j + \Delta U_d(i+1)] - \sum_{j \in (RA)_i} \lambda_j [r_j + \Delta U_d(i+1)] \\
&+ [\mu_{i+1} - \mu_i] \Delta U_d(i) + \mu_i \Delta^2 U_d(i-1).
\end{aligned}$$

Applying the same reasoning as when $i = 0$ and using the induction hypothesis, the right-hand side is strictly less than zero. Thus $\Delta^2 U_d(s) < 0$. ■

https://doi.org/10.1017/S0269964899133047 Published online by Cambridge University Press

We assert that under the assumption of irreducibility there are not any stationary, deterministic gain optimal policies that do not satisfy the gain optimality equations.

PROPOSITION 3: *Suppose all states in S communicate for each stationary, deterministic policy. Every stationary, deterministic gain optimal policy satisfies the gain optimality equations.*

PROOF: We will prove the proposition by contradiction. Suppose there exists a stationary, deterministic gain optimal policy $(d')^\infty$ that does not satisfy the gain optimality equations. Notice that Proposition 1 guarantees that if a policy satisfies the gain optimality equations with vectors, say g and h, then the same policy also satisfies the gain optimality equations with its corresponding gain and bias vectors g_d and h_d. Hence, let d^∞ be a gain optimal policy that satisfies the gain optimality equations with gain and bias vectors g_d and h_d, respectively. Since $g_{d'} = g_d$ and d' does not satisfy the optimality equations we have $r_{d'}(s) + P_{d'}h_d(s) < r_d(s) + P_d h_d(s)$ for some s which by Lemma 1 implies $g_{d'} < g_d$. This contradicts the optimality of $(d')^\infty$. The result follows. ■

It is not hard to see that, given the concavity of U in i implied by Lemma 4, we have the existence of optimal trunk reservation policies for each control level. Moreover, using Proposition 3, we are able to state a stronger result.

PROPOSITION 4: *In the class of stationary, deterministic policies, only trunk reservation policies are gain optimal.*

PROOF: Suppose d is an optimal decision rule and consider an arbitrary customer class $j \neq 1$. Let $i = i^*$ be the smallest such i such that $d(i,j) = \{\text{reject}\}$. The gain optimality equations yield $r_j + \Delta U_d(i^*) \leq 0$. By the concavity of U_d implied in Lemma 4 we also have that $r_j + \Delta U_d(i') < 0$ for all $i' > i^*$. Hence, $d(i',j) = \{\text{reject}\}$ as well. Since j was arbitrary, this is true for each class. Thus, d is of trunk reservation form. ■

We are now able to rewrite the optimality criterion.

PROPOSITION 5: *Suppose that L^∞ is gain optimal, then for each c*

$$r_c + U_L(i+1) - U_L(i) \begin{cases} \geq 0 & 0 \leq i < L_c, \\ \leq 0 & L_c \leq i \leq m-1. \end{cases}$$

PROOF: Since L is gain optimal we know that it must satisfy the gain optimality equations. The result follows. ■

Remark 3: Intuitively, we can think of $\Delta U(i)$ as the amount it "costs" to add another customer to the system. Thus, if the offered reward, say, r_c, is greater than this, we should accept the customer. Otherwise, we would reject the customer. Hence, we have that the criterion for acceptance is $r_c + \Delta U(i) \geq 0$.

https://doi.org/10.1017/S0269964899133047 Published online by Cambridge University Press

Returning to Example 1 recall that when $r_2 = 0.74439$ we had two gain optimal policies, two and three. In fact, the gain for both policies is 0.213191. However, notice that since

$$h_2 = \{2.44331, 1.81277, 1.12968, 0.385291, -0.467473\}^T$$
$$h_3 = \{2.49891, 1.86837, 1.18528, 0.440894, -0.41187\}^T$$

$h_3 \geq h_2$ componentwise. Hence, there is only one bias optimal trunk reservation policy. In order to understand this better, we note that since there are only two customer classes, there cannot be more than two gain optimal trunk reservation policies. Further, if there are two such policies they must occur consecutively. For notational convenience, let $(L_*)_c$ and $(L^*)_c$ be the smallest and largest gain optimal trunk reservation control levels for customer class c. Clearly, $(L_*)_c \leq (L^*)_c$.

PROPOSITION 6: *Suppose $(L_*)_c$ is the smallest gain optimal control level for customer class c. Then the only other control level for customer class c that can be optimal is $(L_*)_c + 1$.*

PROOF: The assertion is trivially true if $(L_*)_c = m$ or $m - 1$. Thus, assume $(L_*)_c \leq m - 2$. We will consider the case when $(L_*)_c = 0$ separately, so for now assume $(L_*)_c > 0$. We know from Proposition 3 that $L_* \in H$. Note that since U_{L_*} is strictly concave, any state (i,c) such that $(L_*)_c < i < m$, $r_c + U_{L_*}(i+1) - U_{L_*}(i) < 0$ (strictly), which yields $H(i,c) = \{\text{reject}\}$. Finally, we claim that $r_c + U_{L_*}((L_*)_c) - U_{L_*}((L_*)_c - 1) > 0$. To show this suppose it were equal to zero. This would imply that $H((L_*)_c - 1, c) = \{\text{reject, accept}\}$. Thus, $(L_*)_c - 1 \in H((L_*)_c - 1, j)$. This implies that $(L_*)_c - 1$ is gain optimal for customer class c. This contradicts the assumption that $(L_*)_c$ is the smallest gain optimal control level for customer class c. For $(L_*)_c = 0$ notice that by the same reasoning as above, {reject} must be optimal for customer class c for all higher system capacities. The result follows. ■

In the previous proof, we have a characterization of $H(i,c)$ for $i > (L_*)_c$, where $(L_*)_c$ is the smallest optimal trunk reservation decision rule for customer class c. The same logic can be applied to determine $H(i,c)$ for $i \leq (L_*)_c$. Suppose that both $(L_*)_c$ and $(L_*)_c + 1$ are optimal. Thus, we have, $U_{L_*}((L_*)_c + 1) - U_{L_*}((L_*)_c) = -r_c$. Now suppose there exists $i < (L_*)_c$ such that $\{\text{reject}\} \in H(i,c)$. Since H is independent of g and h we have $U_{L_*}(i+1) - U_{L_*}(i) = -r_c$. This contradicts Lemma 4. Hence, $H(i,c) = \{\text{accept}\}$ when $i < L_*$. If both $(L_*)_c$ and $(L_*)_c + 1$ are optimal we know that optimal policies which use either control level for class c customers are in H and, thus, we can admit or not admit class c customers in state L_*. So we have proven the following proposition that completely characterizes H.

PROPOSITION 7: *If $(L^*)_c$ and $(L^*)_c - 1$ are gain optimal trunk reservation decision rules for customer class c,*

$$H(i,c) = \begin{cases} \{\text{accept}\} & 0 \leq i < (L^*)_c - 1, \\ \{\text{accept, reject}\} & i = (L^*)_c - 1, \\ \{\text{reject}\} & (L^*)_c - 1 < i \leq m, \end{cases}$$

https://doi.org/10.1017/S0269964899133047 Published online by Cambridge University Press

and if $(L^*)_c$ is the only gain optimal trunk reservation rule for customer class c,

$$H(i,c) = \begin{cases} \{\text{accept}\} & 0 \le i < (L^*)_c, \\ \{\text{reject}\} & (L^*)_c \le i \le m. \end{cases}$$

7. BIAS OPTIMALITY RESULTS

The previous results have reduced our search for stationary, deterministic bias optimal policies to a much smaller set. Since it is well known that such policies exist (see Proposition 2) we now concern ourselves with which policy, if two gain optimal policies exist, is bias optimal. Here again, we restate a useful result (slightly modified) from Haviv and Puterman [3].

PROPOSITION 8: *Suppose g and h satisfy the gain optimality equations and that there exists a vector w for which*

$$w = \max_{d \in H} \{-h + P_d w\}, \tag{7}$$

then, if $d^* \in H$ attains the maximum in (7), $(d^*)^\infty$ is bias optimal. Moreover, suppose the process is unichain and that for some $d \in H$, where g and h satisfy the gain optimality equations, there exists a vector w with

$$w = -h + P_d w. \tag{8}$$

If

$$P_{d'} w \ge P_d w \tag{9}$$

for some $d' \in H$ with strict inequality in at least one recurrent state in the Markov chain generated by d', then $h^{(d')^\infty} \ge h^{(d)^\infty}$ with strict inequality in at least one component.

Next, we will show that if we have two gain optimal (trunk reservation) control levels for a particular customer class *c*, only the larger one is bias optimal. That is, bias optimality is sensitive enough to distinguish between them.

PROPOSITION 9: *If gain optimal trunk reservation policies d^∞ and $(d')^\infty$ are identical except for customer class j, where d^∞ uses control level L_j and $(d')^\infty$ uses control level $L_j + 1$, then $h_{d'} \ge h_d$ with strict inequality in at least one element.*

PROOF: Let g_d and h_d satisfy the gain optimality equations. Assume for now that $L_j > 0$. Suppose w_d is a vector that satisfies $w = -h_d + P_d w$. We will show that $P_{d'} w_d \ge P_d w_d$ with strict inequality in at least one component and apply the second part of Proposition 8. Recall that we have ordered the states lexicographically. It is clear that, as $P_{d'}$ and P_d only differ in the (L_j, j)th row, we need only consider that component. Hence, we need to show that

$$(P_{d'} w_d)(L_j, j) > (P_d w_d)(L_j, j),$$

https://doi.org/10.1017/S0269964899133047 Published online by Cambridge University Press

or equivalently,

$$\sum_{k=1}^{\ell} \lambda_k w_d(L_j+1,k) + \mu_{L_j+1} w_d(L_j,0) + \left[1 - \left(\sum_{k=1}^{\ell} \lambda_k + \mu_{L_j+1}\right)\right] w_d(L_j+1,0)$$
$$> \sum_{k=1}^{\ell} \lambda_k w_d(L_j,k) + \mu_{L_j} w_d(L_j-1,0) + \left[1 - \left(\sum_{k=1}^{\ell} \lambda_k + \mu_{L_j}\right)\right] w_d(L_j,0).$$

To this end, we derive the following set of equations for an arbitrary class $c \geq 1$ with control level L_c under d^∞,

$$-\Delta h_d(i,c) = \begin{cases} \Delta w_d(i,c) - \mu_{i+1}\Delta w_d(i,0) - \sum_{k=1}^{\ell} \lambda_k \Delta w_d(i+1,k) & \\ \quad - \left[1 - \left(\sum_{k=1}^{\ell} \lambda_k + \mu_{i+2}\right)\right] \Delta w_d(i+1,0) & i < L_c - 1 \\ \Delta w_d(i,c) & i = L_c - 1 \\ \Delta w_d(i,c) - \mu_i \Delta w_d(i-1,0) - \sum_{k=1}^{\ell} \lambda_k \Delta w_d(i,k) & \\ \quad - \left[1 - \left(\sum_{k=1}^{\ell} \lambda_k + \mu_{i+1}\right)\right] \Delta w_d(i,0) & L_c \leq i \leq m-1. \end{cases} \tag{10}$$

Using (8) we know for $i < L_c - 1$,

$$w_d(i,c) = -h_d(i,c) + \sum_{k=1}^{\ell} \lambda_k w_d(i+1,k) + \mu_{i+1} w_d(i,0)$$
$$+ \left[1 - \left(\sum_{k=1}^{\ell} \lambda_k + \mu_{i+1}\right)\right] w_d(i+1,0),$$

$$w_d(i+1,c) = -h_d(i+1,c) + \sum_{k=1}^{\ell} \lambda_k w_d(i+2,k) + \mu_{i+2} w_d(i+1,0)$$
$$+ \left[1 - \left(\sum_{k=1}^{\ell} \lambda_k + \mu_{i+2}\right)\right] w_d(i+2,0).$$

Hence,

$$\Delta w_d(i,c) = -\Delta h_d(i,c) + \sum_{k=1}^{\ell} \lambda_k \Delta w_d(i+1,k) - \mu_{i+2}\Delta w_d(i+1,0)$$
$$+ \mu_{i+1}\Delta w_d(i,0) + \left[1 - \sum_{k=1}^{\ell} \lambda_k\right] \Delta w_d(i+1,0),$$

https://doi.org/10.1017/S0269964899133047 Published online by Cambridge University Press

which yields, for $i < L_c - 1$

$$-\Delta h_d(i,c) = \Delta w_d(i,c) - \mu_{i+1}\Delta w_d(i,0) - \sum_{k=1}^{\ell} \lambda_k \Delta w_d(i+1,k) - \left[1 - \left(\sum_{k=1}^{\ell} \lambda_k + \mu_{i+2}\right)\right]\Delta w_d(i+1,0).$$

Now consider when $i = L_c - 1$

$$w_d(L_c-1,c) = -h_d(L_c-1,c) + \sum_{k=1}^{\ell} \lambda_k w_d(L_c,k) + \mu_{L_c} w_d(L_c-1,0) + \left[1 - \left(\sum_{k=1}^{\ell} \lambda_k + \mu_{L_c}\right)\right] w_d(L_c,0),$$

$$w_d(L_c,c) = -h_d(L_c,c) + \sum_{k=1}^{\ell} \lambda_k w_d(L_c,k) + \mu_{L_c} w_d(L_c-1,0) + \left[1 - \left(\sum_{k=1}^{\ell} \lambda_k + \mu_{L_c}\right)\right] w_d(L_c,0).$$

So, we have

$$\Delta w_d(L_c-1,c) = -\Delta h_d(L_c-1,c).$$

Now consider $L_c \le i \le m-1$,

$$w_d(i,c) = -h_d(i,c) + \sum_{k=1}^{\ell} \lambda_k w_d(i,k) + \mu_i w_d(i-1,0) + \left[1 - \left(\sum_{k=1}^{\ell} \lambda_k + \mu_i\right)\right] w_d(i,0),$$

$$w_d(i+1,c) = -h_d(i+1,c) + \sum_{k=1}^{\ell} \lambda_k w_d(i+1,k) + \mu_{i+1} w_d(i,0) + \left[1 - \left(\sum_{k=1}^{\ell} \lambda_k + \mu_{i+1}\right)\right] w_d(i+1,0).$$

Thus,

$$-\Delta h_d(i,c) = \Delta w_d(i,c) - \mu_i \Delta w_d(i-1,0) - \sum_{k=1}^{\ell} \lambda_k \Delta w_d(i,k) - \left[1 - \left(\sum_{k=1}^{\ell} \lambda_k + \mu_{i+1}\right)\right]\Delta w_d(i,0).$$

https://doi.org/10.1017/S0269964899133047 Published online by Cambridge University Press

Now, for class 0 we have,

$$w_d(i,0) = -h_d(i,0) + \sum_{k=1}^{\ell} \lambda_k w_d(i,k) + \mu_i w_d(i-1,0) + \left[1 - \left(\sum_{k=1}^{\ell} \lambda_k + \mu_i\right)\right] w_d(i,0),$$

$$w_d(i+1,0) = -h_d(i+1,0) + \sum_{k=1}^{\ell} \lambda_k w_d(i+1,k) + \mu_{i+1} w_d(i,0) + \left[1 - \left(\sum_{k=1}^{\ell} \lambda_k + \mu_{i+1}\right)\right] w_d(i+1,0).$$

Hence,

$$-\Delta h_d(i,0) = \Delta w_d(i,0) - \mu_i \Delta w_d(i-1,0) - \sum_{k=1}^{\ell} \lambda_k \Delta w_d(i,k) - \left[1 - \left(\sum_{k=1}^{\ell} \lambda_k + \mu_{i+1}\right)\right] \Delta w_d(i,0).$$

Since $\sum_{k=1}^{\ell} \lambda_k + \mu_m = 1$, (10) can be rewritten in matrix notation as

$$(I - Q)\Delta w_d = -\Delta h_d \tag{11}$$

where Q is defined by

$$Q((i_1,c),(i_2,c')) = \begin{cases} \mu_{i_1} & c = 0,\ i_2 = i_1 - 1,\ c' = 0, \\ \lambda_{c'} & c = 0,\ i_2 = i_1,\ c' \neq 0, \\ \mu_m - \mu_{i_1+1} & c = 0,\ i_2 = i_1,\ c' = 0, \\ \mu_{i_1+1} & i_1 < L_c,\ c \neq 0,\ i_2 = i_1,\ c' = 0, \\ \lambda_{c'} & i_1 < L_c,\ c \neq 0,\ i_2 = i_1 + 1,\ c' \neq 0, \\ \mu_m - \mu_{i_1+2} & i_1 < L_c,\ c \neq 0,\ i_2 = i_1 + 1,\ c' = 0, \\ \mu_{i_1} & L_c \leq i_1 \leq m-1,\ c \neq 0,\ i_2 = i_1 - 1,\ c' = 0, \\ \lambda_{c'} & L_c \leq i_1 \leq m-1,\ c \neq 0,\ i_2 = i_1,\ c' \neq 0, \\ \mu_m - \mu_{i_1+1} & L_c \leq i_1 \leq m-1,\ c \neq 0,\ i_2 = i_1,\ c' = 0, \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

Consider state (0,0). Since we have assumed that $\mu_0 = 0$ and $\mu_1 > 0$, this row of Q cannot sum to 1. Q can be thought of as the transition matrix for the transient states

https://doi.org/10.1017/S0269964899133047 Published online by Cambridge University Press

of a finite-state, Markov chain. Hence, we get from Puterman [7, App. A, Prop. A.3] that $I - Q$ is invertible, with

$$(I - Q)^{-1} = \sum_{n=0}^{\infty} Q^n.$$

Since Q has all nonnegative elements, $(I - Q)^{-1}$ must also be nonnegative. Furthermore, from (11),

$$\begin{aligned} \Delta w_d &= -(I - Q)^{-1} \Delta h_d \\ &= -\sum_{n=0}^{\infty} Q^n \Delta h_d. \end{aligned} \tag{13}$$

Now, from Lemma 3 we know that $-\Delta h_d > 0$. Further, note that the first two terms of the sum in Eq. (13) are $-\Delta h_d - Q\Delta h_d$, and that Q is nonnegative with at least one strictly nonnegative element in each row. Hence,

$$\begin{aligned} \Delta w_d &= -\sum_{n=0}^{\infty} Q^n \Delta h_d \\ &> -\Delta h_d. \end{aligned}$$

To complete the proof, recall

$$\Delta w_d(L_*, j) + \Delta h_d(L_*, j) = (P_{d'} w_d)(L_*, j) - (P_d w_d)(L_*, j) > 0.$$

Applying the second part of Proposition 8, we obtain the result. if $L_* = 0$ the analysis is similar. ■

The previous proposition leads to the following corollary.

COROLLARY 1: *There can only be one stationary, deterministic bias optimal policy and that policy uses the highest gain optimal control limit for each customer class.*

Of course, there are other more sensitive optimality criterion including $\infty -$ *optimality* or *Blackwell optimality.*

DEFINITION 4: *A policy π^* is called Blackwell optimal if there exists λ^* such that $0 \le \lambda^* < 1$, $v_\lambda^{\pi^*} \ge v_\lambda^{\pi}$, for all π and $\lambda^* \le \lambda < 1$, where v_λ^{π} is the total discounted reward of the Markov decision process when using policy π with discount rate λ.*

PROOF OF THEOREM 1: The existence of gain, bias, and Blackwell optimal policies is a direct consequence of Proposition 2 and the fact that Blackwell [1] showed the existence of a stationary policy (in the finite state case). Applying Proposition 6 yields that there can be at most two gain optimal control levels for each customer class, and that they must occur consecutively. Further, Proposition 9 distinguishes between two gain optimal control levels using bias optimality. Finally, it is well known that Blackwell optimality implies bias optimality. Hence, we have shown that, if there is only one gain optimal control level for each class, that policy is also

https://doi.org/10.1017/S0269964899133047 Published online by Cambridge University Press

Blackwell optimal, and, if there are classes in which there are two gain optimal control levels, the higher one is Blackwell optimal. ■

Acknowledgments

We would like to thank Dr. Martin I. Reiman of Bell Laboratories, for suggesting the model to us and for all of his help and support. We would also like to thank Prof. Martin L. Puterman for his editorial comments.

References

1. Blackwell, D. (1962). Discrete dynamic programming. *Annals of Mathematical Statistics* 33: 719–726.
2. Feinberg, E. & Reiman, M. (1994). Optimality of randomized trunk reservation. *Probability in the Engineering and Informational Sciences* 8: 463–489.
3. Haviv, M. & Puterman, M.L. (1998). Bias optimality in controlled queueing systems. *Journal of Applied Probability* 35: 136–150.
4. Lippman, S.A. (1975). Applying a new device in the optimization of exponential queueing systems. *Operations Research* 23(4): 687–712.
5. Lippman, S.A. & Ross, S.M. (1971). The streetwalker's dilemma: A job shop model. *SIAM Journal of Applied Mathematics* 20(3): 336–342.
6. Miller, B.L. (1969). A queueing reward system with several customer classes. *Management Science* 16(3): 234–245.
7. Puterman, M.L. (1994). *Markov decision processes: Discrete stochastic dynamic programming*. New York: John Wiley and Sons, Inc.
8. Reiman, M.I. (1989). Asymptotically optimal trunk reservation for large trunk groups. In *Proceedings of the 28th Conference on Decision and Control*. *IEEE* 2536–2541.
9. Stidham, S., Jr. (1985). Optimal control of admission to a queueing system. *IEEE Transactions on Automatic Control* AC-30(8): 705–713.

https://doi.org/10.1017/S0269964899133047 Published online by Cambridge University Press