Bias Optimal Admission Control Policies for a Multi-Class Non-Stationary Queueing System[1]

Mark E. Lewis
Department of Industrial and Operations Engineering
University of Michigan, 1205 Beal Avenue, Ann Arbor, MI 48109-2117
melewis@engin.umich.edu (734) 763-0519

Hayriye Ayhan and Robert D. Foley
School of Industrial and Systems Engineering
Georgia Institute of Technology, Atlanta, Georgia 30332-0205
hayhan@isye.gatech.edu (404) 894-2308
rfoley@isye.gatech.edu (404) 894-2309

submitted July 9, 2001, revised January 17, 2002

[1]Subject classifications: Dynamic programming, applications: Sensitive optimality. Queues, nonstationary. Optimal control: monotone policies

Abstract

We consider a finite capacity queueing system where arriving customers offer rewards which are paid upon acceptance into the system. The gatekeeper, whose objective is to "maximize" rewards decides if the reward offered is sufficient to accept or reject the arriving customer. Suppose the arrival rates, service rates, and system capacity are changing over time in a known manner. We show that all bias optimal (a refinement of long-run average reward optimal) policies are of threshold form. Furthermore, we give sufficient conditions for the bias optimal policy to be monotonic in time. We show, via a counterexample, that if these conditions are violated, the optimal policy may not be monotonic in time or of threshold form.

1 Introduction

Suppose customers arrive to a service system with limited capacity and request entry from the gatekeeper. The gatekeeper, knowing the current number of customers in the system and the reward offered by the current arrival, must decide whether to accept or reject the arrival. It is reasonable to suspect that the gatekeeper's entrance requirements become more stringent as the number of customers in the system increases. Of course, when the system is full, the gatekeeper rejects all arrivals. After receiving service, the customers depart from the system freeing up space for new arrivals.

Although we are primarily interested in non-stationary models, we first describe the *stationary admission control problem*. Assume that the system has a maximum capacity of m customers. The customers arrive according to a marked Poisson process with rate λ. The mark associated with each arrival is the amount that the customer is willing to pay. Suppose that there are ℓ customer classes. The marks are independent, discrete random variables and equal to r_j with probability λ_j/λ for $j = 1, 2, \ldots, \ell$. Thus, the customers willing to pay r_j, which we sometimes refer to as Class j customers, form a Poisson process with arrival rate λ_j. Without loss of generality, we assume that $r_1 > r_2 > \cdots > r_\ell > 0$. The customers' service requirements are i.i.d. exponentially distributed random variables with mean 1. When there are i customers in the system, the server works at rate μ_i for $i = 0, 1, \ldots, m$. Let the service rates be such that $\mu_0 = 0$ and $\mu_1 > 0$. We assume that μ_i is non-decreasing and concave in i; that is,

$$\mu_i \leq \mu_{i+1} \text{ for } i = 0, \ldots, m-1, \tag{1}$$

$$\Delta\mu_{i+1} \leq \Delta\mu_i \text{ for } i = 0, \ldots, m-2, \tag{2}$$

where for a real-valued function $f(i)$, $\Delta f(i) \equiv f(i+1) - f(i)$ for $i = 0, 1, \ldots, m-1$. Note that a system with s identical exponential servers is simply a special case with $\mu_i = \mu \min\{i, s\}$. Later, we will apply Δ to functions of more than one variable, but it will always be applied to the first variable, which will represent the number of customers. For example, $\Delta f(i, j) \equiv f(i+1, j) - f(i, j)$.

The goal of this paper is to analyze the following *non-stationary admission control problem* in which the system capacity, and the arrival and service rates may change over time. More precisely, let time 0 be the initial time, and let $m(t)$ be the capacity of the system at time t for $t \geq 0$. Let $\hat{m} \equiv \sup_{t\geq 0} m(t) \leq \infty$. Unless otherwise specified, we assume that $m(\cdot)$ is non-decreasing. There is still only one server, and we let $\mu_i(t)$ denote the service rate at time t when there are i customers in the system, $i = 0, \ldots, \hat{m}$. We assume that for each t, $\mu_i(t)$ is a non-decreasing, concave function of i, and for each i, $\mu_i(\cdot)$ is bounded and measurable. Of course, $\mu_0(t) \equiv 0$ for all t. Let $\lambda_j(t)$ denote the arrival rate of Class j customers at time t. For each j, $\lambda_j(\cdot)$ is also assumed to be bounded and measurable. Hence, we have $\Psi \equiv \sup_{t\geq 0}\{\mu_{\hat{m}}(t) + \sum_{j=1}^{\ell} \lambda_j(t)\} < \infty$. A Class j customer offers reward r_j for $j = 1, 2, \ldots, \ell$ with $r_1 > r_2 > \cdots > r_\ell$. We assume that there is some finite time $T \geq 0$ after which the arrival, service, and capacity processes are constant. Thus, we can define,

$$m^\infty \equiv m(t) > 0 \text{ for } t \geq T, \tag{3}$$

$$\mu_i^\infty \equiv \mu_i(t) > 0 \text{ for } t \geq T \text{ and } i = 1, 2, \ldots, \hat{m}, \tag{4}$$

$$\lambda_j^\infty \equiv \lambda_j(t) > 0 \text{ for } t \geq T \text{ and } j = 1, 2, \ldots, \ell. \tag{5}$$

It is simple to construct examples in which a long-run average or *gain* optimal policy is arbitrarily bad for some large, but finite, length of time. For example, we could reject all customers until some fixed time $T' < \infty$ and then use a gain optimal policy for $t \geq T'$ and the policy remains gain optimal. This *underselective* nature of the long-run average optimality criterion is alleviated by choosing more sensitive optimality criteria.

The stationary problem described above is solved in Lewis, et. al [11]. Here we extend those results. First, we show that the existence of optimal policies that are monotone in the number of customers in the system continues to hold for the non-stationary problem for *any* bounded, measurable arrival and service rate functions. Furthermore, under some monotonicity assumptions on these rates, we can also get monotonicity of the optimal policies in time. This considerably reduces the search for optimal policies to an easily computable set.

We will be primarily interested in *bias optimality*, which can be considered as a refinement of gain optimality. Haviv and Puterman [5] and Lewis et. al [11] used bias optimality to distinguish among gain optimal policies in controlling stationary queueing systems. Furthermore, in the class of deterministic, stationary policies, Hernández-Lerma and Lasserre (see Theorem 10.3.11) have given conditions for bias optimality, *opportunity cost optimality, weakly overtaking optimality*, and *optimality in the sense of Dutta [2]*, to be equivalent. Since our model satisfies these conditions, our results hold for each of the above criterion.

Threshold policies are often used to control access to finite capacity queueing systems. For example in the telecommunications literature the type of policy considered here might be called a *trunk reservation policy*; in a multi-server system with limited (or no) waiting room space is *reserved* for high class customers (see e.g. [20]). Typically, threshold policies are defined to have constant control levels; however, we allow time dependent control levels. This is an example of a monotone policy since the admittance rule is monotonic in the number of customers in the system.

Definition 1 *A threshold policy with threshold levels $m(t) - k_j(t)$ accepts Class j customers if and only if there are strictly less than $k_j(t)$ customers in the system at time t for $j = 1, 2, \ldots, \ell$. We refer to the level, $k_j(t)$, at which we begin to reject customers of Class j as the control level for Class j customers at time t.*

Despite the fact that it seems quite intuitive that an optimal threshold policy exists, in Section 7 we present a counter-example in which this does not hold.

Let $\lambda(t)$, $\mu(t)$, and $k(t)$ denote the vectors of arrival rates, service rates, and control levels at time t. Define $\Delta\mu(t)$ to be the vector with i^{th} component $\Delta\mu_i(t)$ for $i = 0, \ldots, \hat{m} - 1$. A vector-valued function of t is said to be non-decreasing (non-increasing) in t if each component is non-decreasing (non-increasing). Our main results are summarized in the following theorem.

Theorem 1 *For the non-stationary admission control problem, the following hold:*

1. *All bias optimal policies are threshold policies. Furthermore, bias optimal policies are constant on $[T, \infty)$.*

2. *If $\mu(t)$, and $\Delta\mu(t)$ are non-decreasing and $\lambda(t)$ is non-increasing in t, then all bias optimal threshold policies have control levels $k(t)$ non-decreasing in t and constant on $[T, \infty)$.*

3. *If $m(t)$ is constant, $\mu(t)$ and $\Delta\mu(t)$ non-increasing, and $\lambda(t)$ is non-decreasing in t, all bias optimal threshold policies have control levels $k(t)$ non-increasing in t and constant on $[T, \infty)$.*

The asymmetry between the two monotonicity results will be justified by a counterexample given in the last section. The counterexample satisfies the assumptions of the latter monotonicity result except $m(t)$ is non-increasing (recall our standing assumption that $m(t)$ is non-decreasing). Not only is the optimal policy not monotonic, it is not even of threshold form!

Note that the second and third results encompass the case that the service or arrival rates are monotonically changing while the other is constant. This may be seen in peak call times in telecommunications systems or when servers are interchanged during "off-peak" times. Furthermore, this is different from what was previously considered since all of the literature that we are aware of focuses only on a non-stationary arrival process (cf. [3, 17, 18]).

The rest of the paper is organized as follows: Section 2 contains a survey of related literature. The problem is formulated as a Markov Decision Process in Section 3. We give formal definitions of the optimality criteria gain and bias in Section 4. Section 5 is devoted to proving the optimality of threshold policies, while sufficient conditions for the monotonicity of such policies are given in Section 6. We illustrate our results with examples in Section 7.

2 Literature Survey

Threshold policies were originally reported by Yadin and Naor [24] for a server with vacations and proved to be optimal by Heyman [8]. This was later applied to a finite customer class, finite capacity, queueing system by Miller [19] and continued by Lippman and Ross [16] for a single-server with uncountable customer classes. Helm and Waldmann [6] showed in a general setting if enough information is included in the state space the existence of control limit policies is still guaranteed. A good source for these models is the survey paper of Stidham [21]. More recently, the idea of a threshold policy has been used in several papers in which the state space is multi-dimensional. The intuition is the same in that the policies are still monotone, but it is then termed a *monotone switching curve*, see for example [1, 4, 10].

By contrast, non-stationary queueing systems have not received much attention in the literature. Massey and Whitt [17] analyzed a non-stationary $M/M/s/s$ system and discussed an alternative method of approximating loss probabilities. The same authors also have made interesting insights as to when congestion is "felt" in finite capacity queueing systems when the arrival rates are slowly changing over time [18]. For more on these topics see the references cited in those papers. One will note that Massey and Whitt's work is basically focused on analyzing systems while we focus on *optimization*. The only work that we know of that considers optimization of a non-stationary queueing system is the recent work of Fu et al. [3]. In [3] the authors consider a server staffing problem in which decisions are made periodically to change the staffing level. The arrival process is assumed non-stationary and a finite horizon problem is considered. We allow for non-stationary arrival and service processes as well as allowing the system capacity to change. Furthermore, we prove monotonicity of the optimal policy in the number of customers in the system *and* in time.

The study of bias as a theoretical construct began with Veinott [22] and was continued in an expository chapter by the same author [23]. The recent paper by Haviv and Puterman [5] rekindled interest in the bias and explains how bias can be used to alleviate the problem of multiple gain optimal policies in a stationary single server, single class model. Lewis, et al. [11] discussed a stationary version of the model we consider, and related results for both the stationary and non-stationary versions appear in [9]. Bias optimality on a finite state and action space is considered in Lewis and Puterman [13] while the extension of bias to general state and action space models is discussed in Chapter 10 of Hernández-Lerma and Lasserre [7]. The question of the relationship of bias to implicit discounting was considered by Lewis and Puterman [12, 14].

3 Markov Decision Process Formulation

We formulate our model as a stationary, discrete-time Markov Decision Process as follows. We embed at a random sequence of times $\tau_0 < \tau_1 < \ldots$ Although decisions only need to be made at times when a customer arrives and the system is not full, it will be convenient to embed more frequently than just at arrival times. Since $\Psi < \infty$, we can uniformize the process (see Lippman [15]) so that there is a sequence of times forming a Poisson process with rate Ψ, and all arrival and departure times are a subset of this Poisson process. In addition, we include times 0 and T so that $\tau_0 = 0$ and there exists some random N with

$\tau_{N+1} = T$, almost surely, where N is a Poisson random variable with mean ΨT. Note that N is the number of embedded transitions before reaching time T.

The state at time τ_n will be denoted by X_n. The states are ordered triples (i, j, z). The first coordinate i gives the number of customers in the system before processing the transaction; i.e., for $\tau_n < \infty$, the number of customers just prior to τ_n. The middle coordinate denotes the type of transaction being attempted at time τ_n. If $j \in \{1, 2, \ldots, \ell\}$, a customer offering reward r_j is requesting admission. If $j = -1$, a customer is departing, and $j = 0$, is for "dummy" transitions due to uniformization. The last coordinate $z \equiv \min(\tau_n, T)$. Thus, the state space of $X_0, X_1, \ldots$ is $S = \{0, \ldots, \hat{m}\} \times \{-1, 0, 1, \ldots, l\} \times [0, T]$.

The set of actions available in state $s = (i, j, z) \in S$ is

$$A_s = \begin{cases} \{0, 1\} & \text{if } i < m(z), j > 0, \\ \{0\} & \text{if } i \geq m(z), j > 0, \\ \{j\} & \text{if } j \leq 0 \end{cases}$$

where the action can be interpreted as the resulting change in the queue length. The only exception to this interpretation is that if action -1 were to be taken in state $s = (0, -1, z)$, the number of customers in the system would remain zero. Thus, if the state at embedded time t is $s = (i, j, z)$ and action $a \in A_s$ is selected, the number of customers in the system just after t is $(i + a)^+$. Furthermore, the reward $r(s, a)$ obtained at t for taking action a in state $s = (i, j, z)$ is r_j if $j > 0$ and $a = 1$; otherwise, $r(s, a) = 0$.

To complete the description of the Markov decision process, we give the transition kernel from state $s = (i, j, z)$ under action a to any Borel set $B \subset S$

$$Q(B \mid s, a) = \sum_{i'=0}^{\hat{m}} \sum_{j'=-1}^{\ell} \left[I_B(i', j', T) q_d(i', j' \mid s, a) + \int_0^{T-z} I_B(i', j', z + v) q_c(i', j', v \mid s, a) \Psi e^{-\Psi v} \, dv \right]$$

where I_B is the indicator function of B,

$$q_d(i', j' \mid s, a) = \begin{cases} \lambda_{j'}^{\infty} / \Psi & i' = i + a, j' > 0, z = T, \\ \mu_{i'}^{\infty} / \Psi & i' = i + a, j' = -1, z = T, \\ [\Psi - \mu_{i'}^{\infty} - \sum_{j'=1}^{\ell} \lambda_{j'}^{\infty}] / \Psi & i' = i + a, \ j' = 0, z = T, \\ e^{-\Psi(T-z)} & i' = i + a, j' = 0, z < T, \\ 0 & \text{otherwise}, \end{cases}$$

and

$$q_c(i', j', v \mid s, a) = \begin{cases} \lambda_{j'}(z+v)/\Psi & i' = i+a, j' > 0, z < z+v < T, \\ \mu_{i'}(z+v)/\Psi & i' = i+a, j' = -1, z < z+v < T, \\ [\Psi - \mu_{i'}(z+v) - \sum_{j'=1}^{\ell} \lambda_{j'}(z+v)]/\Psi & i' = i+a, j' = 0, z < z+v < T, \\ 0 & \text{otherwise.} \end{cases}$$

4 Gain and Bias Optimality

In this section we formalize the definitions of gain and bias (similar to AC-optimal and bias in Hernández-Lerma and Lasserre [7] except rewards instead of costs) and discuss some related ideas. Let X_k be the state of the system at stage k and d_k be the decision rule at stage k under a particular policy π. The $k-$stage expected reward under policy π given $X_0 = s$ is

$$J_k(\pi, s) \equiv E_s^\pi \left[\sum_{n=0}^{k-1} r(X_n, d_n(X_n)) \right].$$

The *gain* under policy π starting from state s is

$$g(\pi, s) = \liminf_{k \to \infty} J_k(\pi, s)/k.$$

Hence, the *optimal gain* is

$$g^*(s) = \sup_{\Pi} g(\pi, s)$$

where Π is the set of all (non-anticipating) policies. A policy π^* is called *gain optimal* if $g(\pi^*, s) = g^*(s)$ for all $s \in S$.

Several assumptions will need to be satisfied to guarantee the existence of bias optimal policies. In the interest of saving space we will not reiterate these assumptions here except to state that the non-stationary admission control problem as formulated in Section 3 satisfies Assumptions 10.2.1, 10.2.2 (with $w \equiv 1$), and 10.3.5 (with λ being counting measure on $S_T \equiv \{s = (i, j, z) \in S \mid z = T\}$) of Hernández-Lerma and Lasserre [7]. We refer to these technical assumptions collectively as the HL-Assumptions. More important to our analysis are several results from Hernández-Lerma and Lasserre [7], which follow from these assumptions, that will be stated for completeness.

Let d^∞ denote a stationary, deterministic policy that uses the deterministic decision rule d at all decision epochs, and let $g_d(s)$ denote the gain under d^∞ starting from state s. Under the HL-Assumptions for any stationary, deterministic policy d^∞, there exists a constant $g_d = g_d(s)$ for all $s \in S$ and a bounded function h referred to as the *bias* with

$$h(d,s) = \sum_{k=0}^{\infty} E_s^{d^\infty}[r(X_k, d(X_k)) - g_d]. \tag{6}$$

Let G denote the set of all stationary, deterministic decision rules d such that d^∞ is gain optimal; i.e., such that $g_d(s) = g^*(s)$. The optimal bias function is $h^*(s) \equiv \sup_G h(d,s)$. A stationary, deterministic policy $(d^*)^\infty$ is called *bias optimal* if $d \in G$ and

$$h(d^*, s) = h^*(s)$$

for all $s \in S$.

We next make a few comments about bias optimality. Note that the bias is only defined for stationary policies, but we are considering a non-stationary system. To resolve this, we have included the current time in the state space. To understand what the bias represents, notice that (6) implies that bias optimal is equivalent to maximizing the total expected reward in a problem where the rewards at each stage have been decreased by the gain. For further discussion of this and several other interpretations of the bias function see [13, 14].

It is well-known that if g, h satisfy the *average optimality equations*

$$g + h(s) = \max_{a \in A_s} \left[r(s,a) + \int_S h(s') Q(ds' \mid s, a) \right], \tag{7}$$

then $g = g^*$ and $h = h^* + k1$, where k is a constant, and 1 is the vector of all ones. We say a decision rule d is *canonical* if for all states s, $d(s)$ attains the maximum in (7). Let $\mathbb{F}_{ca}$ denote the set of all canonical decision rules, and let $\mathbb{F}_{bias}$ denote the set of all deterministic decision rules such that d^∞ is bias optimal. The following theorem appears as Theorem 10.3.6 in Hernández-Lerma and Lasserre [7].

Theorem 2 *Suppose the HL-Assumptions hold. Then*

1. *There exists a canonical policy.*

2. *If d^∞ is a stationary, deterministic average optimal policy that induces a Markov chain with stationary distribution ϕ_d, then (g^*, w_d) satisfy* (7) *ϕ_d-almost everywhere, where w_d is the solution to Poisson's equation*

$$g^* + w_d(s) = r(s, d(s)) + \int_S w_d(s')Q(ds' \mid s, d(s)). \tag{8}$$

The HL-Assumptions and Theorem 2 imply that $G \neq \emptyset$. We next state the parts of Theorem 10.3.10 of Hernández-Lerma and Lasserre [7] that are necessary for our analysis.

Theorem 3 *Suppose the HL-Assumptions hold. Then*

1. *There exists a bias optimal decision rule $d \in \mathbb{F}_{bias}$.*

2. *If d^∞ is bias optimal, (g^*, h_d) satisfies the average optimality equations* (7)*.*

3. *The following statements are equivalent:*

   (a) d^∞ is bias optimal

   (b) d^∞ is a canonical decision rule and

$$\int_S \hat{h}(y)Q^*(y|x, d(x)) = 0 \quad \textit{for all } x \in S,$$

   where Q^* is the stationary distribution of the Markov chain generated by d^∞ and $\hat{h}$ (along with g^*) is a solution to the average optimality equations.

Hence the HL-Assumptions imply the existence of a stationary, deterministic policy that is bias optimal and from Part 3 of Theorem 3, $\mathbb{F}_{bias} \subset \mathbb{F}_{ca}$.

5 Threshold Policies

The main purpose of this section is to show that we can restrict attention to policies of threshold form. We will show that all canonical decision rules $\mathbb{F}_{ca}$ are of threshold form. The average optimality equations (7) can be rewritten for our model as

$$g^* + h^*(i,j,z) = \begin{cases} \max\{r_j + U(i+1,z), U(i,z)\} & \text{if } j>0, i<m(z), \\ U(i,z) & \text{if } j>0, i \geq m(z), \\ U(i+j,z) & \text{otherwise} \end{cases} \tag{9}$$

where for $z < T$

$$\begin{aligned} U(i,z) = \int_0^{T-z} \Bigg[& \sum_{j=1}^{\ell} \frac{\lambda_j(z+v)}{\Psi} h^*(i,j,z+v) + \frac{\mu_i(z+v)}{\Psi} h^*(i,-1,z+v) \\ & + [1 - (\sum_{j=1}^{\ell} \frac{\lambda_j(z+v)}{\Psi} + \frac{\mu_i(z+v)}{\Psi})] h^*(i,0,z+v) \Bigg] \Psi e^{-\Psi v}\, dv + h^*(i,0,T) e^{-\Psi(T-z)}, \end{aligned} \tag{10}$$

and

$$U(i,T) = \sum_{j=1}^{\ell} \frac{\lambda_j^\infty}{\Psi} h^*(i,j,T) + \frac{\mu_i^\infty}{\Psi} h^*(i,-1,T) + [1 - \frac{(\sum_{j=1}^{\ell} \lambda_j^\infty + \mu_i^\infty)}{\Psi}] h^*(i,0,T). \tag{11}$$

To show that canonical policies are threshold policies, we show that $U(i,z)$ is strictly concave in i. Note that (9) implies that a canonical policy accepts a type j customer arriving at time t if $r_j + U(i+1,z) > U(i,z)$ where $z = t \wedge T$. Similarly, a canonical policy rejects if the inequality is reversed, and can either accept or reject when equality holds. If $U(i,z)$ is strictly concave, i.e., $U(i+1,z) - U(i,z)$ is strictly decreasing, any canonical policy must be of threshold form.

To show concavity, we construct two sequences of concave functions h_n and U_n, show that the sequences converge to h^* and U, and observe that h^* and U inherit concavity. A short additional argument gives strict concavity. Recall

$$S_T = \{s = (i,j,z) \in S | z = T\}. \tag{12}$$

Note that the optimal bias function $h^*(s)$ is known for $s \in S_T$ since it is the solution to the stationary admission control problem analyzed in [11]. Furthermore, it was shown (see [11, Lemma 4]) that for every j, $h^*(i,j,T)$ is strictly decreasing and strictly concave in i.

Using h^* on S_T, we recursively define two sequences of functions h_n and U_n as follows:

$$h_0(i,j,z) \equiv U_0(i,z) \equiv h^*(i,0,T)$$

and for $n \geq 1$ with $s = (i,j,z)$,

$$h_n(s) \equiv \max_{a \in A_s} \left[r(s,a) - g^* + \int_S h_{n-1}(s') Q(ds' \mid s,a) \right], \tag{13}$$

or equivalently,

$$h_n(i,j,z) = \begin{cases} \max\{r_j + U_{n-1}(i+1,z), U_{n-1}(i,z)\} - g^* & \text{if } j > 0, i < m(z), \\ U_{n-1}(i,z) - g^* & \text{if } j > 0, i \geq m(z), \\ U_{n-1}(i+j,z) - g^* & \text{otherwise}, \end{cases} \tag{14}$$

and

$$U_n(i,z) \equiv \int_0^{T-z} \left[\sum_{j=1}^{\ell} \frac{\lambda_j(z+v)}{\Psi} h_n(i,j,z+v) + \frac{\mu_i(z+v)}{\Psi} h_n(i,-1,z+v) \right.$$
$$\left. + [1 - (\sum_{j=1}^{\ell} \frac{\lambda_j(z+v)}{\Psi} + \frac{\mu_i(z+v)}{\Psi})] h_n(i,0,z+v) \right] \Psi e^{-\Psi v}\, dv + h^*(i,0,T) e^{-\Psi(T-z)}. \tag{15}$$

Lemma 1 *For the non-stationary admission control problem,* $\lim_{n\to\infty} U_n = U$ *and* $\lim_{n\to\infty} h_n = h^*$.

Proof. First note that for $s \in S_T$, we have $h_n(s) = h^*(s)$ and $U_n(s) = U(s)$. To complete the proof, it suffices to show that $h_n(s)$ converges for $s \notin S_T$. This is sufficient since from (13), $\lim_{n\to\infty} h_n(s)$ satisfies the average optimality equation (7), which from Theorem 10.3.7 of Hernández-Lerma and Lasserre [7] has a unique solution up to addition by a constant. Thus, $\lim_{n\to\infty} h_n(s) = h^*(s) + c$ for some constant c and all $s \in S$. However, in our case c must be zero since $\lim_{n\to\infty} h_n(s) = h^*(s)$ on S_T. This also implies that U_n converges to U since $U_n(i,z) = h_n(i,0,z) + g^*$.

For $s = (i,j,z) \notin S_T$, define $J_n(s) \equiv h_n(s) + g^* \, \mathbb{E}[(N_{T-z} + 1) \wedge n]$ where N_{T-z} is a Poisson random variable with mean $(T-z)\Psi$. To show that $h_n(s)$ converges, we will show that $J_n(s)$ converges to a finite value. Since their difference converges to $g^*[(T-z)\Psi + 1]$, $h_n(s)$ must converge.

For $s = (i, j, z) \notin S_T$, rewrite (13) as

$$h_n(s) \equiv \max_{a \in A_s} \left[r(s,a) - g^* + \int_{S_T^c} h_{n-1}(s') Q(ds' \mid s,a) + \int_{S_T} h^*(s') Q(ds' \mid s,a) \right]. \tag{16}$$

Substitute $J_n(s) - g^* \, \mathbb{E}[(N_{T-z} + 1) \wedge n]$ for $h_n(s)$, similarly for $h_{n-1}(s')$, where $s' = (i'j', z+v)$, and use the identity

$$g^* + g^* \int_{v=0}^{T-z} \mathbb{E}[(N_{T-z-v} + 1) \wedge (n-1)] \Psi e^{-\Psi v} \, dv = g^* \, \mathbb{E}[(N_{T-z} + 1) \wedge n]$$

to obtain

$$J_n(s) = \max_{a \in A_s} \left[r(s,a) + \int_{S_T^c} J_{n-1}(s') Q(ds' \mid s,a) + \int_{S_T} h^*(s') Q(ds' \mid s,a) \right]. \tag{17}$$

Since $J_n(s)$ can be interpreted as the total expected reward allowing $(N_{T-z} + 1) \wedge n$ decisions and a terminal reward h^* on S_T, it is straightforward to show inductively in (17) that $J_n(s)$ is non-decreasing in n. Furthermore, since r_1 is the maximum reward obtained at any decision epoch, $J_n(s)$ is bounded above by $(\Psi T + 1) r_1 + h^*(0, 1, T)$. Hence, $J_n(s)$ converges, and the proof is complete. ∎

Lemma 2 *For the non-stationary admission control problem, $U(\cdot, z)$ is strictly decreasing and strictly concave.*

Proof. When we write $h_n \in DC_m$, we mean that for all $z \geq 0$, $j = -1, 0, \ldots, \ell$, we have

$$\Delta h_n(i,j,z) \leq 0 \qquad\qquad i = 0, \ldots, m(z) - 1,$$

$$\Delta^2 h_n(i,j,z) \leq 0 \qquad\qquad i = 0, \ldots, m(z) - 2.$$

In other words, h_n is a non-increasing, concave function in i. Similarly define, $h^* \in DC_m$, $U_n \in DC_m$, and $U \in DC_m$. We will show that for all n, h_n and U_n are in DC_m. Clearly U_0 and h_0 are in DC_m since $h^*(i,j,T)$ is strictly decreasing and strictly concave by [11, Lemma 4]. We first prove that if $U_{n-1} \in DC_m$, then $h_n \in DC_m$. Second, we show if $h_n \in DC_m$, then $U_n \in DC_m$. Hence, by induction, $h_n \in DC_m$ and $U_n \in DC_m$ for $n \geq 0$.

Assume $U_{n-1} \in DC_m$. Then from (14), Δh_n and $\Delta^2 h_n$ are not positive in all cases on the r.h.s. of (14) except possibly when $j > 0$ and $i < m(z)$. To consider this situation, fix $j > 0$, $z \in [0, T]$, $n \geq 1$, and define $i^* \equiv \min\{i \geq 0 | r_j + \Delta U_{n-1}(i, z) \leq 0\}$. By the concavity of U_{n-1}, $r_j + U_{n-1}(i+1, z) < U_{n-1}(i, z)$ if and only if $i < i^*$. Hence,

$$h_n(i, j, z) = \begin{cases} r_j + U_{n-1}(i+1, z) - g^* & \text{if } i < i^* \wedge m(z), \\ U_{n-1}(i, z) - g^* & \text{if } i \geq i^* \wedge m(z). \end{cases} \tag{18}$$

From (18) Δh_n is negative from the inductive hypothesis except possibly

if $i = [i^* \wedge m(z)] - 1$. However,

$$\Delta h_n([i^* \wedge m(z)] - 1, j, z) = -r_j < 0$$

Similarly, $\Delta^2 h_n$ is negative from the inductive hypothesis except possibly if $i = [i^* \wedge m(z)] - 2$ or if $i = [i^* \wedge m(z)] - 1$. Note that,

$$\Delta^2 h_n([i^* \wedge m(z)] - 2, j, z) = -[r_j + \Delta U_{n-1}([i^* \wedge m(z)] - 1, z)] \, < \, 0$$

since $[i^* \wedge m(z)] - 1 < i^*$. Now,

$$\Delta^2 h_n([i^* \wedge m(z)] - 1, j, z) = r_j + \Delta U_{n-1}([i^* \wedge m(z)], z). \tag{19}$$

If $m(z) > i^*$, then the r.h.s. of (19) is $r_j + \Delta U_{n-1}(i^*, z) \leq 0$. If $m(z) \leq i^*$, nothing needs to be shown since there are no requirements on $\Delta^2 h_n(m(z) - 1, j, z)$ to be in DC_m. Thus, we have shown that if $U_{n-1} \in DC_m$, then $h_n \in DC_m$.

Now, we show if U_{n-1} and h_n are in DC_m, then $U_n \in DC_m$. By taking differences on both sides of

(15), it suffices to show that $H_n \in DC_m$ where

$$H_n(i, z+v) \equiv \frac{\mu_i(z+v)}{\Psi} h_n(i, -1, z+v) + [1 - (\sum_{j=1}^{\ell} \frac{\lambda_j(z+v)}{\Psi} + \frac{\mu_i(z+v)}{\Psi})] h_n(i, 0, z+v)$$

$$= \frac{\mu_i(z+v)}{\Psi} U_{n-1}(i-1, z+v) + [1 - (\sum_{j=1}^{\ell} \frac{\lambda_j(z+v)}{\Psi} + \frac{\mu_i(z+v)}{\Psi})] U_{n-1}(i, z+v)$$

In fact, if $H_n \in DC_m$, it follows from (15) that U is strictly decreasing and strictly concave since we will have proven that $\Delta U_n(i, z) \leq \Delta h^*(i, 0, T) e^{-\Psi(T-z)} < 0$ and $\Delta^2 U_n(i, z) \leq \Delta^2 h^*(i, 0, T) e^{-\Psi(T-z)} < 0$.

Using $\Delta(a_i \cdot b_i) = a_{i+1} \Delta b_i + (\Delta a_i) b_i$, yields

$$\begin{aligned}
\Delta H_n(i, z) &= \frac{\mu_{i+1}(z)}{\Psi} \Delta U_{n-1}(i-1, z) + (\frac{\Delta \mu_i(z)}{\Psi}) U_{n-1}(i-1, z) \\
&\quad + [1 - (\sum_{j=1}^{\ell} \frac{\lambda_j(z)}{\Psi} + \frac{\mu_{i+1}(z)}{\Psi})] \Delta U_{n-1}(i, z) - (\frac{\Delta \mu_i(z)}{\Psi}) U_{n-1}(i, z) \\
&= \frac{\mu_{i+1}(z)}{\Psi} \Delta U_{n-1}(i-1, z) - (\frac{\Delta \mu_i(z)}{\Psi})(\Delta U_{n-1}(i-1, z)) \\
&\quad + [1 - (\sum_{j=1}^{\ell} \frac{\lambda_j(z)}{\Psi} + \frac{\mu_{i+1}(z)}{\Psi})] \Delta U_{n-1}(i, z) \\
&= \frac{\mu_i(z)}{\Psi} \Delta U_{n-1}(i-1, z) + [1 - (\sum_{j=1}^{\ell} \frac{\lambda_j(z)}{\Psi} + \frac{\mu_{i+1}(z)}{\Psi})] \Delta U_{n-1}(i, z) \\
&\leq 0
\end{aligned}$$

Similarly,

$$
\begin{aligned}
\Delta^2 H_n(i,z) &= \frac{\mu_{i+1}(z)}{\Psi}\Delta^2 U_{n-1}(i-1,z) + (\frac{\Delta\mu_i(z)}{\Psi})(\Delta U_{n-1}(i-1,z)) \\
&\quad +[1-(\sum_{j=1}^{\ell}\frac{\lambda_j(z)}{\Psi}+\frac{\mu_{i+2}(z)}{\Psi})]\Delta^2 U_{n-1}(i,z) - (\frac{\Delta\mu_{i+1}(z)}{\Psi})(\Delta U_{n-1}(i,z)) \\
&= \frac{\mu_{i+1}(z)}{\Psi}\Delta^2 U_{n-1}(i-1,z) + [1-(\sum_{j=1}^{\ell}\frac{\lambda_j(z)}{\Psi}+\frac{\mu_{i+2}(z)}{\Psi})]\Delta^2 U_{n-1}(i,z) \\
&\quad -(\frac{\Delta\mu_i(z)}{\Psi})(\Delta^2 U_{n-1}(i-1,z)) - (\frac{\Delta^2\mu_i(z)}{\Psi})(\Delta U_{n-1}(i,z)) \\
&= \frac{\mu_i(z)}{\Psi}\Delta^2 U_{n-1}(i-1,z) + [1-(\sum_{j=1}^{\ell}\frac{\lambda_j(z)}{\Psi}+\frac{\mu_{i+2}(z)}{\Psi})]\Delta^2 U_{n-1}(i,z) \\
&\quad -(\frac{\Delta^2\mu_i(z)}{\Psi})(\Delta U_{n-1}(i,z)) \\
&\leq 0
\end{aligned}
$$

The last inequality follows since $U_{n-1} \in DC_m$ and $\mu_i(z)$ is concave. ■

The following proposition summarizes some of the the results relating bias optimal and canonical decision rules.

Proposition 1 *For the non-stationary admission control problem, we have the following:*

1. *Canonical decision rules are of threshold form with*

$$\min\{i \geq 0 | r_j + \Delta U(i,t) \leq 0\} \leq k_j(t) \leq \min\{i \geq 0 | r_j + \Delta U(i,t) < 0\}$$

   for all $t \leq T$.

2. *Canonical decision rules accept Class 1 customers whenever possible; i.e., $k_1(t) \equiv m(t)$ for all $t \geq 0$.*

3. *A canonical decision rule is bias optimal iff $k_j(T) = \min\{i \geq 0 | r_j + \Delta U(i,T) < 0\}$.*

4. *For each fixed t, the vector of control levels $k(t)$ of canonical decision rules (and bias optimal policies) are such that $0 \leq k_\ell(t) \leq \ldots \leq k_1(t) = m(t)$.*

Proof. The first claim follows directly from (9) and Lemma 2. To see that Class 1 customers are accepted whenever possible, note that the difference $U(i+1, z) - U(i, z)$ can be interpreted as the total expected difference in rewards between starting at time z with $i+1$ customers and starting with i customers. Suppose we start two processes at time z on the same probability space. Process 1 begins with $i+1$ customers and Process 2 with i customers. Assume that each process uses the same optimal threshold policy. Since the process on S_T is irreducible, aperiodic and has one finite recurrent class, the two processes will almost surely couple in finite time. Furthermore, since they are using the same threshold policy, any arriving customer that is rejected by Process 2, will also be rejected by Process 1. There are two ways for the processes to couple. If they couple at a departure from Process 1 that is not met by a corresponding departure by Process 2, they have seen the same reward stream and $\Delta U(i, z) = 0$. Let this be event D, and note that, $\Pr(D) > 0$. Similarly, if they couple at an arrival of a Class j customer that is accepted by Process 2, but not by Process 1, $\Delta U(i, z) = -r_j$. Let this event be A_j. Hence,

$$\Delta U(i, z) \quad = \quad -\sum_{j=1}^{\ell} r_j \Pr(A_j) + 0 \Pr(D). \tag{20}$$

Since $\sum_{j=1}^{\ell} \Pr(A_j) < 1$ and $r_1 \geq r_j$ for each j, we have $r_1 + \Delta U(i, z) > 0$. But this is precisely the condition of the average optimality equations that implies it is optimal to accept Class 1 customers if possible.

The third assertion follows from Part 3 of Theorem 3, which states that if a canonical decision rule is bias optimal on the recurrent states, it is bias optimal. The specified decision rule, $k_j(T)$, was shown in [11] to be the bias optimal policy on the recurrent states.

The final claim is clear since the acceptance (rejection) criterion when in state (i, j, z) for $j > 0$ is $r_j + \Delta U(i, z) \geq \ (\leq)\ 0$ and the assumption that $r_1 > r_2 > \cdots > r_\ell$. ∎

6 Monotonicity

In this section we derive sufficient conditions for the bias optimal threshold levels to be monotonic in time.

Theorem 4 *Let $k(t)$ be the vector of bias optimal control levels at time t. Under the assumptions that $\mu(t)$ and $\Delta\mu(t)$ are non-decreasing (non-increasing), $\lambda(t)$ is non-increasing (non-decreasing) in t, and $m(t)$ is non-decreasing (constant) then $k(t)$ is non-decreasing (non-increasing) in t.*

Proof. We prove the first assertion. The second assertion is proved similarly. It suffices to show that $\Delta U(i,z)$ is non-decreasing in z for $0 < z < T$ and $i = 0, 1, \ldots, m(z) - 1$. To prove this, we show that $d\Delta U_n(i,z)/dz \geq 0$ for all n.

In (15), change the region of integration so that v ranges from z to T, take the first difference with respect to i, and differentiate with respect to z using Leibnitz's rule to obtain

$$\begin{aligned}
\frac{d\Delta U_n(i,z)}{dz} \quad = \quad & \int_z^T (\sum_{j=1}^{\ell} \frac{\lambda_j(v)}{\Psi} \Delta h_n(i,j,v) + \frac{\mu_i(v)}{\Psi} \Delta U_{n-1}(i-1,v) \\
& + (1 - (\sum_{j=1}^{\ell} \frac{\lambda_j(v)}{\Psi} + \frac{\mu_{i+1}(v)}{\Psi}))\Delta U_{n-1}(i,v)) e^{-(v-z)}\, dv \\
& + \Delta h^*(i,0,T) e^{-(T-z)} - \left[\sum_{j=1}^{\ell} \frac{\lambda_j(z)}{\Psi} \Delta h_n(i,j,z) + \frac{\mu_i(z)}{\Psi} \Delta U_{n-1}(i-1,z) \right. \\
& \left. + (1 - (\sum_{j=1}^{\ell} \frac{\lambda_j(z)}{\Psi} + \frac{\mu_{i+1}(z)}{\Psi}))\Delta U_{n-1}(i,z) \right]. \qquad (21)
\end{aligned}$$

We prove the derivatives are positive by induction. Since $\Delta U_0(i,z) = \Delta h^*(i,0,T)$ it is clear that the result holds for $n = 0$. Now, assume that it holds for $n - 1 \geq 0$. By adding and subtracting terms involving time

z, and recalling that h^* satisfies the average optimality equations

$$\begin{aligned}\Delta h^*(i,0,T) &= \sum_{j=1}^{\ell}\lambda_j^{\infty}\Delta h^*(i,j,T) + \mu_i^{\infty}\Delta U(i-1,T)\\ &\quad +(1-(\sum_{j=1}^{\ell}\lambda_j^{\infty}+\mu_{i+1}^{\infty}))\Delta U(i,T),\end{aligned}$$

we have

$$\begin{aligned}\frac{d\Delta U_n(i,z)}{dz} &= \int_z^T (\sum_{j=1}^{\ell}\frac{\lambda_j(v)}{\Psi}[\Delta h_n(i,j,v)-\Delta h_n(i,j,z)] + \frac{\mu_i(v)}{\Psi}[\Delta U_{n-1}(i-1,v)-\Delta U_{n-1}(i-1,z)]\\ &\quad +(1-(\sum_{j=1}^{\ell}\frac{\lambda_j(v)}{\Psi}+\frac{\mu_{i+1}(v)}{\Psi}))[\Delta U_{n-1}(i,v)-\Delta U_{n-1}(i,z)])e^{-(v-z)}\,dv\\ &\quad +\int_z^T \kappa_n(i,v)\,dv + \kappa_n(i,T)e^{-(T-z)} \qquad (22)\end{aligned}$$

where

$$\begin{aligned}\kappa_n(i,v) &\equiv \sum_{j=1}^{\ell}[\frac{\lambda_j(v)-\lambda_j(z)}{\Psi}][\Delta h_n(i,j,z))] + [\frac{\mu_i(v)-\mu_i(z)}{\Psi}][\Delta U_{n-1}(i-1,z)]\\ &\quad -(\sum_{j=1}^{\ell}[\frac{\lambda_j(v)-\lambda_j(z)}{\Psi}] + [\frac{\mu_{i+1}(v)-\mu_{i+1}(z)}{\Psi}])[\Delta U_{n-1}(i,z)]. \qquad (23)\end{aligned}$$

The integrand in the first integral of (22) is non-negative since $\Delta h_n(i,j,v)$, $\Delta U_{n-1}(i,v)$ and $\Delta U_{n-1}(i-1,v)$ are non-decreasing in v—the latter two by the induction hypothesis, while the first from the following argument.

To show that $\Delta h_n(i,j,y)$ is non-negative for fixed i and time $y \le T$ let $A_i(y)$ represent the set of customer classes that are accepted by a canonical policy. Similarly, let $R_i(y)$ be the set of rejected customer classes by a canonical policy. Hence, we denote the set of customer classes that are accepted by canonical policies both when there are i and $i+1$ in the system by $A_iA_{i+1}(y)$. Similarly since canonical policies are threshold policies define $A_iR_{i+1}(y)$ and $R_iR_{i+1}(y)$. Now notice there are three possibilities for $\Delta h_n(i,j,y)$,

$$\Delta h_n(i,j,y) = \begin{cases} \Delta U_{n-1}(i+1,y) & \text{if } j \in A_iA_{i+1}(y), \\ -r_j & \text{if } j \in A_iR_{i+1}(y), \\ \Delta U_{n-1}(i,y) & \text{if } j \in R_iR_{i+1}(y). \end{cases} \tag{24}$$

When the optimal actions in state (i,j,v) and (i,j,z) coincide, it is easy to see from (24) that the induction hypothesis guarantees that the first term is non-negative. Furthermore, since $\Delta U_{n-1}(i,z) \le \Delta U_{n-1}(i,v)$ accepting in (i,j,z) implies accepting in (i,j,v). Hence, there are three possibilities remaining to consider; when $j \in A_iR_{i+1}(z) \cap A_iA_{i+1}(v)$, when $j \in R_iR_{i+1}(z) \cap A_iA_{i+1}(v)$, and when $j \in R_iR_{i+1}(z) \cap A_iR_{i+1}(v)$. In the first case $\Delta h_n(i,j,v) - \Delta h_n(i,j,z) = r_j + \Delta U_{n-1}(i+1,v)$ which is non-negative since "accept" is optimal in (i,j,v). The second case yields $\Delta h_n(i,j,v) - \Delta h_n(i,j,z) = r_j + \Delta U_{n-1}(i+1,v) - (r_j + \Delta U_{n-1}(i,z))$ which again is non-negative since accept is optimal in (i,j,v) and reject is optimal in (i,j,z). Finally, for the last case, $\Delta h_n(i,j,v) - \Delta h_n(i,j,z) = -(r_j + \Delta U_{n-1}(i,z))$ which is non-negative since reject is optimal in (i,j,z).

To complete the argument that (22) is non-negative, we need only show that $\kappa_n(i,v) \ge 0$. Using the above notation,

$$\begin{aligned}
\kappa_n(i,v) &= -\sum_{j \in A_iR_{i+1}(z)} \frac{(\lambda_j(v) - \lambda_j(z))}{\Psi} r_j + \sum_{j \in A_iA_{i+1}(z)} \frac{(\lambda_j(v) - \lambda_j(z))}{\Psi} \Delta U_{n-1}(i+1,z) \\
&\quad + \sum_{j \in R_iR_{i+1}(z)} \frac{(\lambda_j(v) - \lambda_j(z))}{\Psi} \Delta U_{n-1}(i,z) + \frac{(\mu_i(v) - \mu_i(z))}{\Psi} \Delta U_{n-1}(i-1,z) \\
&\quad -(\sum_{j=1}^{\ell} \frac{(\lambda_j(v) - \lambda_j(z))}{\Psi} + \frac{(\mu_{i+1}(v) - \mu_{i+1}(z))}{\Psi})[\Delta U_{n-1}(i,z)] \\
&= -\sum_{j \in A_iR_{i+1}(z)} \frac{(\lambda_j(v) - \lambda_j(z))}{\Psi} (r_j + \Delta U_{n-1}(i,z)) + \sum_{j \in A_iA_{i+1}(z)} \frac{(\lambda_j(v) - \lambda_j(z))}{\Psi} \Delta^2 U_{n-1}(i,z)) \\
&\quad - \frac{(\mu_{i+1}(v) - \mu_{i+1}(z))}{\Psi} \Delta^2 U_{n-1}(i-1,z) - \frac{(\Delta\mu_i(v) - \Delta\mu_i(z))}{\Psi} \Delta U_{n-1}(i-1,z) \\
&\ge 0
\end{aligned}$$

where the inequality follows from the assumptions on $\mu(t), \Delta\mu(t), \lambda(t)$, the fact that U_{n-1} is concave and non-increasing, and the definition of the set $A_i R_{i+1}(z)$.

Hence, $\Delta U_n(i,z)$ is non-decreasing. Letting $n \to \infty$ we have $\Delta U(i,z)$ is non-decreasing. Since it is optimal to accept customers when in state (i,j,z) if $r_j + \Delta U(i,z) \geq 0$, the optimal control level can only increase. ∎

7 Examples and Counterexamples

We have shown that not only are all bias optimal policies threshold policies, but also, given the right structure of the problem, these policies can be chosen to be monotonic in time. We illustrate these concepts with some simple examples.

Example 1

Suppose that $\ell = 2$ with $r_1 = 1$ and $r_2 = .75$. Further let $\mu_i(t) = .035i$ and $m(t) = 5$ for all t. Then if

$$\lambda_1(t) = \begin{cases} .65 & \text{if } t \leq 2\pi, \\ .45 & \text{if } t > 2\pi, \end{cases} \tag{25}$$

and

$$\lambda_2(t) = \begin{cases} .25 & \text{if } t \leq 2\pi, \\ .20 & \text{if } t > 2\pi \end{cases} \tag{26}$$

we have the conditions of part 2 of Theorem 1 which imply the existence of bias optimal threshold decision rules that are monotonically non-decreasing in time. Recall that this was evidenced by the monotonicity of ΔU and the concavity of U. Figure 1 is a plot of $\Delta U(i,z)$ versus time. Note that $\Delta U(i,z)$ is non-decreasing in z. Further, since $\Delta U(i+1,z) < \Delta U(i,z)$ for each $i = 0, 1, \ldots, 4$, concavity holds.

Example 2

Suppose the parameters are the same as Example 1 except

$$\lambda_1(t) = \begin{cases} .5 + .5\sin(t) & \text{if } t \leq 2\pi, \\ .65 & \text{if } t > 2\pi \end{cases} \tag{27}$$

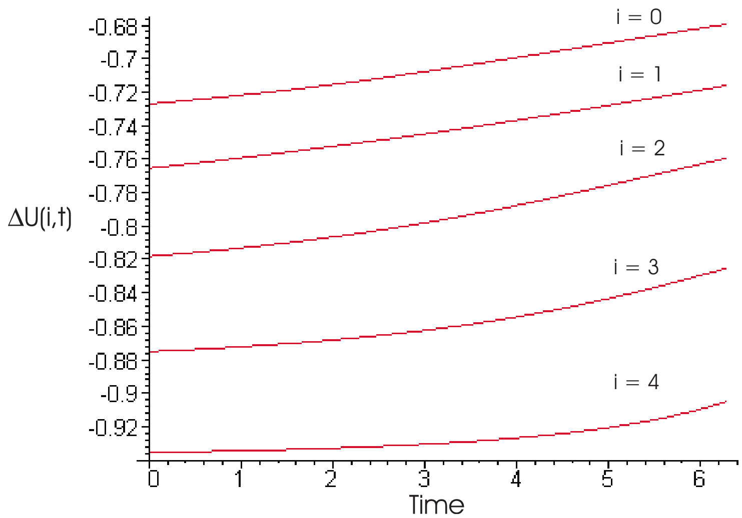

Figure 1: An example of the concavity of U and the monotonicity of $\Delta U(i,t)$.

and

$$\lambda_2(t) \quad = \quad \begin{cases} 1+\sin(t) & \text{if } t \leq 2\pi, \\ .25 & \text{if } t > 2\pi. \end{cases} \tag{28}$$

Figure 2 shows that the concavity continues to hold implying the existence of bias optimal threshold policies, however, the monotonicity fails. This is to be expected since the arrival process is not monotonic. It is also interesting to note that $r_2 + \Delta U(1,t) < 0$ until $t \approx 1.8$, then $r_2 + \Delta U(1,t) > 0$ until $t \approx 4.5$ when it is again non-negative. This implies that the optimal control level for Class 2 customers changes from zero to 1 and back to zero at these times, respectively.

One might conjecture that the assumption that $m(z)$ is constant in part 3 of Theorem 1 is overly restrictive and all that is really needed is that $m(z)$ be non-increasing. However, the following example shows with $m(z)$ non-increasing we may lose both concavity and monotonicity. In fact, the bias optimal policies may not even be of threshold form!

To handle the fact that there may be more than $m(z)$ customers in the system we assume that customers that were already accepted are allowed to stay and complete service. However, no more arrivals can be accepted until the number of customers in the system falls below the current capacity.

Example 3

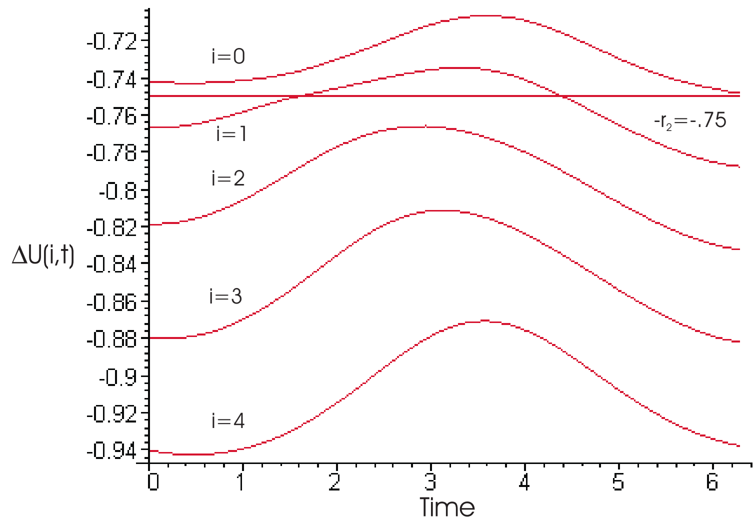

Figure 2: An example of the concavity of U where the monotonicity does not hold.

Suppose the parameters are the same as Example 1 except

$$m(t) = \begin{cases} 5 & \text{if } t \leq 2\pi, \\ 3 & \text{if } t > 2\pi, \end{cases} \tag{29}$$

$$\lambda_1(t) = \begin{cases} .10 & \text{if } t \leq 2\pi, \\ .15 & \text{if } t > 2\pi, \end{cases} \tag{30}$$

and

$$\lambda_2(t) = \begin{cases} .15 & \text{if } t \leq 2\pi, \\ .35 & \text{if } t > 2\pi. \end{cases} \tag{31}$$

Figure 3 shows that $\Delta U(3, z)$ is not monotonic. Since the two curves cross, $U(i, z)$ is not concave either. Furthermore, since $r_2 + \Delta U(3, z) > 0$ and $r_2 + \Delta U(2, z) < 0$ for z larger than approximately 5.4, we have the surprising conclusion that we reject when there are 2 customers in the system, but accept when there are 3. The bias optimal policy is **not** even of threshold form!

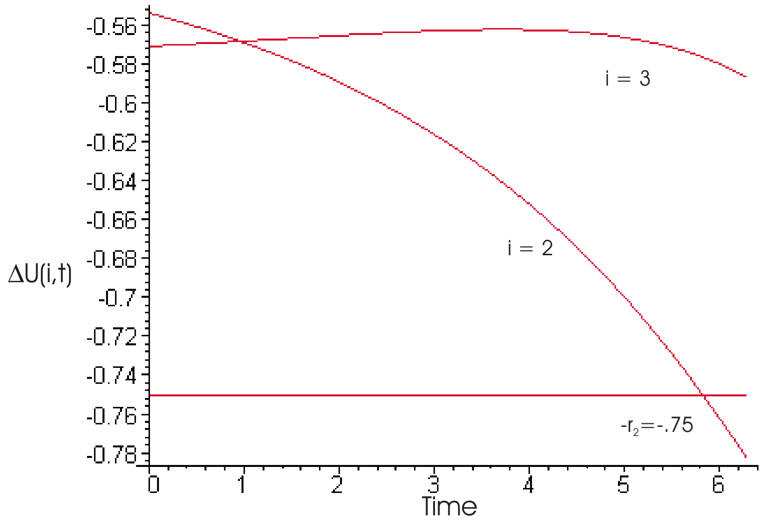

Figure 3: An example when the concavity of U and the monotonicity of $\Delta U(i, z)$ fail.

8 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 the referee for several insightful and helpful comments. This work is partially supported by NSF grant: DMI-9908321.

References

[1] Eitan Altman and Shaler Stidham, Jr. Optimality of monotone policies for two-action markovian decision processes, with applications to control of queues with delayed information. *Queueing Systems*, 3-4(21):267–291, 1995.

[2] Prajit K. Dutta. What do discounted optima converge to? A theory of discount rate asymptotics in economic models. *Journal of Economic Theory*, 55(1):64–94, 1991.

[3] Michael C. Fu, Steven I. Marcus, and I-Jeng Wang. Monotone optimal policies for a transient queueing staffing problem. *Operations Research*, 48(2):327–331, March-April 2000.

[4] Hussein A. Ghoneim and Shaler Stidham, Jr. Control of arrivals to two queues in series. *European Journal of Operational Research*, 21:399–409, 1985.

[5] Moshe Haviv and Martin L. Puterman. Bias optimality in controlled queueing systems. *Journal of Applied Probability*, 35:136–150, 1998.

[6] Werner E. Helm and Karl-Heinz Waldmann. Optimal control of arrivals to multi-server queues in a random environment. *Journal of Applied Probability*, 21:602–615, 1984.

[7] Onésimo Hernández-Lerma and Jean B. Lasserre. *Further Topics on Discrete-Time Markov Control Processes*. Springer-Verlag Inc., New York, 1999.

[8] Daniel P. Heyman. Optimal operating policies for m/g/1 queueing systems. *Operations Research*, 16:362–382, 1968.

[9] Mark E. Lewis. *Bias Optimality in a Two-Class Nonstationary Queueing System*. Doctoral dissertation, Georgia Institute of Technology, Atlanta, Georgia, September 1998.

[10] Mark E. Lewis. Average optimal policies in a controlled queueing system with dual admission control. *Journal of Applied Probability*, 38(2):369–385, June 2001.

[11] Mark E. Lewis, Hayriye Ayhan, and Robert D. Foley. Bias optimality in a queue with admission control. *Probability in the Engineering and Informational Sciences*, 13:309–327, 1999.

[12] Mark E. Lewis and Martin L. Puterman. A note on bias optimality in controlled queueing systems. *Journal of Applied Probability*, 37(1):300–305, 2000.

[13] Mark E. Lewis and Martin L. Puterman. Bias optimality. In Eugene Feinberg and Adam Shwartz, editors, *Handbook of Markov Decision Processes: Methods and Applications*. Kluwer Academic Publishers, 2001. in press.

[14] Mark E. Lewis and Martin L. Puterman. A probabilistic analysis of bias optimality in unichain Markov decision processes. *IEEE Transactions on Automatic Control*, 46(1):96–100, January 2001.

[15] Steven A. Lippman. Applying a new device in the optimization of exponential queueing systems. *Operations Research*, 23(4):687–712, 1975.

[16] Steven A. Lippman and Sheldon M. Ross. The streetwalker's dilemma: A job shop model. *SIAM Journal of Applied Mathematics*, 20(3):336–342, 1971.

[17] Willam A. Massey and Ward Whitt. Stationary-process approximations for the nonstationary erlang loss model. *Operations Research*, 44(6):976–983, November-December 1996.

[18] Willam A. Massey and Ward Whitt. Peak congestion in multi-service systems with slowly varying arrival rates. *QUESTA*, 25:157–172, 1997.

[19] Bruce L. Miller. A queueing reward system with several customer classes. *Management Science*, 16(3):234–245, 1969.

[20] Martin I. Reiman. Asymptotically optimal trunk reservation for large trunk groups. In *Proceedings of the 28th Conference on Decision and Control*, pages 2536–2541, 1989.

[21] Shaler Stidham, Jr. Optimal control of admission to a queueing system. *IEEE Transactions on Automatic Control*, AC-30(8):705–713, 1985.

[22] Arthur F. Veinott. On finding optimal policies in discrete dynamic programming with no discounting. *Annals of Mathematical Statistics*, 37(5):1284–1294, October 1966.

[23] Arthur F. Veinott, Jr. Markov decision chains. In *Studies in Optimization*, volume 10 of *Studies in Mathematics*, pages 124–159. Mathematics Association of America, 1974.

[24] M. Yadin and P. Naor. Queueing systems with a removable service station. *Operations Research Quarterly*, 14(4):393–405, 1963.