UNIVERSIDAD CARLOS III DE MADRID working papers

Working Paper 08-72
Statistics and Econometrics Series 25
December 2008

Departamento de Estadística
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624-98-49

MARGINAL PRODUCTIVITY INDEX POLICIES FOR PROBLEMS OF ADMISSION CONTROL AND ROUTING TO PARALLEL QUEUES WITH DELAY

Peter Jacko and José Niño-Mora

Abstract

In this paper we consider the problem of admission control of Bernoulli arrivals to a buffer with geometric server, in which the controller's actions take effect one period after the actual change in the queue length. An optimal policy in terms of marginal productivity indices (MPI) is derived for this problem under the following three performance objectives: (i) minimization of the expected total discounted sum of holding costs and rejection costs, (ii) minimization of the expected time-average sum of holding costs and rejection costs, and (iii) maximization of the expected time-average number of job completions. Our employment of existing theoretical and algorithmic results on restless bandit indexation together with some new results yields a fast algorithm that computes the MPI for a queue with a buffer size of I performing only O(I) arithmetic operations. Such MPI values can be used both to immediately obtain the optimal thresholds for the admission control problem, and to design an index policy for the routing problem (with possible admission control) in the multi-queue system. Thus, this paper further addresses the problem of designing and computing a tractable heuristic policy for dynamic job admission control and/or routing in a discrete time Markovian model of parallel loss queues with one-period delayed state observation and/or action implementation, which comes close to optimizing an infinite-horizon problem under the above three objectives. Our approach seems to be tractable also for the analogous problems with larger delays and, more generally, for arbitrary restless bandits with delays.

Keywords: admission control, routing, parallel queues, delayed information, delayed action implementation, index policy, restless bandits, marginal productivity index

Department of Statistics, Universidad Carlos III de Madrid, Av. de Universidad 30, 289 11 Leganés (Madrid), Spain, e-mail: peter.jacko@uc3m.es, jnimora@alum.mit.edu

Marginal Productivity Index Policies for Problems of Admission Control and Routing to Parallel Queues with Delay *

Peter Jacko & José Niño-Mora

Department of Statistics
Universidad Carlos III de Madrid
Av. de Universidad 30
289 11 Leganés (Madrid), Spain
peter.jacko@uc3m.es, jnimora@alum.mit.edu

Abstract. In this paper we consider the problem of admission control of Bernoulli arrivals to a buffer with geometric server, in which the controller's actions take effect one period after the actual change in the queue length. An optimal policy in terms of marginal productivity indices (MPI) is derived for this problem under the following three performance objectives: (i) minimization of the expected total discounted sum of holding costs and rejection costs, (ii) minimization of the expected time-average sum of holding costs and rejection costs, and (iii) maximization of the expected time-average number of job completions. Our employment of existing theoretical and algorithmic results on restless bandit indexation together with some new results yields a fast algorithm that computes the MPI for a queue with a buffer size of I performing only $\mathcal{O}(I)$ arithmetic operations. Such MPI values can be used both to immediately obtain the optimal thresholds for the admission control problem, and to design an index policy for the routing problem (with possible admission control) in the multi-queue system. Thus, this paper further addresses the problem of designing and computing a tractable heuristic policy for dynamic job admission control and/or routing in a discrete time Markovian model of parallel loss queues with one-period delayed state observation and/or action implementation, which comes close to optimizing an infinite-horizon problem under the above three objectives. Our approach seems to be tractable also for the analogous problems with larger delays and, more generally, for arbitrary restless bandits with delays.

Key words: admission control, routing, parallel queues, delayed information, delayed action implementation, index policy, restless bandits, marginal productivity index

December 19, 2008

* This research is part of work carried out towards the completion of graduate student P. Jacko's doctorate under the direction of J. Niño-Mora, who acknowledges funding support by the Spanish Ministry of Education and Science under projects MTM2004-02334, MTM2007-63140, and an I3 faculty endowment grant, by the European Union's Network of Excellence Euro-FGI, and by the Autonomous Community of Madrid under grant CCG07-UC3M/ESP-3389. P. Jacko's work has been supported by the Spanish Ministry of Education and Science under predoctoral research training grant BES-2005-7026, attached to project MTM2004-02334 and funded by the European Social Fund. We are grateful to Sofía Villar for bringing to our attention the Möbius transformation, crucial for obtaining closed-form solutions of the recurrences in this paper.

1 Introduction

This paper addresses the problem of designing and computing a tractable heuristic policy for dynamic job admission control and/or routing in a discrete time Markovian model of parallel loss queues with one-period delayed state observation and/or action implementation, which comes close to optimizing an infinite-horizon problem under several objectives. Two versions of the model are considered, depending on whether the admission control capability is enabled or not. The queue servers may be endowed with finite or infinite buffer space.

We consider the following three performance objectives: (i) minimization of the expected total discounted sum of holding costs and rejection costs, (ii) minimization of the expected time-average sum of holding costs and rejection costs, and (iii) maximization of the expected time-average number of job completions. Holding costs are assumed to be convex and nondecreasing in the number of jobs queued in the buffer space.

Such problems are relevant in a variety of application domains, most notably in the operation of packet-switched communication networks and distributed computer systems. In such systems there are nonnegligible propagation delays, which force the controller to take decisions based on stale system state information and cannot take effect before a time lag. Additional recent applications include long-distance-controlled robots, and situations in which an advanced processing of observations is necessary.

As for our considering joint admission control and routing problems, instead of restricting attention to the conventional pure-routing case, the motivation is that it allows the system designer to take into account the tradeoff between rejection and holding costs. The key insight is that, when the system is heavily congested, denying access to further arrivals until the congestion is sufficiently reduced can substantially decrease holding costs at a relatively small expense in terms of increased rejection costs.

The above problems are naturally formulated as *partially observed Markov decision processes* (POMDPs), which in turn are readily reformulated as conventional *Markov decision processes* (MDPs) by redefining the state of each queue as the *augmented state* build up of the last observed queue length and the actions applied since then (Brooks and Leondes, 1972). Computation of optimal policies for the resultant multidimensional MDPs by solving the associated *dynamic programming* (DP) equations is, however, hindered by the curse of dimensionality in large-scale models. We will thus focus attention on the more realistic and practical goals of designing and computing well-grounded heuristic policies that are readily implementable. Since in such problems the controller must dynamically assess the relative values of alternative rejection and routing actions, it is intuitively appealing to do so based on an *index policy*, defined after the model description below.

1.1 Model Description

Time is slotted into discrete-time *epochs* $t = 0, 1, 2, \ldots$ The system consists of N independent parallel queues with servers and a gate. Queue $n \in \mathcal{N} := \{1, \ldots, N\}$ is endowed with a (possibly infinite) buffer with room for holding $I_n \geq 1$ jobs waiting or in service, and has a single geometric server, which serves jobs in FCFS order and completes the service of a job at the end of a period with probability $0 < \mu_n < 1$.

Jobs arrive to the system as a Bernoulli stream with probability $0 < \lambda \leq 1$ of arrival at the beginning of each period. Upon a job's arrival to the gate, a central controller (gatekeeper) must decide: (i) in the case that admission control is enabled, whether to admit the job or to reject it (admission function); and, if admitted, (ii) to which of N queues in parallel to route the job for service (router function). We assume that a customer that is admitted and routed to an empty queue starts to be served immediately, and therefore may leave the system at the end of the same period if the service is completed.

Denote by $X_n(t)$ the *state* of queue n at the beginning of period t, given by the number of jobs it holds waiting or in service, and by $a_n(t) \in \{0, 1\}$ the *action* indicator that takes the value 1 when a job arriving at time t *is not to be routed to queue* n. We assume that at the start of period t the controller does not know the current state, but has information on previous states and actions, knowing in particular $X_n(t-1)$ and $a_n(t-1)$ for each queue n. Thus, we deal with the problem with a *one-period delay*.

Action choice is based on adoption of an admission and routing policy (if admission function is enabled), or just a routing policy (if it is not), denoted by π. This is to be chosen from the corresponding class Π of (possibly randomized) policies that use previous state and action information.

1.2 Performance Objectives

For two of our performance objectives defined below we will assume that the system incurs per-queue *holding costs* at rate $c_n(i_n)$ per period during which i_n jobs are held in queue n, such that $c_n(i_n)$ is convex and nondecreasing in i_n for each queue n. The system further incurs *loss costs* at rate ν per rejected job, due either to active rejection (not admitting) or to forced rejection (when an admitted job finds the buffer to which it is routed full).

We will find it convenient to formulate the *overal cost* incurred in a period in which the joint system state is $\boldsymbol{i} = (i_n)$ and action $\boldsymbol{a} = (a_n)$ prevails as a constant plus a term that is separably additive across queues, using the identity

$$\sum_{n \in \mathcal{N}} c_n(i_n) + \nu\lambda \left[1 - \sum_{n \in \mathcal{N}} (1 - a_n)\right] = -(N-1)\lambda\nu + \sum_{n \in \mathcal{N}} \left[c_n(i_n) + \nu\lambda a_n\right].$$

Note that the term $1 - \sum_{n \in \mathcal{N}} (1 - a_n)$ in the above equality takes the value 1 if an arrival is to be rejected ($a_n = 1$ for every queue n), and takes the value 0 otherwise ($a_n = 0$ for exactly one queue n).

For the third performance objective we will denote by $c'_n(a_n, i_n)$ the expected rate of *job completions* per period during which i_n jobs are held in queue n and action a_n prevails, i.e.,

$$c'_n(a_n, i_n) := \begin{cases} 0, & \text{if } i_n = 0 \text{ and } a_n = 1, \\ \lambda\mu_n, & \text{if } i_n = 0 \text{ and } a_n = 0, \\ \mu_n, & \text{if } i_n \geq 1. \end{cases}$$

Let $\mathbb{E}^{\pi}_{(\boldsymbol{a},\boldsymbol{i})}[\cdot]$ denote expectation under policy π conditioned on the initial previous joint action and state vectors being equal to $\boldsymbol{a}(-1) := \boldsymbol{a} = (a_n)$ and $\boldsymbol{X}(-1) := \boldsymbol{i} = (i_n)$. The operation of such a system raises the following performance optimization problems:

(i) find a policy minimizing the expected total discounted sum of holding costs and rejection costs,

$$\min_{\pi\in\Pi} \mathbb{E}^{\pi}_{(\boldsymbol{a},\boldsymbol{i})} \left[\sum_{t=0}^{\infty} \sum_{n\in\mathcal{N}} \{c_n(X_n(t)) + \nu\lambda a_n(t)\} \beta^t \right], \tag{1}$$

where $0 < \beta < 1$ is the discount factor;

(ii) find a policy minimizing the expected time-average sum of holding costs and rejection costs,

$$\min_{\pi\in\Pi} \limsup_{T\to\infty} \frac{1}{T} \mathbb{E}^{\pi}_{(\boldsymbol{a},\boldsymbol{i})} \left[\sum_{t=0}^{T} \sum_{n\in\mathcal{N}} \{c_n(X_n(t)) + \nu\lambda a_n(t)\} \right]; \tag{2}$$

(iii) find a policy maximizing the expected time-average number of job completions,

$$\max_{\pi\in\Pi} \liminf_{T\to\infty} \frac{1}{T} \mathbb{E}^{\pi}_{(\boldsymbol{a},\boldsymbol{i})} \left[\sum_{t=0}^{T} \sum_{n\in\mathcal{N}} c'_n(a_n(t), X_n(t)) \right]. \tag{3}$$

1.3 Index Policies

As mentioned earlier, a way to formulate present model as an MDP is to redefine the state of each queue n as the augmented state $\widetilde{X}_n(t) := (a_n(t-1), X_n(t-1))$, and use the joint state and action process $\widetilde{\boldsymbol{X}}(t) := (\widetilde{X}_n(t))$ and $\boldsymbol{a}(t) := (a_n(t))$.

In the present model, index policies are based on attaching to each queue n a numeric index $\nu_n(a_n, i_n)$, which can be thought of as a *measure of undesirability for routing a job to queue* n, given as a function of the queue's augmented state, which we denote by (a_n, i_n) and emphasize that it refers to the observed action-state pair at the previous period. We note that we allow the index to be undefined for certain (uncontrollable) states. Further, under the time-average criteria (2) and (3), in case that the (first-order) index $\nu_n(a_n, i_n)$ be defined and constant, we define a second-order index $\gamma_n(a_n, i_n)$. (When the first-order index is undefined, then the second-order index is undefined as well.)

The resultant index policy prescribes the following actions, when at time t the augmented state of each queue n is known to be $\widetilde{X}_n(t) = (a_n, i_n)$:

- under objective (1),
  - in the problem version with admission control capability, the policy prescribes to admit an arriving job if $\nu > \nu_n(a_n, i_n)$ for at least one queue n such that $\nu_n(a_n, i_n)$ is defined, i.e., if the cost of rejecting the job exceeds the undesirability of routing it to some queue; otherwise, the job is rejected;
  - if admitted, the job is routed to a queue
    - of *lowest index* $\nu_n(a_n, i_n)$, breaking ties arbitrarily, among those queues n for which $\nu_n(a_n, i_n)$ is defined and $\nu > \nu_n(a_n, i_n)$, if at least one such queue exists;
    - of *undefined index* $\nu_n(a_n, i_n)$, breaking ties arbitrarily, if there is no queue n for which $\nu_n(a_n, i_n)$ is defined and $\nu > \nu_n(a_n, i_n)$ and if at least one queue with undefined index exists;
    - of *lowest index* $\nu_n(a_n, i_n)$, breaking ties arbitrarily, in all the remaining cases;
- under objectives (2) and (3),

- in the problem version with admission control capability, the policy prescribes to admit an arriving job if $\nu > \nu_n(a_n, i_n)$ for at least one queue n such that $\nu_n(a_n, i_n)$ is defined, i.e., if the cost of rejecting the job exceeds the undesirability of routing it to some queue; otherwise, the job is rejected;
- if admitted, the job is routed to a queue
  - ∗ of *lexicographically lowest index pair* $(\nu_n(a_n, i_n), \gamma_n(a_n, i_n))$, breaking ties arbitrarily, among those queues n for which $\nu_n(a_n, i_n)$ and $\gamma_n(a_n, i_n)$ are defined and $\nu > \nu_n(a_n, i_n)$, if at least one such queue exists;
  - ∗ of *undefined index* $\nu_n(a_n, i_n)$, breaking ties arbitrarily, if there is no queue n for which $\nu_n(a_n, i_n)$ and $\gamma_n(a_n, i_n)$ are defined and $\nu > \nu_n(a_n, i_n)$, and if at least one queue with undefined index exists;
  - ∗ of *lexicographically lowest index pair* $(\nu_n(a_n, i_n), \gamma_n(a_n, i_n))$, breaking ties arbitrarily, in all the remaining cases.

Note that such policies may well prescribe to admit and route a job to a queue that is actually full, unbeknownst to the controller, in which case the job will be blocked and hence rejected.

For the special case of the pure-routing problem under objectives (1) and (2) in which there are two symmetric infinite-buffer queues and linear holding cost ($N = 2$, $\mu_n \equiv \mu$, and $c_n(i) \equiv i$), it was shown in Kuri and Kumar (1995) that an index policy is optimal: the *Join the Shortest Expected Queue* (JSEQ) rule, where the JSEQ index of a queue n is defined as

$$\nu_n^{\text{JSEQ}}(a_n, i_n) := \begin{cases} i_n - \mu, & \text{if } a_n = 1, i_n \geq 1, \\ 0, & \text{if } a_n = 1, i_n = 0, \\ i_n + \lambda - \mu, & \text{if } a_n = 0, i_n \geq 1, \\ \lambda(1-\mu), & \text{if } a_n = 0, i_n = 0, \end{cases}$$

where the index represents the expected value of $X_n(t)$ conditioned on $(a_n(t-1), X_n(t-1)) = (a_n, i_n)$. Such a result partially extends to queues with delays classical results in (Winston, 1977; Hordijk and Koole, 1990) for symmetric queues without delays on optimality of the *Join the Shortest (Nonfull) Queue* (JSQ) rule.

For the case of routing to two nonsymmetric queues with infinite buffers, in which index policies need no longer be optimal, Artiges (1995) showed (in a variation on the above model) that the optimal routing policy is characterized by a monotone switching curve, extending a classical result in Hajek (1984) for a model without delayed information. Still, one can easily devise a variety of heuristic routing index rules by defining indices based on ad hoc arguments, analogously to the *Shortest Expected Delay* routing rule in Houck (1987). Yet, a drawback of such conventional indices, which typically measure a queue's expected weighted load, is that they only give a routing rule, being of no use to obtain a reasonable combined admission control and routing rule as outlined above, since consideration of rejection costs does not play a role in their definition.

We are thus led to address the issue of how to define appropriate indices $\nu_n(a_n, i_n)$ for the above admission control and routing problems. Instead of proposing some ad hoc index via heuristic arguments, we will deploy a unifying fundamental design principle for priority allocation policies in *multiarmed*

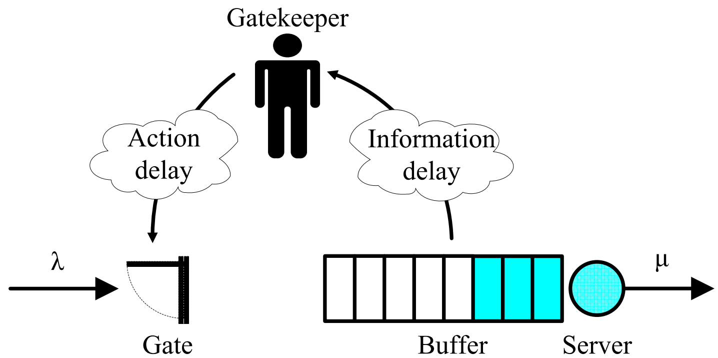


Figure 1. A design of the single-queue admission control problem with delay. The gatekeeper's work consists shutting and opening the entry gate and thus of rejecting of some of the arriving customers.

restless bandit problems (MARBPs), of which (1), (2), and (3) are special cases, based on the economically intuitive concept of *marginal productivity index* (MPI).

Such an approach was introduced in Whittle (1988), and has been developed and applied in a variety of models by the second author in work including Niño-Mora (2001, 2002, 2006b,a), which was reviewed in Niño-Mora (2007b). In particular, Niño-Mora (2002, 2007c) introduced such an approach to the design of index policies for admission control and routing to parallel exponential queues without delayed information. As for use of MPI policies for problems with delayed state information, they were introduced in Niño-Mora (2007d) in the setting of a dynamic scheduling model.

In the present setting, and focusing for concreteness on discounted problem (1) under combined admission control and routing, such a restless bandit indexation approach is based on decoupling the problem into individual single-queue admission control subproblems, one for each queue $n \in \mathcal{N}$:

$$\min_{\pi_n \in \Pi_n} \mathbb{E}^{\pi_n}_{(a_n, i_n)} \left[\sum_{t=0}^{\infty} \{c_n(X_n(t)) + \nu\lambda a_n(t)\} \beta^t \right], \tag{4}$$

where Π_n denotes the class of admission control policies based on one-period delayed state observation for operating queue n *in isolation*, and $\mathbb{E}^{\pi_n}_{(a_n, i_n)}[\cdot]$ denotes expectation conditioned on the initial observed state and action pair being equal to $\widetilde{X}_n(0) := (a_n(-1), X_n(-1)) = (a_n, i_n)$. Note that, in such a setting, taking action $a_n(t) = 1$ at period t means denying access to potential arrivals, which can be conveniently visualized as the action of *shutting the queue's entry gate* which is taken by a *gatekeeper* (see Figure 1).

Problem (4) is a single *restless bandit problem* (RBP), i.e., a binary-action ($a_n(t) = 1$: active; $a_n(t) = 0$: passive) MDP, on which we can deploy the powerful theoretical and algorithmic results available for *restless bandit indexation* (cf. Niño-Mora, 2007b). Let us say that problem (4) is *indexable* if there exists an index $\nu_n^{\mathrm{MPI}}(a_n, i_n)$ that characterizes its optimal policies for every real value of rejection cost parameter ν, as follows: it is optimal to take the active action (shut the entry gate) in augmented state $\widetilde{X}_n(t) = (a_n, i_n)$ if $\nu_n^{\mathrm{MPI}}(a_n, i_n) \geq \nu$ and it is optimal to take the passive action (open the entry gate) in augmented state $\widetilde{X}_n(t) = (a_n, i_n)$ if $\nu_n^{\mathrm{MPI}}(a_n, i_n) \leq \nu$.

In such a case, we term $\nu_n^{\mathrm{MPI}}(a_n, i_n)$ the queue's MPI, due to its economic interpretation as a measure of the rate of marginal reduction in expected holding cost relative to the marginal increase in expected rejections that results from shutting the gate in state (i_n, a_n) instead of opening it, which characterizes the expected holding cost versus rejections tradeoff curve. Such is the index we propose to use as the basis for designing an index rule for admission control and/or routing for the multi-queue problems of concern.

1.4 Contributions

Two issues need thus be addressed: (i) show that problem (4) is indeed indexable; and (ii) design an efficient index-computing algorithm. As for the first issue, we will deploy the sufficient indexability conditions based on *partial conservation laws* (PCLs) introduced in Niño-Mora (2001, 2002). Such conditions require one to identify a family of stationary deterministic policies among which an optimal policy for problem (4) exists for every value of the parameter ν.

For such a purpose, we draw on results in Altman and Nain (1992) and Kuri and Kumar (1995) that characterize the structure of optimal policies for such a single-queue admission control problem (in an infinite-buffer model) with one-period delayed state information. Such work shows that it suffices to consider policies that are characterized by two thresholds $k_1 \geq k_0 \geq 0$, as follows: if the previous observed number of jobs in the system was i and the previous action was to open, i.e., $a = 0$ (resp. shut, i.e., $a = 1$) the queue's entry gate, the (k_0, k_1)-policy prescribes to shut the gate iff $i > k_0$ (resp. iff $i > k_1$).

The intuition behind such a result is that, if it is optimal to shut the entry gate given that it was previously shut, then, other things being equal, it should also be optimal to shut it when it was previously open, as in the latter case the actual number of jobs in the system cannot be smaller than in the former. It is further shown in Altman and Nain (1992) that one need only consider threshold pairs that differ in at most one unit: $0 \leq k_1 - k_0 \leq 1$. Note that, in order to be consistent with such *bi-threshold policies*, the MPI $\nu_n^{\mathrm{MPI}}(a_n, i_n)$ must be monotone nondecreasing in i_n for both $a_n \in \{0, 1\}$, and must satisfy $\nu_n^{\mathrm{MPI}}(0, i_n) \geq \nu_n^{\mathrm{MPI}}(1, i_n)$.

As for the second issue, that of index computation, provided PCL-indexability is established relative to such a family of policies, one can use the adaptive-greedy index algorithm introduced in Niño-Mora (2001, 2002) to compute the MPI. Using the general fast-pivoting implementation given in Niño-Mora (2007a) such an algorithm has a cubic arithmetic operation complexity in the number of restless bandit states, which in the present setting corresponds to an $\mathcal{O}(I^3)$ operation count. While tractable, such a complexity can be overly burdensome for online computation in high-speed communication switches.

Relative to the above two issues, this paper presents the following contributions: (i) it shows that problem (4) is PCL-indexable relative to bi-threshold policies, which ensures both existence of the MPI and the validity of the adaptive-greedy index algorithm for its computation; (ii) by exploiting special structure, a substantially faster index algorithm is presented that computes the MPI in $\mathcal{O}(I)$ operations; (iii) we present an algorithm to calculate the MPI for any particular state by performing at most $\mathcal{O}(\log_2 I)$ arithmetic operations; (iv) validity of the same algorithm for the MPI under the time-average criterion to be used for (2) is established; (v) the MPI to be used for (3) is obtained by the same algorithm and shown to be constant, and the second-order MPI is derived.

An extensive computational study testing the performance of the proposed index policies will be included in the final version of this paper.

2 MDP Formulations

In order to see the analogy, in this section we formulate as a Markov decision process (MDP) both the admission control problem without delay and the admission control problem with a one-period delay. Since the problem considered in the following is a single-queue case, we drop the subscript n from the notation. For concreteness, in this section we focus on the objective (1); for the problem of maximum expected number of job completions one would simply replace the holding cost c_i for the completion reward $-c'(a, i)$.

2.1 Admission Control Problem

First we formulate as an MDP the no-delay admission control problem. Let $X(t)$ be the state process, denoting the *queue length* (including customers in service, if any) at time epoch t. If $a(t)$ denotes the action process, then the task at time epoch t is to choose between closing the gate ($a(t) = 1$) and letting the gate open ($a(t) = 0$). The MDP elements are as follows:

- The *action space* is denoted by $\mathcal{A} := \{0, 1\}$.
- The *state space* is $\mathcal{I} := \{0, 1, \ldots, I\}$, where state $i \in \mathcal{I}$ represents the number of customers in the buffer or in service.
- Denoting by $\zeta := \lambda(1-\mu)$, $\eta := \mu(1-\lambda)$, and $\varepsilon := 1 - \zeta - \eta$, the *one-period transition probabilities* $p^a_{ij} := \mathbb{P}\left[X(t) = j | X(t-1) = i, a(t-1) = a\right]$ from state $1 \le i \le I-1$ to state j under action a are

$$p^0_{ij} = \begin{cases} \eta & \text{if } j = i-1 \\ \varepsilon & \text{if } j = i \\ \zeta & \text{if } j = i+1 \end{cases} \qquad p^1_{ij} = \begin{cases} \mu & \text{if } j = i-1 \\ 1-\mu & \text{if } j = i \end{cases} \tag{5}$$

  and for the boundary cases, $p^1_{00} = 1$, and

$$p^0_{0j} = \begin{cases} 1-\zeta & \text{if } j = 0 \\ \zeta & \text{if } j = 1 \end{cases} \qquad p^a_{Ij} = \begin{cases} \mu & \text{if } j = I-1 \\ 1-\mu & \text{if } j = I \end{cases} \tag{6}$$

  The remaining transition probabilities are zero.
- If the queue length is $i \in \mathcal{I}$ and action $a \in \mathcal{A}$ is chosen, then the gatekeeper's *one-period reward* is defined as the negative of the expected holding cost at the current epoch,

$$R^a_i := -c_i.$$

  At the same time, the gatekeeper's *one-period work* is defined as the expected number of rejected customers during the current period,

$$W^1_i := \lambda \qquad W^0_i := \begin{cases} \lambda & \text{if } i = I \\ 0 & \text{otherwise} \end{cases}$$

Thus, for rejection cost (gatekeeper's wage) ν, the *one-period overal cost* is

$$-R_i^a + \nu W_i^a = c_i + \lambda \nu a + (1-a)\mathbf{1}\{i = I\}\lambda \nu,$$

where $\mathbf{1}\{Y\}$ is the $0/1$ indicator function of statement Y.

Given the definition above, we call state I *uncontrollable*, because in this state both the actions result in identical consequences (for having identical one-period reward, one-period work, and transition probabilities), and there is actually no decision to make. This is not the case for the remaining states, henceforth called *controllable*.

Finally, to ease later reference we summarize here our model parameters assumptions:

$$0 < \beta < 1, \qquad 0 < \lambda \leq 1, \qquad 0 < \mu < 1, \qquad 0 \leq \eta < 1, \qquad 0 < \varepsilon < 1, \qquad 0 < \zeta < 1. \tag{7}$$

2.2 Admission Control Problem with Delay

In this subsection we follow the classic reformulation as MDP of problems with a discrete-time delay, which is a special case of *partially observed MDPs*, by augmenting the state space (Brooks and Leondes, 1972).

In the admission control problem with delay[1], the decision at epoch t is based on $\widetilde{X}(t) := (a(t-1), X(t-1))$, which is henceforth called an *augmented state* process. Thus, $\widetilde{X}(t)$ is the observed state at time epoch t, while $X(t)$ is the actual (hidden) queue length process. The MDP elements of the admission control problem with delay are as follows:

- The *action space* is $\mathcal{A}$ as in the no-delay problem.
- Recall that in the no-delay problem, state I is uncontrollable. Consequently, states $(0, I)$ and $(1, I)$ in the problem with delay are *duplicates*, having identical one-period reward, one-period work, and transition probabilities, so they can and should be merged into a unique state $(*, I)$. We therefore define the *augmented state space*

$$\widetilde{\mathcal{I}} := (\mathcal{A} \times \{0, 1, \ldots, I-1\}) \cup \{(*, I)\}.$$

- The *one-period transition probabilities* are

$$\begin{aligned} p^{a'}_{(a,i),(b,j)} &:= \mathbb{P}\left[\widetilde{X}(t+1) = (b,j) | \widetilde{X}(t) = (a,i), a(t) = a'\right] \\ &= \mathbb{P}\left[X(t) = j, a(t) = b | X(t-1) = i, a(t-1) = a, a(t) = a'\right] \\ &= p^a_{ij} \cdot \mathbf{1}\{a' = b\}. \end{aligned}$$

For the merged state $(*, I)$, we have $p^{a'}_{(a,i),(*,I)} := p^{a'}_{(a,i),(0,I)} + p^{a'}_{(a,i),(1,I)} = p^a_{ij}$.

[1] We use the 'tilded' notation for the delayed version when not doing so might be confusing; note that state-dependent quantities are easy to distinguish since the original state is uni-dimensional, while the augmented state of the delayed problem is bi-dimensional.

- If the current-epoch augmented state is (a, i), then the gatekeeper's *one-period reward* is defined as the negative of the expected holding cost at the current epoch,

$$\overline{R}^b_{(a,i)} := \mathbb{E}\left[R^b_{X(t)}|a(t-1) = a, X(t-1) = i\right].$$

  Similarly, the gatekeeper's *one-period work* is defined as the expected number of rejected customers during the current period,

$$\overline{W}^b_{(a,i)} := \mathbb{E}\left[W^b_{X(t)}|a(t-1) = a, X(t-1) = i\right].$$

  Thus, for rejection cost (gatekeeper's wage) ν, the *one-period overal cost* is $-\overline{R}^b_{(a,i)} + \nu\overline{W}^b_{(a,i)}$.

The above one-period reward and one-period work can be explicitly stated as follows:

$$\overline{R}^b_{(a,i)} := \begin{cases} 0, & \text{if } (a,i) = (1,0), \\ -[(1-\zeta)c_0 + \zeta c_1], & \text{if } (a,i) = (0,0), \\ -[\mu c_{i-1} + (1-\mu)c_i], & \text{if } a = 1 \text{ and } 1 \le i \le I-1, \\ -[\eta c_{i-1} + \varepsilon c_i + \zeta c_{i+1}], & \text{if } a = 0 \text{ and } 1 \le i \le I-1, \\ -[\mu c_{I-1} + (1-\mu)c_I], & \text{if } (a,i) = (*, I). \end{cases}$$

$$\overline{W}^b_{(a,i)} := \begin{cases} \lambda, & \text{if } b = 1, \\ \zeta, & \text{if } b = 0 \text{ and } (a,i) = (*, I), \\ \zeta\lambda, & \text{if } b = 0 \text{ and } (a,i) = (0, I-1), \\ 0, & \text{otherwise.} \end{cases}$$

To evaluate a policy π under the discounted criterion, we consider the following two measures. Let $\overline{g}^\pi_{(a,i)}$ be the *expected total β-discounted work* (or, the expected total β-discounted number of rejected customers) if starting from state $(a(-1), X(-1)) := (a, i)$ under policy π,

$$\overline{g}^\pi_{(a,i)} := \mathbb{E}^\pi_{(a,i)}\left[\sum_{t=0}^{\infty} \beta^t \overline{W}^{a(t)}_{(a(t-1),X(t-1))}\right].$$

Analogously for $\overline{f}^\pi_{(a,i)}$, the *expected total β-discounted reward* if starting from state $(a(-1), X(-1)) := (a, i)$ under policy π,

$$\overline{f}^\pi_{(a,i)} := \mathbb{E}^\pi_{(a,i)}\left[\sum_{t=0}^{\infty} \beta^t \overline{R}^{a(t)}_{(a(t-1),X(t-1))}\right].$$

If the rejection cost ν is interpreted as the wage paid to gatekeeper for each rejected customer, then the objective is to solve the following ν-wage problem for each ν:

$$\min_{\pi \in \Pi} -\overline{f}^{\pi}_{(a,i)} + \nu \overline{g}^{\pi}_{(a,i)}, \tag{8}$$

where Π is the set of all non-anticipative control policies.

Next we present alternative, simpler definitions of one-period work and one-period reward, and show that they lead to an equivalent problem. These alternative definitions capture the reward and work one period backwards comparing to the original ones. If the current-epoch augmented state is (a, i), then the alternative gatekeeper's one-period reward is defined as the negative of the expected holding cost at the previous epoch,

$$R^b_{(a,i)} := \beta(-c_i/\beta) = -c_i. \tag{9}$$

Similarly, the alternative gatekeeper's one-period work is defined as the expected number of rejected customers during the previous period,

$$W^b_{(1,i)} := \lambda \qquad\qquad W^b_{(0,i)} := \begin{cases} \lambda & \text{if } i = I \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

Notice that we have $R^b_{(a,i)} = R^a_i$ and $W^b_{(a,i)} = W^a_i$.

Then, the alternative expected total β-discounted work if starting from state $(a, i) := (a(-1), X(-1))$ under policy π is

$$g^{\pi}_{(a,i)} := \mathbb{E}^{\pi}_{(a,i)} \left[\sum_{t=0}^{\infty} \beta^t W^{a(t)}_{(a(t-1),X(t-1))} \right]. \tag{11}$$

Analogously, the alternative expected total β-discounted reward if starting from state $(a, i) := (a(-1), X(-1))$ under policy π is

$$f^{\pi}_{(a,i)} := \mathbb{E}^{\pi}_{(a,i)} \left[\sum_{t=0}^{\infty} \beta^t R^{a(t)}_{(a(t-1),X(t-1))} \right]. \tag{12}$$

Then, the alternative objective is

$$\min_{\pi \in \Pi} -f^{\pi}_{(a,i)} + \nu g^{\pi}_{(a,i)}, \tag{13}$$

where Π is the set of all non-anticipative control policies, and the following proposition demonstrates its equivalence to (8).

Proposition 1.

(i) For any state $(a, i) \in \widetilde{\mathcal{I}}$ and any policy $\pi \in \Pi$, $f^{\pi}_{(a,i)} = R^a_i + \beta \overline{f}^{\pi}_{(a,i)}$.
(ii) For any state $(a, i) \in \widetilde{\mathcal{I}}$ and any policy $\pi \in \Pi$, $g^{\pi}_{(a,i)} = W^a_i + \beta \overline{g}^{\pi}_{(a,i)}$.

(iii) Problems (8) *and* (13) *are equivalent.*

Proof. (i) Using the above definitions, we can write

$$\overline{f}^{\pi}_{(a,i)} = \mathbb{E}^{\pi}_{(a,i)} \left[\sum_{t=0}^{\infty} \beta^t \, \mathbb{E} \left[R^{a(t)}_{X(t)} | a(t-1), X(t-1) \right] \right] = \mathbb{E}^{\pi}_{(a,i)} \left[\sum_{t=0}^{\infty} \beta^t R^{a(t)}_{X(t)} \right],$$

where the last equality follows from Fubini's theorem and from the law of total expectation.

On the other hand, we have

$$f^{\pi}_{(a,i)} = \mathbb{E}^{\pi}_{(a,i)} \left[\sum_{t=0}^{\infty} \beta^t R^{a(t-1)}_{X(t-1)} \right],$$

hence we obtain identity $f^{\pi}_{(a,i)} = R^a_i + \beta \overline{f}^{\pi}_{(a,i)}$. □

(ii) Analogously to (i). □

(iii) Since, for all (a,i), $-R^a_i + \nu W^a_i$ is constant, $\beta > 0$, and the identities in (i) and (ii) hold, (13) is equivalent to (8). □

Notice that the alternative one-period reward $R^b_{(a,i)}$ and the one-period work $W^b_{(a,i)}$ are independent of the current-epoch action (superscript b), therefore we will omit the superscript in the remaining sections.

3 Restless Bandit Indexation

In the previous section we have formulated the admission control problem with delay as a binary-action Markov decision process (MDP), i.e., a *restless bandit*, where shutting the entry gate corresponds to the active action, and opening it as the passive action.

We next address such a problem by deploying a restless bandit indexation approach, following the seminal idea introduced in Whittle (1988) and developed by the second author, in work surveyed in Niño-Mora (2007b). We focus on the finite-buffer problem under the discounted criterion. The solution to the problem under the time-average criterion is treated in subsection 4.6.

MDP theory ensures existence of an optimal policy that is stationary, deterministic and independent of the initial state. We represent a stationary deterministic policy in terms of an *active set* $\mathcal{S} \subseteq \widetilde{\mathcal{I}}$, i.e., the set of states in which it prescribes to shut the gate; in the remaining states it prescribes to let the gate open. The problem to find an optimal admission control policy is thus reduced to finding an optimal active set,

$$\min_{\mathcal{S} \subseteq \widetilde{\mathcal{I}}} -f^{\mathcal{S}}_{(a,i)} + \nu g^{\mathcal{S}}_{(a,i)}. \tag{14}$$

For every rejection cost ν, the optimal policy is characterized by the unique solution vector $(v^*_{(a,i)}(\nu))_{(a,i)\in\widetilde{\mathcal{I}}}$ to the *Bellman equations*

$$v^*_{(a,i)}(\nu) = \min_{a'\in\mathcal{A}} \left[-R_{(a,i)} + \nu W_{(a,i)} - \beta \sum_{(b,j)\in\widetilde{\mathcal{I}}} p^{a'}_{(a,i),(b,j)} v^*_{(b,j)}(\nu) \right], \qquad (a,i)\in\widetilde{\mathcal{I}} \tag{15}$$

where $v^*_{(a,i)}(\nu)$ denotes the optimal value of (13) starting at (a,i) under rejection cost ν. Hence, there exists a *maximal optimal active set* (i.e., a set of states in which it is optimal to close the gate) $\mathcal{S}^*(\nu) \subseteq \widetilde{\mathcal{I}}$ for (13), which is characterized by

$$\mathcal{S}^*(\nu) := \left\{ (a,i)\in\widetilde{\mathcal{I}} : \sum_{(b,j)\in\widetilde{\mathcal{I}}} p^0_{(a,i),(b,j)} v^*_{(b,j)}(\nu) \leq \sum_{(b,j)\in\widetilde{\mathcal{I}}} p^1_{(a,i),(b,j)} v^*_{(b,j)}(\nu) \right\}.$$

Problem (14) can be viewed as a bi-criteria parametric optimization problem. Intuitively, if the rejection cost $\nu \to -\infty$, the optimal active set should be $\widetilde{\mathcal{I}}$, whereas if the rejection cost $\nu \to \infty$, the optimal active set should be the empty set. In fact, we set out to show a stronger, so-called *indexability* property: Active sets $\mathcal{S}^*(\nu)$ diminish monotonically from $\widetilde{\mathcal{I}}$ to the empty set as the rejection cost ν increases from $-\infty$ to ∞. Such a property was introduced in Whittle (1988) for the restless bandits with one-periods works equal to 1 under the active action, and equal to 0 under the passive action, and extended to restless bandits without these limitations in Niño-Mora (2002).

Such an indexability property is equivalent to existence of break-even values $\nu^{\mathrm{MPI}}_{(a,i)}$ of the rejection cost ν attached to augmented states $(a,i)\in\widetilde{\mathcal{I}}$, which characterize the optimal policies for (14) as follows: it is optimal to take the active action when the system occupies augmented state (a,i) if $\nu^{\mathrm{MPI}}_{(a,i)} \geq \nu$, and it is optimal to take the passive action when the system occupies augmented state (a,i) if $\nu^{\mathrm{MPI}}_{(a,i)} \leq \nu$. Since we have defined $\mathcal{S}^*(\nu)$ as the *maximal* optimal active set, state $(a,i)\in\mathcal{S}^*(\nu)$ if $\nu = \nu_{(a,i)}$, though this choice is arbitrary. We will refer to index $\nu^{\mathrm{MPI}}_{(a,i)}$ as the *marginal productivity index* (MPI), after its economic interpretation as the marginal productivity of work at state (a,i), as elucidated in Niño-Mora (2002, 2006b).

3.1 Exploiting Special Structure

While one could test numerically whether a given instance is indexable and calculate the indices $\nu_{(a,i)}$ for all $(a,i)\in\widetilde{\mathcal{I}}$, we aim instead to establish analytically indexability of the admission control problem with delay in general. This will further allow us to achieve our second objective of obtaining a fast way of computing the indices. In this subsection we present how to exploit special structure of the model by aligning indexability to a known family of optimal bi-threshold policies.

Suppose that we postulate a family $\mathcal{F} \subseteq 2^{\widetilde{\mathcal{I}}}$ of active sets, satisfying certain connectivity conditions (see Niño-Mora (2007b) for the details). Before presenting such a family for the admission control problem with delay, we review a test (deployed in section 4) to verify whether a postulated family $\mathcal{F}$ can be used to establish indexability, via the sufficient conditions termed PCL($\mathcal{F}$)-indexability introduced in Niño-Mora (2001, 2002).

```
Ŝ_0 := Ĩ;
for k = 1 to 2I+1 do
    pick (a_k, i_k) ∈ arg min { ν^{Ŝ_{k-1}}_{(a,i)} : (a,i) ∈ Ŝ_{k-1} and Ŝ_{k-1} \ {(a,i)} ∈ F };
    ν̂_{(a_k,i_k)} := ν^{Ŝ_{k-1}}_{(a_k,i_k)};
    Ŝ_k := Ŝ_{k-1} \ {(a_k,i_k)};
end {for};
{Output {Ŝ_k}_{k=0}^{2I+1}, {ν̂_{(a_k,i_k)}}_{k=1}^{2I+1}}
```

Figure 2. Algorithmic scheme of $AG_{\mathcal{F}}$.

Let policy $\langle a, \mathcal{S}\rangle$ be the policy where action a is applied in the current period and policy $\mathcal{S}$ proceeds. Notice that policy $\langle a, \mathcal{S}\rangle$ implies that the next-epoch augmented state will be (a, j) for some state $j \in \mathcal{I}$. We define the *marginal work* of closing the gate instead of letting it open (or, of rejecting possible customers instead of admitting them), if starting from state (a, i) under active-set policy $\mathcal{S}$, as

$$w^{\mathcal{S}}_{(a,i)} := g^{\langle 1,\mathcal{S}\rangle}_{(a,i)} - g^{\langle 0,\mathcal{S}\rangle}_{(a,i)}, \tag{16}$$

i.e., as the increment in total work that results from closing the gate instead of opening it at current epoch. Analogously, we define the *marginal reward*,

$$r^{\mathcal{S}}_{(a,i)} := f^{\langle 1,\mathcal{S}\rangle}_{(a,i)} - f^{\langle 0,\mathcal{S}\rangle}_{(a,i)}, \tag{17}$$

as the analogous increment in total reward. Finally, we define the *marginal productivity rate*

$$\nu^{\mathcal{S}}_{(a,i)} := \frac{r^{\mathcal{S}}_{(a,i)}}{w^{\mathcal{S}}_{(a,i)}}, \tag{18}$$

provided that the denominator does not vanish. As we will see, the denominator is positive for the admission control problem with delay. It can be shown that if the indices exist, then $\nu_{(a,i)} = \nu^{\mathcal{S}}_{(a,i)}$ for some active set $\mathcal{S}$, and therefore the indices are appropriately called the *marginal productivity indices*.

In Figure 2 is given a scheme of the *adaptive-greedy* algorithm $AG_{\mathcal{F}}$, which calculates the candidates for the maximal optimal active sets $\{\widehat{\mathcal{S}}_k\}_{k=0}^{2I+1}$ and the candidates for the marginal productivity indices $\{\widehat{\nu}_{i_k}\}_{k=1}^{2I+1}$. It is greedy, since in each step it picks the state with the lowest marginal productivity rate $\nu^{\widehat{\mathcal{S}}_{k-1}}_{(a_k,i_k)}$ (out of the feasible ones), and it is adaptive, because in each step it updates the marginal productivity rates for the actual active set $\widehat{\mathcal{S}}_{k-1}$.

Now we are ready to define PCL($\mathcal{F}$)-indexability, based on partial conservation laws (PCL), which determines both the computational and analytical value of the adaptive-greedy algorithm $AG_{\mathcal{F}}$.

Definition 1 (PCL($\mathcal{F}$)-indexability). *The admission control problem with delay is called* PCL($\mathcal{F}$)-indexable, *if* *(i) [Positive Marginal Works under $\mathcal{F}$] for each active set $\mathcal{S} \in \mathcal{F}$ and for each controllable state $(a,i) \in \widetilde{\mathcal{I}}$, the marginal work $w^{\mathcal{S}}_{(a,i)} > 0$;*

and either of the following conditions holds:

(ii) for every rejection cost ν, there exists an optimal active set $\mathcal{S} \in \mathcal{F}$;
(ii') the output $\{\widehat{\nu}_{(a_k,i_k)}\}_{k=1}^{2I+1}$ of the algorithm $AG_{\mathcal{F}}$ are marginal productivity indices in nondecreasing order.

Niño-Mora (2001, 2002, 2007b) introduced variants of PCL($\mathcal{F}$)-indexability and proved that PCL($\mathcal{F}$)-indexability implies indexability, i.e., the existence of marginal productivity indices, which are calculated as $\{\widehat{\nu}_{(a_k,i_k)}\}_{k=1}^{2I+1}$ by the adaptive-greedy algorithm $AG_{\mathcal{F}}$. To ease later reference, we summarize the above in the following theorem.

Theorem 1. *If marginal works are positive under $\mathcal{F}$ (cf. Definition 1(i)) for problem (13), then for that problem the following statements are equivalent:*

(i) for every rejection cost ν, there exists a maximal optimal active set $\mathcal{S} \in \mathcal{F}$;
(ii) the problem is indexable and all active sets $\mathcal{S}^*(\nu) \in \mathcal{F}$;
(iii) the output $\{\widehat{\nu}_{(a_k,i_k)}\}_{k=1}^{2I+1}$ of the algorithm $AG_{\mathcal{F}}$ are marginal productivity indices in nondecreasing order.

Proof. ADD A PROOF IN THE APPENDIX. □

In section 4 we show that for a certain family $\mathcal{F}$ (defined below), Definition 1(i) holds and, given the existing results, Theorem 1(i) is true. In this way indexability of the admission control problem with delay will be established, and the algorithm $AG_{\mathcal{F}}$ can be used to obtain the indices.

Definition 1(i) has an intuitive interpretation (cf. Niño-Mora, 2002, Proposition 6.2): positivity of marginal work $w^{\mathcal{S}}_{(a,i)}$ (where $\mathcal{S} \in \mathcal{F}$ and state $(a,i) \in \widetilde{\mathcal{I}}$ is controllable) is equivalent to monotonicity of total work,

$$\begin{aligned} g^{\mathcal{S}\setminus\{(a,i)\}}_{(a,i)} &< g^{\mathcal{S}}_{(a,i)}, && \text{if } (a,i) \in \mathcal{S}, \\ g^{\mathcal{S}}_{(a,i)} &< g^{\mathcal{S}\cup\{(a,i)\}}_{(a,i)}, && \text{if } (a,i) \notin \mathcal{S}. \end{aligned}$$

Informally stated, rejecting in a larger number of states corresponds to a larger expected total discounted number of rejected customers. Definition 1(i) is a natural assumption in many models, though, in general, it is neither a sufficient nor a necessary condition for indexability.

3.2 Postulated Active-Set Family

We use the results of Altman and Nain (1992), who characterized the optimal bi-threshold policies, and identify an active-set family $\mathcal{F}$ for which Theorem 1(i) holds. A bi-threshold active-set policy with open-gate threshold K_0 and closed-gate threshold K_1 will be denoted by

$$\widetilde{\mathcal{I}}_{K_0,K_1} := \{(0,K_0),(0,K_0+1),\ldots,(0,I)\} \cup \{(1,K_1),(1,K_1+1),\ldots,(1,I)\}, \tag{19}$$

which is well-defined for all $0 \le K_0, K_1 \le I+1$ except the active sets $\widetilde{\mathcal{I}}_{I+1,I}$ and $\widetilde{\mathcal{I}}_{I,I+1}$, because states $(1, I)$ and $(0, I)$ are duplicates, and by definition either both or none of them can belong to $\widetilde{\mathcal{I}}_{K_0,K_1}$.

In words, active set $\widetilde{\mathcal{I}}_{K_0,K_1}$ prescribes to open or close the gate depending on the previous-epoch action and previous-epoch state. If the gate was open in the previous period, then we open the gate if and only if the queue length in the previous epoch was equal to or larger than the open-gate threshold K_0. Similarly, if the gate was closed in the previous period, then we open the gate if and only if the queue length in the previous epoch was equal to or larger than the closed-gate threshold K_1.

Intuitively, if an active set $\widetilde{\mathcal{I}}_{K_0,K_1}$ is optimal for some rejection cost ν, then $K_0 \le K_1$. Indeed, for a given previous-epoch queue length, we would be less prone to close the gate if it was closed than if it was open in the preceding period, because the queue length could not get larger under a closed gate, and therefore the rejection costs become relatively more harmful than the holding costs. On the other hand, it can be shown that $K_1 \le K_0 + 1$ (see below). Thus, the postulated family of optimal active sets for the admission control problem with delay is

$$\mathcal{F} := \{\widetilde{\mathcal{I}}_{K,K} : K = 0, 1, \ldots, I+1\} \cup \{\widetilde{\mathcal{I}}_{K,K+1} : K = 0, 1, \ldots, I-1\}. \tag{20}$$

Theorem 2 (Altman and Nain (1992), Theorem 3.1). *If the holding cost c_i is nondecreasing and convex on $\mathcal{I}$, then $\mathcal{F}$ as defined in (20) contains an optimal active set for every rejection cost ν.*

Though the above result was shown for the problem with infinite buffer, it directly applies to the finite-buffer variant. Notice that if a bi-threshold policy is optimal for the infinite-buffer problem, then it is also optimal for all problems with buffer equal to or larger than both the thresholds. If the buffer is smaller than the larger optimal threshold (K_1), then it is optimal open the gate all the time.

For active-set family $\mathcal{F}$ given in (20), picking (a_k, i_k) becomes trivial, because there is only a unique feasible augmented state in each step. For instance, in step $k = 1$, only state $(1, 0)$ both belongs to $\widehat{\mathcal{S}}_0$ and $\widehat{\mathcal{S}}_0 \setminus \{(1, 0)\} = \widetilde{\mathcal{I}}_{0,1} \in \mathcal{F}$, since $\widehat{\mathcal{S}}_0 := \widetilde{\mathcal{I}} = \widetilde{\mathcal{I}}_{0,0}$. Similarly, in step $k = 2$, only state $(0, 0)$ both belongs to $\widehat{\mathcal{S}}_1$ and $\widehat{\mathcal{S}}_1 \setminus \{(0, 0)\} = \widetilde{\mathcal{I}}_{1,1} \in \mathcal{F}$. In general, $(a_k, i_k) = (0, (k/2) - 1)$ for all even $1 \le k \le 2I$, and $(a_k, i_k) = (1, (k-1)/2)$ for all odd $1 \le k \le 2I$. Finally, in step $k = 2I + 1$, the picked state is $(*, I)$.

To summarize, the sequence of candidate active sets $\{\widehat{\mathcal{S}}_k\}_{k=0}^{2I+1}$ in algorithm $AG_{\mathcal{F}}$ under active-set family $\mathcal{F}$ given in (20) is

$$\begin{aligned}\widehat{\mathcal{S}}_0 = \widetilde{\mathcal{I}} = \widetilde{\mathcal{I}}_{0,0}, \widehat{\mathcal{S}}_1 = \widetilde{\mathcal{I}}_{0,1}, \widehat{\mathcal{S}}_2 = \widetilde{\mathcal{I}}_{1,1}, \widehat{\mathcal{S}}_3 = \widetilde{\mathcal{I}}_{1,2}, \widehat{\mathcal{S}}_4 = \widetilde{\mathcal{I}}_{2,2}, \ldots \\ \ldots, \widehat{\mathcal{S}}_{2I-1} = \widetilde{\mathcal{I}}_{I-1,I}, \widehat{\mathcal{S}}_{2I} = \widetilde{\mathcal{I}}_{I,I}, \widehat{\mathcal{S}}_{2I+1} = \widetilde{\mathcal{I}}_{I+1,I+1} = \emptyset,\end{aligned} \tag{21}$$

and the sequence of picked states $\{(a_k, i_k)\}_{k=1}^{2I+1}$ is

$$\begin{aligned}(a_1, i_1) = (1, 0), (a_2, i_2) = (0, 0), (a_3, i_3) = (1, 1), (a_4, i_4) = (0, 1), \ldots \\ \ldots, (a_{2I-1}, i_{2I-1}) = (1, I-1), (a_{2I}, i_{2I}) = (0, I-1), (a_{2I+1}, i_{2I+1}) = (*, I).\end{aligned}$$

Given the above, in Figure 3 we present the reduction of the algorithmic scheme $AG_{\mathcal{F}}$ as it applies to the postulated family $\mathcal{F}$ given in (20). Notice that the computational complexity remains at the same level since the main difficulty lies in the calculation of $\nu_{(a_k,i_k)}^{\widehat{\mathcal{S}}_{k-1}}$, for which no computational details are

```
for K = 1 to I do
    ν̂_(1,K−1) := ν_(1,K−1)^{Ĩ_{K−1,K−1}};
    ν̂_(0,K−1) := ν_(0,K−1)^{Ĩ_{K−1,K}};
end {for};
ν̂_(*,I) := ν_(*,I)^{Ĩ_{I,I}};
{Output {ν̂_(a,i)}_{(a,i)∈Ĩ}}
```

Figure 3. Algorithmic scheme of $AG_{\mathcal{F}}$ under active-set family $\mathcal{F}$ given in (20).

given. Therefore we also call them algorithmic schemes, not algorithms. The goal of this paper is to establish the validity of $AG_{\mathcal{F}}$ for our problem and to develop its implementation of low computational complexity.

4 Results

In this section we focus on the admission control problem with delay to a buffer (i.e., $I \geq 2$) under the discounted criterion. The case $I = 1$, referring to the admission control problem with delay to server with no dedicated buffer, is treated in subsection 4.5. The results under the time-average criterion are summarized in subsection 4.6.

Our main results are twofold. First, we prove the positivity of marginal works (cf. Definition 1(i)) for $\mathcal{F}$ given in (20), so that the algorithm $AG_{\mathcal{F}}$ can be applied to compute the indices. Second, we simplify $AG_{\mathcal{F}}$ obtaining a procedure that performs only a linear number of arithmetic operations to compute all the indices and the optimal thresholds.

Let us introduce a more compact notation. For any augmented-state-dependent variable $x_{(a,i)}$, we will use the backward difference operator in the first dimension, i.e., the *action-difference operator*,

$$\Delta_1 x_{(1,i)} := x_{(1,i)} - x_{(0,i)} \tag{22}$$

and in the second dimension, i.e., the *state-difference operator*,

$$\Delta_2 x_{(a,i)} := x_{(a,i)} - x_{(a,i-1)} \tag{23}$$

whenever the right-hand side expressions are defined. For definiteness, we further let $\Delta_2 x_{(a,0)} := 0$ for $a \in \mathcal{A}$. Directly from these definitions we obtain the following auxiliary identity,

$$\Delta_2 x_{(1,i)} - \Delta_2 x_{(0,i)} = \Delta_1 x_{(1,i)} - \Delta_1 x_{(1,i-1)}. \tag{24}$$

In the following we list our main results, drawing on the technical analysis of work measures presented in the appendix (section A).

Proposition 2.

(i) The marginal works in problem (14) are positive under the active-set family $\mathcal{F}$ given in (20), i.e., Definition 1(i) holds.

(ii) If the holding cost c_i is nondecreasing and convex on $\mathcal{I}$, then the admission control problem with delay in (14) is PCL($\mathcal{F}$)-indexable, and therefore it is indexable and algorithm $AG_{\mathcal{F}}$ calculates the marginal productivity indices for this problem.

Proof.

(i) By Proposition A2(iii) and Proposition A3(iv), $\Delta_1 g^{\mathcal{S}}_{(1,i)} > 0$ for all $0 \leq i \leq I-1$ and $\Delta_1 g^{\mathcal{S}}_{(1,I)} = 0$ under every active set $\mathcal{S} \in \mathcal{F}$. Then, Lemma A4 establishes the positivity of marginal works for all states. □

(ii) Due to (i), Theorem 1(i)-(iii) are equivalent. The validity of Theorem 1(i) was established in Theorem 2, therefore Theorem 1(iii) holds, and implies the above claim. □

4.1 A Fast Algorithm for Calculation of All Marginal Productivity Indices

Suppose that the holding cost c_i is nondecreasing and convex on $\mathcal{I}$. In the following we develop an algorithm for calculation of *all* marginal productivity indices in $\mathcal{O}(I)$, which is two orders of magnitude faster that the best general implementation of algorithm $AG_{\mathcal{F}}$ performing $\mathcal{O}(I^3)$ arithmetic operations.

The algorithmic scheme $AG_{\mathcal{F}}$ in Figure 3 is exhibited in its *bottom-up* version, as it calculates the marginal productivity indices in nondecreasing order (cf. Definition 1(ii')). This is closely related to our definition of indexability in section 3 as the property that "active sets $\mathcal{S}^*(\nu)$ diminish monotonically from $\widetilde{\mathcal{I}}$ to the empty set as the rejection cost ν increases from $-\infty$ to ∞," being emulated by the bottom-up version of the algorithm. Notice that we could equivalently define indexability as "active sets $\mathcal{S}^*(\nu)$ expand monotonically from the empty set to $\widetilde{\mathcal{I}}$ as the rejection cost ν decreases from ∞ to $-\infty$." This intuitively leads to consideration of algorithm $AG_{\mathcal{F}}$ in its equivalent, *top-down* version, starting with the empty set and calculating the indices in nonincreasing order.

In other words, while the bottom-up version of algorithm $AG_{\mathcal{F}}$ traverses the active-set family $\mathcal{F}$ in the order (cf. (21))

$$\widetilde{\mathcal{I}}_{0,0}, \widetilde{\mathcal{I}}_{0,1}, \widetilde{\mathcal{I}}_{1,1}, \widetilde{\mathcal{I}}_{1,2}, \ldots, \widetilde{\mathcal{I}}_{I-1,I}, \widetilde{\mathcal{I}}_{I,I}, \widetilde{\mathcal{I}}_{I+1,I+1},$$

the top-down version does that in the reverse order

$$\widetilde{\mathcal{I}}_{I+1,I+1}, \widetilde{\mathcal{I}}_{I,I}, \widetilde{\mathcal{I}}_{I-1,I}, \ldots, \widetilde{\mathcal{I}}_{1,2}, \widetilde{\mathcal{I}}_{1,1}, \widetilde{\mathcal{I}}_{0,1}, \widetilde{\mathcal{I}}_{0,0}.$$

For instance, index $\nu_{(1,0)}$ is calculated as the marginal productivity rate $\nu^{\widetilde{\mathcal{I}}_{0,0}}_{(1,0)}$ in the bottom-up version, while the same index is calculated as the marginal productivity rate $\nu^{\widetilde{\mathcal{I}}_{0,1}}_{(1,0)}$ in the top-down version. In fact, Niño-Mora (2002, Theorem 6.4(b)) implies that $\nu^{\widehat{\mathcal{S}}_{k-1}}_{(a_k,i_k)} = \nu^{\widehat{\mathcal{S}}_k}_{(a_k,i_k)}$, using the notation of Figure 2. Thus, since the active set of type $\widetilde{\mathcal{I}}_{K,K}$ is efficient every two steps of the algorithm (except for the last step, where $\widetilde{\mathcal{I}}_{I+1,I+1}$ follows $\widetilde{\mathcal{I}}_{I,I}$), we can formulate the indices in terms of marginal productivity rates under active sets $\widetilde{\mathcal{I}}_{K,K}$ only. Such an algorithmic scheme is presented in Figure 4.

```
ν_(1,0) := ν_(1,0)^{Ĩ_{0,0}};
for K = 1 to I − 1 do
    ν_(0,K−1) := ν_(0,K−1)^{Ĩ_{K,K}};
    ν_(1,K) := ν_(1,K)^{Ĩ_{K,K}};
end {for};
ν_(0,I−1) := ν_(0,I−1)^{Ĩ_{I,I}};
ν_(*,I) := ν_(*,I)^{Ĩ_{I+1,I+1}};
{Output {ν_(a,i)}_{(a,i)∈Ĩ}}
```

Figure 4. Algorithmic scheme of calculation of marginal productivity indices for the admission control problem with delay in terms of active sets $\widetilde{\mathcal{I}}_{K,K}$ only.

Next we develop an efficient implementation of the algorithmic scheme $AG_{\mathcal{F}}$, which we present in Figure 5. The algorithm FA is two orders of magnitude faster than the best existing general implementation of the algorithm $AG_{\mathcal{F}}$. We characterize the marginal productivity indices calculated as indicated in Figure 4 in terms of closed-form expressions of pivot state-differences given in Lemma B5 and Lemma A7. Notice that one iteration of the algorithm solves the problem (14) for the entire range of the real-valued rejection cost parameter ν.

Proposition 3.

(i) The algorithm FA in Figure 5 computes the marginal productivity indices for problem (14) under the discounted criterion.

(ii) The algorithm FA in Figure 5 performs $\mathcal{O}(I)$ arithmetic operations.

Proof. (i) The algorithm FA is an implementation of expressions of marginal productivity rates $\nu_{(a,i)}^{\widetilde{\mathcal{I}}_{K,K}}$ developed below into the algorithmic scheme given in Figure 4.

The marginal productivity rates are, by definition (18), computed as the ratio of marginal rewards to marginal works. These quantities are given in Lemma B3 and Lemma A4, respectively, in terms of their action-differences, which can be expressed in terms of pivot state-differences due to Proposition B1 and Proposition A1. The pivot state-differences $\Delta_2 f_{(0,K)}^{\widetilde{\mathcal{I}}_{K,K}}, \Delta_2 f_{(1,K)}^{\widetilde{\mathcal{I}}_{K,K}}, \Delta_2 g_{(0,K)}^{\widetilde{\mathcal{I}}_{K,K}}, \Delta_2 g_{(1,K)}^{\widetilde{\mathcal{I}}_{K,K}}$ given in Lemma B5 and Lemma A7 are in Figure 5 briefly denoted by f^0, f^1, g^0, g^1, respectively. □

(ii) The number of arithmetic operations in "Initialization" and in "Finalization" is constant (with respect to I). Similarly, at each step of "Loop", a constant number of arithmetic operations is performed. Since there are $\mathcal{O}(I)$ steps of "Loop", the overal complexity of the algorithm is $\mathcal{O}(I)$. □

Once the optimal index policy is known, the optimal thresholds for a given rejection cost ν can easily be obtained. The optimal open-gate threshold is

$$K_0 := \min\{i \in \mathcal{I} : \nu_{(0,i)} \geq \nu\}.$$

$$\{\texttt{Input } I, \lambda, \mu, \beta\}$$

$$\{\texttt{Initialization}\}$$

$$\zeta := \lambda(1-\mu); \quad \eta := \mu(1-\lambda); \quad \varepsilon := 1 - \zeta - \eta;$$

$$A_0 := 0; \quad A'_0 := \beta\zeta; \quad B := \beta\mu/(1-\beta+\beta\mu); \quad B' := \beta\zeta B + \beta(\mu - \eta);$$

$$C''_{0,0} := \Delta_1 R_{(1,0)} - \zeta B \Delta_2 R_{(1,1)}/\mu; \quad D_0 := 0; \quad D'_0 := \Delta_1 R_{(1,0)};$$

$$\nu_{(1,0)} := C''_{0,0}/\lambda;$$

$$\{\texttt{Loop}\}$$

$$\texttt{for } K = 1 \texttt{ to } I-1 \texttt{ do}$$

$$A_K := \beta\zeta/[1-\beta+\beta\zeta+\beta\eta(1-A_{K-1})]; \quad A'_K := \beta\zeta + \beta(\mu-\eta)A_K; \quad Z_K := A_K A'_{K-1}/A'_K;$$

$$C''_{K,K} := \Delta_1 R_{(1,K)} - \zeta B \Delta_2 R_{(1,K+1)}/\mu;$$

$$C''_{K,K+1} := \Delta_1 R_{(1,K+1)} - \zeta B^2 \Delta_2 R_{(1,K+1)}/\mu - \zeta B \Delta_2 R_{(1,K+2)}/\mu;$$

$$D_K := \left(\beta\eta D_{K-1} - \Delta_2 R_{(0,K)}\right) A_K/(\beta\zeta); \quad D'_K := \Delta_1 R_{(1,K)} + \beta(\mu-\eta) D_K;$$

$$f^0 := -\frac{\frac{\beta\zeta}{A_K} D_K - \Delta_1 R_{(1,K)} + (1-\beta) C''_{K,K+1} - (\beta D'_{K-1} + \Delta_2 R_{(1,K)}) B' + \beta\mu(\beta\zeta D_{K-1} B + C''_{K,K+1} - \Delta_1 R_{(1,K-1)})}{\frac{A'_K}{A_K} + \beta A'_{K-1} B' + \beta\zeta\beta\mu(1 - B A_{K-1})};$$

$$f^1 := -\frac{\frac{\beta\zeta}{A_K} D_K - \Delta_1 R_{(1,K)} + (1-\beta) D'_{K-1} - (\beta C''_{K,K+1} + \Delta_2 R_{(1,K)}) A'_{K-1} + \beta\mu(\beta\zeta D_{K-1} + (C''_{K,K+1} - \Delta_1 R_{(1,K-1)}) A_{K-1})}{\frac{A'_K}{A_K} + \beta A'_{K-1} B' + \beta\zeta\beta\mu(1 - B A_{K-1})};$$

$$g^0 := \frac{\beta\lambda\,(1+B')}{\frac{A'_K}{A_K} + \beta A'_{K-1} B' + \beta\zeta\beta\mu(1 - B A_{K-1})}; \quad g^1 := \frac{1 + A'_{K-1}}{1+B'} g^0;$$

$$\texttt{if } K = 1 \texttt{ then}$$

$$\nu_{(0,0)} := \frac{(1-\zeta) D'_0 + \zeta C''_{1,1} - (1-\zeta) A'_0 f^0 - \zeta B' f^1}{\lambda - (1-\zeta) A'_0 g^0 - \zeta B' g^1};$$

$$\texttt{else}$$

$$\nu_{(0,K-1)} := \frac{\eta\left[D'_{K-2} + D_{K-1} A'_{K-2}\right] + \varepsilon D'_{K-1} + \zeta C''_{K,K} - [\eta Z_{K-1} + \varepsilon] A'_{K-1} f^0 - \zeta B' f^1}{\lambda - [\eta Z_{K-1} + \varepsilon] A'_{K-1} g^0 - \zeta B' g^1};$$

$$\texttt{end } \{\texttt{if}\};$$

$$\nu_{(1,K)} := \frac{\mu D'_{K-1} + (1-\mu) C''_{K,K} - \mu A'_{K-1} f^0 - (1-\mu) B' f^1}{\lambda - \mu A'_{K-1} g^0 - (1-\mu) B' g^1};$$

$$\texttt{end } \{\texttt{for}\};$$

$$\{\texttt{Finalization}\}$$

$$A_I := \beta\zeta/[1-\beta+\beta\zeta+\beta\eta(1-A_{I-1})]; \quad A'_I := \beta\zeta + \beta(\mu-\eta) A_I; \quad Z_I := A_I A'_{I-1}/A'_I;$$

$$D_I := \left(\beta\eta D_{I-1} - \Delta_2 R_{(0,I)}\right) A_I/(\beta\zeta);$$

$$f^0 := -\frac{\frac{\beta\zeta}{A_I} D_I - \beta\mu D'_{I-1}}{\frac{A'_I}{A_I} + \beta\mu A'_{I-1}}; \quad g^0 := \frac{\lambda(1+\beta\mu)}{\frac{A'_I}{A_I} + \beta\mu A'_{I-1}};$$

$$\nu_{(0,I-1)} := \frac{\eta\left[D'_{I-2} + D_{I-1} A'_{I-2}\right] + \varepsilon D'_{I-1} - [\eta Z_{I-1} + \varepsilon] A'_{I-1} f^0}{(\eta+\varepsilon)\lambda - [\eta Z_{I-1} + \varepsilon] A'_{I-1} g^0};$$

$$\nu_{(*,I)} := \frac{D'_{I-1} + \frac{\beta\zeta}{A_I} D_I Z_I}{\lambda(1 - Z_I)};$$

$$\{\texttt{Output } \left\{\nu_{(a,i)}\right\}_{(a,i)\in\widetilde{\mathcal{I}}}\}$$

Figure 5. Fast algorithm FA for calculation of an optimal index policy under general rewards.

Similarly, the optimal closed-gate threshold is

$$K_1 := \min\{i \in \mathcal{I} : \nu_{(1,i)} \geq \nu\}.$$

If $\nu > \nu_{(*,I)}$, then $K_0 := I + 1$ and $K_1 := I + 1$.

4.2 A Fast Algorithm for Calculation of One Marginal Productivity Index

In this subsection we develop an algorithm for calculation of *one* marginal productivity index, say of state (a, K) in isolation. We show that it performs at most $\mathcal{O}(\log_2 K)$ arithmetic operations, given that it may require to calculate an integer power, for which the best known algorithm (*exponentiation by squaring*) needs $\mathcal{O}(\log_2 K)$ operations.

The idea is to perform step K of the "Loop" of algorithm FA in Figure 5, which performs a constant number of arithmetic operations, having calculated A_K and D_K using their respective closed-form formulae. If $\eta = 0$, then A_K is constant and requires a constant number of arithmetic operations to calculate. In the following we assume $\eta > 0$ and denote by

$$s := \frac{\zeta}{\eta}, \qquad t := \frac{1 - \beta + \beta\zeta + \beta\eta}{\beta\eta}, \qquad u := \sqrt{t^2 - 4s}, \qquad u_+ := \frac{t+u}{2}, \qquad u_- := \frac{t-u}{2}. \tag{25}$$

The above quantities are well defined given the model parameters assumptions. The following lemma characterizes A_K and D_K in terms of K-th powers of u_+ and u_-.

Lemma 1.

(i) For any $K \geq 1$,

$$A_K = s\frac{u_+^K - u_-^K}{u_+^{K+1} - u_-^{K+1}}.$$

(ii) If $\Delta_2 R_{(1,i)} = R$ *for all* $i \geq 1$*, then for any* $K \geq 1$,

$$D_K = -\frac{R}{1-\beta}\left[\frac{u + (u_+^K - u_-^K)u_- u_+}{u_+^{K+1} - u_-^{K+1}} - 1\right].$$

(iii) Sequences A_K *and (if* $\Delta_2 R_{(1,i)} = R$ *for all* $i \geq 1$*)* D_K *converge as* $K \to \infty$ *to their respective limits*

$$A = \frac{s}{u_+}, \qquad\qquad D = \frac{c}{1-\beta}\left(u_- - 1\right).$$

Proof. (i) We start by developing the formula for A_K. By definition of A_K in (39) and those of s and t in (25) we can write $A_K = \frac{s}{-A_{K-1} + t}$. Notice that this is a well-defined Möbius transformation $m(x) := \frac{0 \cdot x + s}{-1 \cdot x + t}$ represented in the matrix form as

$$\boldsymbol{M} := \begin{pmatrix} 0 & s \\ -1 & t \end{pmatrix}$$

Therefore, A_K expressed in terms of A_0 is given by the K-th functional power of $m(x)$ for $x = A_0$, i.e., $A_K = m^K(A_0) = \dfrac{a_K \cdot A_0 + b_K}{c_K \cdot A_0 + d_K}$. Since $A_0 = 0$ by definition, we obtain $A_K = b_K/d_K$. By properties of Möbius transformation, we have

$$\begin{pmatrix} a_K & b_K \\ c_K & d_K \end{pmatrix} = \boldsymbol{M}^K.$$

Since

$$\boldsymbol{M}^K = \begin{pmatrix} 0 & s \\ -1 & t \end{pmatrix} \begin{pmatrix} a_{K-1} & b_{K-1} \\ c_{K-1} & d_{K-1} \end{pmatrix} = \begin{pmatrix} s \cdot c_{K-1} & s \cdot d_{K-1} \\ -a_{K-1} + t \cdot c_{K-1} & -b_{K-1} + t \cdot d_{K-1} \end{pmatrix}$$

we have $b_K = s \cdot d_{K-1}$ and hence $A_K = sd_{K-1}/d_K$.
Next, we set out to obtain a closed-form solution for sequence d_K. Using the above identity for b_{K-1}, we have $d_K = -s \cdot d_{K-2} + t \cdot d_{K-1}$. This recurrence is one of the two Lucas sequences (cf. Lucas, 1878; Kalman and Mena, 2003) with the initial values $d_0 = 1$ and $d_1 = t$ obtained from the definition of matrix M and the relationship $b_1 = s \cdot d_0$. Its closed-form solution is

$$d_K = \frac{u_+^{K+1} - u_-^{K+1}}{u}$$

and therefore

$$A_K = s \frac{u_+^K - u_-^K}{u_+^{K+1} - u_-^{K+1}}.$$

□

(ii) Now we develop the formula for D_K. By recursive implementation of the definition of D_K in (78) and using definition of s in (25) we have

$$D_K = -\frac{1}{\beta\eta} \left[\frac{A_K}{s} \Delta_2 R_{(1,K)} + \frac{A_K}{s} \frac{A_{K-1}}{s} \Delta_2 R_{(1,K-1)} + \cdots + \frac{A_K}{s} \frac{A_{K-1}}{s} \cdots \frac{A_1}{s} \Delta_2 R_{(1,1)} \right].$$

Since $A_K = sd_{K-1}/d_K$, this simplifies to

$$D_K = -\frac{1}{\beta\eta} \left[\frac{d_{K-1}}{d_K} \Delta_2 R_{(1,K)} + \frac{d_{K-2}}{d_K} \Delta_2 R_{(1,K-1)} + \cdots + \frac{d_0}{d_K} \Delta_2 R_{(1,1)} \right] = -\frac{1}{\beta\eta d_K} \sum_{k=0}^{K-1} d_k \Delta_2 R_{(1,k+1)},$$

which, under constant $\Delta_2 R_{(1,i)} = R$ for all $i \geq 1$, is

$$D_K = -\frac{R}{\beta\eta d_K} \sum_{k=0}^{K-1} d_k.$$

Plugging the expression for d_k and simplifying the constant terms gives

$$\sum_{k=0}^{K-1} d_k = \frac{1}{1 - t + s} \left[1 - \frac{u_+^{K+1}(1 - u_-) - u_-^{K+1}(1 - u_+)}{u} \right].$$

The last two identities together with $(1-t+s)\beta\eta = 1-\beta$ imply

$$D_K = -\frac{R}{1-\beta}\frac{u - u_+^{K+1}(1-u_-) + u_-^{K+1}(1-u_+)}{u_+^{K+1} - u_-^{K+1}} = -\frac{R}{1-\beta}\left[\frac{u + (u_+^K - u_-^K)u_- u_+}{u_+^{K+1} - u_-^{K+1}} - 1\right].$$

□

(iii) The Lucas sequence d_K defined in part (i) satisfies $d_{K+1}/d_K = u_+ > 1$ as $K \to \infty$. Therefore, we obtain the limits

$$A = \frac{s}{u_+}, \qquad D = \frac{c}{1-\beta}(u_- - 1).$$

□

Finally we note that calculation of all marginal productivity indices using this method requires $\mathcal{O}(\log_2(I!))$ arithmetic operations, which is more than the linear number performed by algorithm FA in Figure 5.

4.3 Fast Algorithm under Convex Non-Decreasing Holding Costs in Admission Control Problem with Delay

Under convex non-decreasing holding costs, the immediate reward is $R_{(a,i)} = R_i^a : -c_i$ under any $a \in \mathcal{A}, i \in \mathcal{I}$. Therefore, we have

$$\Delta_1 R_{(1,i)} = 0, \qquad i \geq 0,$$
$$\Delta_2 R_{(0,i)} = \Delta_2 R_{(1,i)} = -c_i + c_{i-1} =: -\Delta c_i, \qquad i \geq 1,$$

and the fast algorithm simplifies to the one shown in Figure 6. This includes the special case of linear holding costs, when $\Delta c_i := c$ for all $i \in \mathcal{I}$.

The algorithm can also be used to derive the *greedy index*, which only looks one period ahead. Such an index is defined as $\nu_{(a,i)}^{\text{GREEDY}} := \lim_{\beta \to 0} \frac{\nu_{(a,i)}}{\beta}$ for all $(a,i) \in \widetilde{\mathcal{I}}$. In the case of linear holding costs, it is straightforward to obtain the greedy index as follows.

$$\begin{aligned}
\nu_{(1,0)}^{\text{GREEDY}} &= c(1-\mu);\\
\nu_{(0,0)}^{\text{GREEDY}} &= c(1-\eta-\mu^2\lambda);\\
\nu_{(1,1)}^{\text{GREEDY}} &= c(1-\mu^2);\\
\nu_{(0,1)}^{\text{GREEDY}} &= c(1-\frac{\mu\eta}{1-\zeta}),\ \text{if } I=2;\\
\nu_{(0,1)}^{\text{GREEDY}} &= c(1-\mu\eta),\ \text{if } I\geq 3;\\
\nu_{(a,i)}^{\text{GREEDY}} &= c,\ \text{for all } a\in\mathcal{A} \text{ and } 2\leq i\leq I;
\end{aligned}$$

4.4 Admission Control Problem with Delay to Server with an Infinite Buffer

In this subsection we assume linear holding costs.

{Input $I, \lambda, \mu, \beta, \{c_i\}_{i\in\mathcal{I}}$}
{Initialization}
$\zeta := \lambda(1-\mu)$; $\eta := \mu(1-\lambda)$; $\varepsilon := 1-\zeta-\eta$;
$A_0 := 0$; $A'_0 := \beta\zeta$; $B := \beta\mu/(1-\beta+\beta\mu)$; $B' := \beta\zeta B + \beta(\mu-\eta)$;
$C''_{0,0} := \zeta B\Delta c_1/\mu$; $D_0 := 0$; $D'_0 := 0$;
$\nu_{(1,0)} := C''_{0,0}/\lambda$;
{Loop}
for $K=1$ to $I-1$ do
  $A_K := \beta\zeta/[1-\beta+\beta\zeta+\beta\eta(1-A_{K-1})]$; $A'_K := \beta\zeta+\beta(\mu-\eta)A_K$; $Z_K := A_K A'_{K-1}/A'_K$;
  $C''_{K,K} := \zeta B\Delta c_{K+1}/\mu$; $C''_{K,K+1} := \zeta B^2\Delta c_{K+1}/\mu + \zeta B\Delta c_{K+2}/\mu$;
  $D_K := (\beta\eta D_{K-1}+\Delta c_K)\,A_K/(\beta\zeta)$; $D'_K := \beta(\mu-\eta)D_K$;

$$f^0 := -\frac{\frac{\beta\zeta}{A_K}D_K + (1-\beta)C''_{K,K+1} - (\beta D'_{K-1}-\Delta c_K)B' + \beta\mu(\beta\zeta D_{K-1}B + C''_{K,K+1})}{\frac{A'_K}{A_K} + \beta A'_{K-1}B' + \beta\zeta\beta\mu(1-BA_{K-1})};$$

$$f^1 := -\frac{\frac{\beta\zeta}{A_K}D_K + (1-\beta)D'_{K-1} - (\beta C''_{K,K+1}-\Delta c_K)A'_{K-1} + \beta\mu(\beta\zeta D_{K-1} + C''_{K,K+1}A_{K-1})}{\frac{A'_K}{A_K} + \beta A'_{K-1}B' + \beta\zeta\beta\mu(1-BA_{K-1})};$$

$$g^0 := \frac{\beta\lambda\,(1+B')}{\frac{A'_K}{A_K} + \beta A'_{K-1}B' + \beta\zeta\beta\mu(1-BA_{K-1})}; \quad g^1 := \frac{1+A'_{K-1}}{1+B'}g^0;$$

  if $K=1$ then

$$\nu_{(0,0)} := \frac{(1-\zeta)D'_0 + \zeta C''_{1,1} - (1-\zeta)A'_0 f^0 - \zeta B' f^1}{\lambda - (1-\zeta)A'_0 g^0 - \zeta B' g^1};$$

  else

$$\nu_{(0,K-1)} := \frac{\eta\left[D'_{K-2} + D_{K-1}A'_{K-2}\right] + \varepsilon D'_{K-1} + \zeta C''_{K,K} - [\eta Z_{K-1}+\varepsilon]\,A'_{K-1}f^0 - \zeta B' f^1}{\lambda - [\eta Z_{K-1}+\varepsilon]\,A'_{K-1}g^0 - \zeta B' g^1};$$

  end {if};

$$\nu_{(1,K)} := \frac{\mu D'_{K-1} + (1-\mu)C''_{K,K} - \mu A'_{K-1}f^0 - (1-\mu)B' f^1}{\lambda - \mu A'_{K-1}g^0 - (1-\mu)B' g^1};$$

end {for};
{Finalization}
$A_I := \beta\zeta/[1-\beta+\beta\zeta+\beta\eta(1-A_{I-1})]$; $A'_I := \beta\zeta+\beta(\mu-\eta)A_I$; $Z_I := A_I A'_{I-1}/A'_I$;
$D_I := (\beta\eta D_{I-1}+\Delta c_I)\,A_I/(\beta\zeta)$;

$$f^0 := -\frac{\frac{\beta\zeta}{A_I}D_I - \beta\mu D'_{I-1}}{\frac{A'_I}{A_I} + \beta\mu A'_{I-1}}; \quad g^0 := \frac{\lambda(1+\beta\mu)}{\frac{A'_I}{A_I} + \beta\mu A'_{I-1}};$$

$$\nu_{(0,I-1)} := \frac{\eta\left[D'_{I-2} + D_{I-1}A'_{I-2}\right] + \varepsilon D'_{I-1} - [\eta Z_{I-1}+\varepsilon]\,A'_{I-1}f^0}{(\eta+\varepsilon)\lambda - [\eta Z_{I-1}+\varepsilon]\,A'_{I-1}g^0};$$

$$\nu_{(*,I)} := \frac{D'_{I-1} + \frac{\beta\zeta}{A_I}D_I Z_I}{\lambda(1-Z_I)};$$

{Output $\left\{\nu_{(a,i)}\right\}_{(a,i)\in\widetilde{\mathcal{I}}}$}

Figure 6. Fast algorithm FA for calculation of an optimal index policy under convex non-decreasing holding costs.

Notice that the indices calculated in the algorithm FA's "Loop" are *independent* of the buffer length I (only the indices of states $(0, I-1)$ and $(*, I)$ in "Finalization" depend on I). In other words, considering two buffers with lengths $I_1 < I_2$, the marginal productivity indices of states $(1, 0), (0, 0), (1, 1), \ldots, (0, I_1-2), (1, I_1-1)$ are the same for both buffers; the indices of states $(0, I_1-1)$ and $(*, I_1)$ would differ, while the remaining states only exist under buffer I_2. Therefore, the algorithm FA can be used to obtain the indices for infinite-length buffer. However, in such a case, "Loop" would never stop.

We present a simple algorithmic check (Figure 7) that can be run before "Loop" (and after "Initialization") to verify whether $K_0 = K_1 = \infty$, i.e., whether it is optimal to let the gate open always. It is because the indices are calculated in nondecreasing order and they converge as the buffer length $I \to \infty$.

Lemma 2. *If the buffer length $I = \infty$, the marginal productivity indices calculated in "Loop" of algorithm FA under the discounted criterion in Figure 5 converge.*

Proof. We prove that A_K and D_K converge, and that their convergence implies the convergence of the marginal productivity indices. Lemma A5(ii) implies that A_K converges to a limit, say, $A \le \beta$ as $K \to \infty$. This limit must satisfy

$$A = \frac{\beta\zeta}{1 - \beta + \beta\zeta + \beta\eta(1 - A)}.$$

This equation has two solutions for A,

$$\frac{1 - \beta + \beta\zeta + \beta\eta \pm \sqrt{(1 - \beta + \beta\zeta + \beta\eta)^2 - 4\beta\eta\beta\zeta}}{2\beta\eta}.$$

However, it can be shown that $1 > \zeta - \beta\eta$ and $\beta < 1$ (which is true by the model parameter assumptions) implies

$$\frac{1 - \beta + \beta\zeta + \beta\eta - \sqrt{(1 - \beta + \beta\zeta + \beta\eta)^2 - 4\beta\eta\beta\zeta}}{2\beta\eta} < \beta < \frac{1 - \beta + \beta\zeta + \beta\eta + \sqrt{(1 - \beta + \beta\zeta + \beta\eta)^2 - 4\beta\eta\beta\zeta}}{2\beta\eta}.$$

By Lemma A5 it must be $A \le \beta$, therefore the limit is

$$A = \frac{1 - \beta + \beta\zeta + \beta\eta - \sqrt{(1 - \beta + \beta\zeta + \beta\eta)^2 - 4\beta\eta\beta\zeta}}{2\beta\eta}.$$

Similarly, it can be shown (see the next subsection) that D_K converges and the limit is therefore

$$D = \frac{cA}{\beta\zeta - \beta\eta A}.$$

As a consequence of the above, the remaining expressions, including those of the marginal productivity indices, converge. □

If the algorithmic check does not confirm the infinite thresholds, the algorithm FA can be run, stopping the loop once an index greater than ν is found and omitting "Finalization" part.

$A := \left[1-\beta+\beta\zeta+\beta\eta-\sqrt{(1-\beta+\beta\zeta+\beta\eta)^2-4\beta\eta\beta\zeta}\right]/(2\beta\eta)$; $A' := \beta\zeta+\beta(\mu-\eta)A$; $D := cA/(\beta\zeta-\beta\eta A)$;

$f^0 := -\dfrac{\frac{\beta\zeta}{A}D+\beta\zeta(c+\beta\mu BD)+\left[c-\beta(\mu-\eta)\beta D\right]B'}{\frac{A'}{A}+\beta A'B'+\beta\zeta\beta\mu(1-BA)}$;

$f^1 := -\dfrac{\frac{\beta\zeta}{A}D+c\beta\zeta BA+\left[\beta\mu\beta\zeta+(1-\beta)\beta(\mu-\eta)\right]D+(c-\beta\zeta\beta C)A'}{\frac{A'}{A}+\beta A'B'+\beta\zeta\beta\mu(1-BA)}$;

$g^0 := \dfrac{\beta\lambda\,(1+B')}{\frac{A'}{A}+\beta A'B'+\beta\zeta\beta\mu(1-BA)}$; $g^1 := \dfrac{1+A'}{1+B'}g^0$;

$\nu_{(1,\infty)} := \dfrac{\left[\beta(1-\mu)\beta\zeta C+\beta\mu\beta(\mu-\eta)D\right]-\beta\mu A'f^0-\beta(1-\mu)B'f^1}{\beta\lambda-\beta\mu A'g^0-\beta(1-\mu)B'g^1}$;

if $\nu \geq \nu_{(1,\infty)}$ then $K_0 := \infty$; $K_1 := \infty$; end {if};

Figure 7. Algorithmic check for the problem with infinite-length buffer.

4.5 Admission Control Problem with Delay to Server with no Dedicated Buffer

In this section we solve the admission control problem with delay for $I = 1$, i.e., no customer is allowed to be queued, except for the one in service. While this problem may not be of intrinsic interest, its solution given next will serve as a basis for the servers assignment problem with delay discussed in **??**. Considered is the linear holding cost case.

Proposition 4. *The marginal productivity index of state $(a, i) \in \widetilde{\mathcal{I}}$ in case $I = 1$ is state-independent and equals*

$$\nu_{(a,i)} := \frac{\zeta\beta C}{\lambda} = \frac{c\lambda\beta(1-\mu)}{1-\beta(1-\mu)}.$$

These indices can be obtained in the same way as the general case $I \geq 2$. The state-independent marginal productivity index means that, given a rejection cost ν, it is optimal either to admit always, or to reject always, regardless of the previous-epoch state and previous-epoch action, i.e., regardless of information available.

4.6 Admission Control Problem with Delay under Time-Average Criterion

Our results extend directly to the admission control with delay under the time-average criterion.

Proposition 5. *By setting $\beta := 1$, the algorithm FA in Figure 5 computes the marginal productivity indices for problem (14) under the time-average criterion.*

Proof. The algorithm FA is valid by setting $\beta := 1$ for the time-average criterion because the marginal productivity indices under that criterion are obtained in the limit $\beta \to 1$ of the marginal productivity indices under the discounted criterion, and the limits of all the expressions exist and are finite. □

In case $I = \infty$, the algorithmic check in Figure 7 is only valid under $\beta < 1$, and therefore is not suitable for the time-average criterion. In fact, it is not necessary to perform such a check, because under the time-average criterion the indices diverge.

4.7 Further Remarks

If in state $(1,0)$, the buffer is empty, because it was empty a period ago and the gate has been closed since then. Therefore, one could expect that the index of state $(1,0)$ is the same as the index of state 0 in the no-delay problem, which is in fact true. Moreover, there is a simple interpretation of that expression.

If the buffer is empty, the expected total β-discounted holding cost is

$$\zeta\beta c\left[1+\beta(1-\mu)+(\beta(1-\mu))^2+\dots\right]=\frac{\beta\zeta c}{1-\beta+\beta\mu},$$

because ζ is the probability that the customer remains in the buffer for more than a period. The above expression is equal to $\lambda\nu_{(1,0)}$, the expected (total β-discounted) rejection cost is if the rejection cost $\nu = \nu_{(1,0)}$. Thus, in state $(1,0)$ it is optimal to close the gate if the expected rejection cost is lower than the expected discounted total holding cost of an admitted customer. Further, in state $(1,0)$ it is optimal to let the gate open if the expected rejection cost is greater than the expected discounted total holding cost of an admitted customer. If the two expected costs are equal, both closing and opening is optimal. It is also clear that under the former condition it is optimal to close the gate in *any* state, and therefore the indices of all states must not be smaller than $\nu_{(1,0)}$.

Figure 8 shows the indices for a number of instances of the admission control problem with delay. An extensive simulation study we have performed suggests a convergence of the indices:

$$\begin{aligned}
\nu_{(1,i)} &\to \nu_{(0,i)} && \text{as } \lambda \to 0,\\
\nu_{(1,i)} &\to \nu_{(0,i-1)} && \text{as } \zeta \to 1,\\
\nu_{(0,i)} &\to \frac{\beta c}{1-\beta} && \text{as } i \to \infty,\\
\nu_{(1,i)} &\to \frac{\beta c}{1-\beta} && \text{as } i \to \infty.
\end{aligned}$$

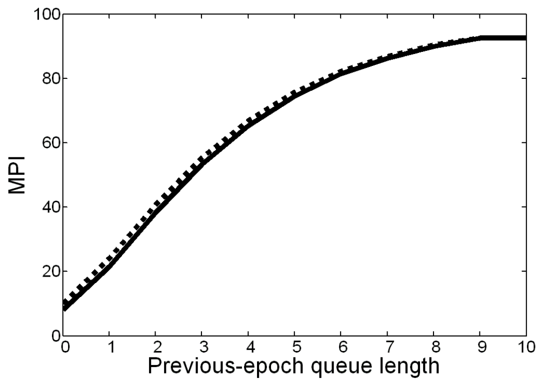
The convergence of the marginal productivity indices to $\beta c/(1-\beta)$ is intuitive. If the buffer is almost full (say, the pervious-epoch queue length is $I-2$), then admitting a customer means to increase the overal holding cost by c at least in the following $I-2$ periods, because the admitted customer cannot leave the system earlier that the previous $I-2$ customers. Therefore, the expected total β-discounted holding cost is at least

$$\beta c\left[1+\beta+\beta^2+\dots+\beta^{I-2}\right]=\frac{\beta c(1-\beta^{I-1})}{1-\beta}.$$

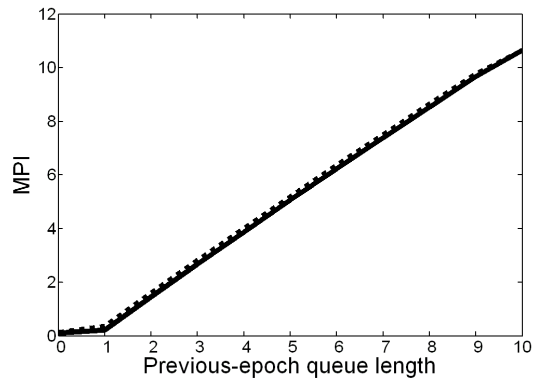
On the other hand, it is not greater than the expected holding cost of remaining in the buffer forever, which is

$$\beta c\left[1+\beta+\beta^2+\dots\right]=\frac{\beta c}{1-\beta}.$$

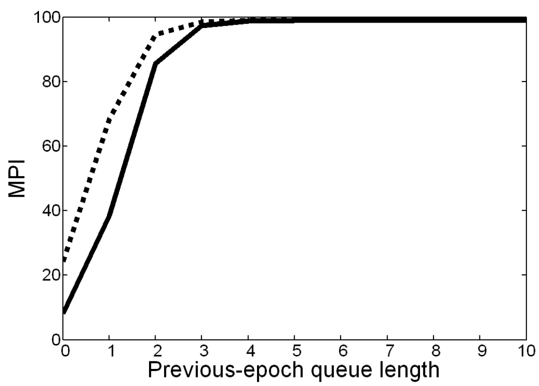
Now it is clear that the marginal productivity indices converge to $\beta c/(1-\beta)$ as $I\to\infty$.



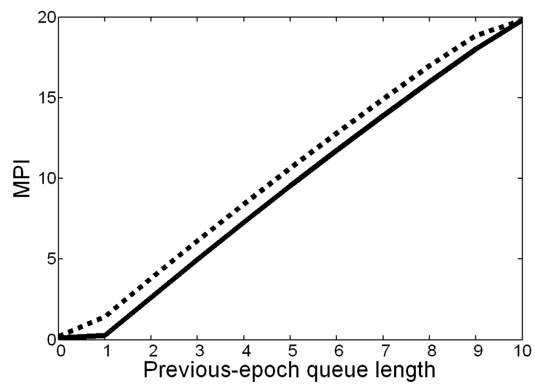
(a) if $\lambda = 0.1, \mu = 0.1$



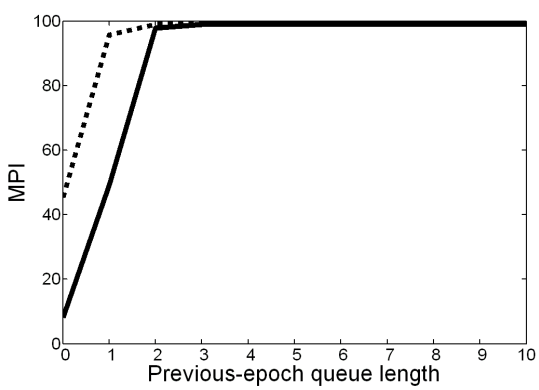
(b) if $\lambda = 0.1, \mu = 0.9$



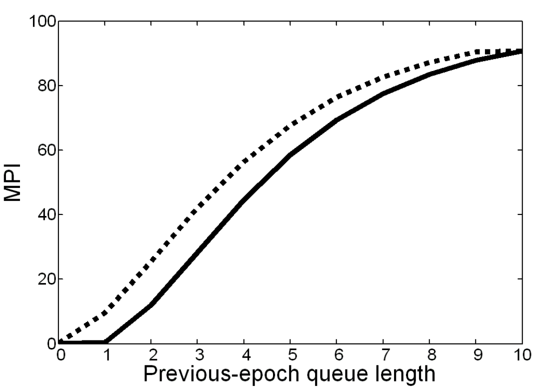
(c) if $\lambda = 0.5, \mu = 0.1$



(d) if $\lambda = 0.5, \mu = 0.9$



(e) if $\lambda = 0.9, \mu = 0.1$



(f) if $\lambda = 0.9, \mu = 0.9$

Figure 8. Optimal marginal productivity indices (MPI) for the admission control problem with delay with parameters $I = 10, c = 1, \beta = 0.99$. The solid line exhibits indices $\nu_{(1,i)}$ and the dotted line exhibits indices $\nu_{(0,i)}$.

5 Fast Algorithm for the Job Completions Problem with Delay

In the job completions problem with delay, we define the one-period rewards in the following way. If the queue length is $i \in \mathcal{I}$ and action $a \in \mathcal{A}$ is chosen, then the gatekeeper's *one-period reward* is defined as the expected number of job completions during the current period,

$$R_{(a,i)} = R_i^a := \begin{cases} 0, & \text{if } i = 0 \text{ and } a = 1, \\ \lambda\mu, & \text{if } i = 0 \text{ and } a = 0, \\ \mu, & \text{if } i \geq 1. \end{cases}$$

Therefore, we have

$$\Delta_1 R_{(1,i)} = \begin{cases} -\lambda\mu, & \text{if } i = 0, \\ 0, & \text{if } i \geq 1, \end{cases} \qquad \Delta_2 R_{(0,i)} = \begin{cases} \mu(1-\lambda), & \text{if } i = 1, \\ 0, & \text{if } i \geq 2, \end{cases} \qquad \Delta_2 R_{(1,i)} = \begin{cases} \mu, & \text{if } i = 1, \\ 0, & \text{if } i \geq 2. \end{cases}$$

The algorithm FA is presented in Figure 9, after substituting the following expressions: $C''_{0,0} := -\lambda B/\beta$; $C''_{K,K} := 0$; $C''_{K,K+1} := 0$.

However, we are only interested in the job completions problem under the long-run average criterion. Then, setting $\beta := 1$ in the fast algorithm, yields the constant index $\nu_{(a,i)} = -1$ for all $(a,i) \in \widetilde{\mathcal{I}}$. Figure 10 shows the simplified quantities that are obtained in this case in the fast algorithm FA. Since such an index is noninformative, we set out to obtain an alternative, second-order index in the following subsection.

5.1 Second-Order Marginal Productivity Index

Since the (first-order) index is noninformative, we proceed by introducing a second-order marginal productivity index $\gamma_{(a,i)}$, based on the Taylor series of $\nu_{(a,i)}$ at $\beta = 1$,

$$\nu_{(a,i)} = -1 + \gamma_{(a,i)}(1-\beta) + \mathcal{O}((1-\beta)^2), \qquad \text{as } \beta \to 1.$$

Thus, $\gamma_{(a,i)} := -\left.\dfrac{\partial \nu_{(a,i)}}{\partial \beta}\right|_{\beta=1}$. As the marginal productivity index policy prescribes to route an arriving customer to the queue of the lowest marginal productivity index, in the case of constant (first-order) indices the customer is to be routed to the queue of the lowest second-order MPI.

{Input I, λ, μ, β}
{Initialization}
$\zeta := \lambda(1-\mu)$; $\eta := \mu(1-\lambda)$; $\varepsilon := 1-\zeta-\eta$;
$A_0 := 0$; $A'_0 := \beta\zeta$; $B := \beta\mu/(1-\beta+\beta\mu)$; $B' := \beta\zeta B + \beta(\mu-\eta)$; $D_0 := 0$; $D'_0 := -\lambda\mu$;
$\nu_{(1,0)} := -B/\beta$;
{Loop}
for $K = 1$ to $I-1$ do
  $A_K := \beta\zeta/[1-\beta+\beta\zeta+\beta\eta(1-A_{K-1})]$; $A'_K := \beta\zeta+\beta(\mu-\eta)A_K$; $Z_K := A_K A'_{K-1}/A'_K$;
  if $K=1$ then $D_1 := -\eta A_1/(\beta\zeta)$; else $D_K := \beta\eta D_{K-1}A_K/(\beta\zeta)$; end {if};
  $D'_K := \beta(\mu-\eta)D_K$;
  $$g^0 := \frac{\beta\lambda(1+B')}{\frac{A'_K}{A_K}+\beta A'_{K-1}B'+\beta\zeta\beta\mu(1-BA_{K-1})}; \quad g^1 := \frac{1+A'_{K-1}}{1+B'}g^0;$$
  if $K=1$ then
    $$f^0 := -\frac{-\eta-(\beta D'_0+\mu)B'+\beta\mu\lambda\mu}{\frac{A'_1}{A_1}+\beta A'_0B'+\beta\zeta\beta\mu};$$
    $$f^1 := -\frac{-\eta+(1-\beta)D'_0-\mu A'_{K-1}}{\frac{A'_1}{A_1}+\beta A'_0B'+\beta\zeta\beta\mu};$$
    $$\nu_{(0,0)} := \frac{(1-\zeta)D'_0-(1-\zeta)A'_0f^0-\zeta B'f^1}{\lambda-(1-\zeta)A'_0g^0-\zeta B'g^1};$$
  else
    $$f^0 := -\frac{\frac{\beta\zeta}{A_K}D_K-\beta D'_{K-1}B'+\beta\mu\beta\zeta D_{K-1}B}{\frac{A'_K}{A_K}+\beta A'_{K-1}B'+\beta\zeta\beta\mu(1-BA_{K-1})};$$
    $$f^1 := -\frac{\frac{\beta\zeta}{A_K}D_K+(1-\beta)D'_{K-1}+\beta\mu\beta\zeta D_{K-1}}{\frac{A'_K}{A_K}+\beta A'_{K-1}B'+\beta\zeta\beta\mu(1-BA_{K-1})};$$
    $$\nu_{(0,K-1)} := \frac{\eta\left[D'_{K-2}+D_{K-1}A'_{K-2}\right]+\varepsilon D'_{K-1}-\left[\eta Z_{K-1}+\varepsilon\right]A'_{K-1}f^0-\zeta B'f^1}{\lambda-\left[\eta Z_{K-1}+\varepsilon\right]A'_{K-1}g^0-\zeta B'g^1};$$
  end {if};
  $$\nu_{(1,K)} := \frac{\mu D'_{K-1}-\mu A'_{K-1}f^0-(1-\mu)B'f^1}{\lambda-\mu A'_{K-1}g^0-(1-\mu)B'g^1};$$
end {for};
{Finalization}
$A_I := \beta\zeta/[1-\beta+\beta\zeta+\beta\eta(1-A_{I-1})]$; $A'_I := \beta\zeta+\beta(\mu-\eta)A_I$; $Z_I := A_IA'_{I-1}/A'_I$;
$D_I := \beta\eta D_{I-1}A_I/(\beta\zeta)$;
$$f^0 := -\frac{\frac{\beta\zeta}{A_I}D_I-\beta\mu D'_{I-1}}{\frac{A'_I}{A_I}+\beta\mu A'_{I-1}}; \quad g^0 := \frac{\lambda(1+\beta\mu)}{\frac{A'_I}{A_I}+\beta\mu A'_{I-1}};$$
$$\nu_{(0,I-1)} := \frac{\eta\left[D'_{I-2}+D_{I-1}A'_{I-2}\right]+\varepsilon D'_{I-1}-\left[\eta Z_{I-1}+\varepsilon\right]A'_{I-1}f^0}{(\eta+\varepsilon)\lambda-\left[\eta Z_{I-1}+\varepsilon\right]A'_{I-1}g^0};$$
$$\nu_{(*,I)} := \frac{D'_{I-1}+\frac{\beta\zeta}{A_I}D_IZ_I}{\lambda(1-Z_I)};$$
{Output $\left\{\nu_{(a,i)}\right\}_{(a,i)\in\widetilde{\mathcal{I}}}$}

Figure 9. Fast algorithm FA for calculation of an optimal index policy for the job completions problem.

{Input I, λ, μ}
{Initialization}
$\zeta := \lambda(1-\mu)$; $\eta := \mu(1-\lambda)$; $\varepsilon := 1-\zeta-\eta$;
$A_0 := 0$; $A'_0 := \zeta$; $B := 1$; $B' := \lambda$; $D_0 := 0$; $D'_0 := A'_0 - B'$; $\nu_{(1,0)} := -1$;
{Loop}
for $K = 1$ to $I-1$ do
  $A_K := \zeta/[\zeta + \eta(1-A_{K-1})]$; $A'_K := \zeta + \lambda\mu A_K$; $Z_K := A_K A'_{K-1}/A'_K$;
  $D_K := A_K - 1$; $D'_K := A'_K - B'$;
  if $K = 1$ then
    $f^0 := \dfrac{\mu(1-\lambda\mu-\lambda^2)}{\lambda(1+\lambda)+\mu(1-\lambda\mu-\lambda^2)}$; $f^1 := \dfrac{\mu(1-\lambda\mu)}{\lambda(1+\lambda)+\mu(1-\lambda\mu-\lambda^2)}$;
  else
    $f^0 := \dfrac{-\mu(1-\lambda\mu-\lambda^2)D_{K-1}}{\lambda(1+\lambda)-\mu(1-\lambda\mu-\lambda^2)D_{K-1}}$; $f^1 := \dfrac{-\mu(1-\lambda\mu)D_{K-1}}{\lambda(1+\lambda)-\mu(1-\lambda\mu-\lambda^2)D_{K-1}}$;
  end {if};
  $g^0 := 1-f^0$; $g^1 := 1-f^1$; $\nu_{(0,K-1)} := -1$; $\nu_{(1,K)} := -1$;
end {for};
{Finalization}
$A_I := \zeta/[\zeta + \eta(1-A_{I-1})]$; $A'_I := \zeta + \lambda\mu A_I$; $Z_I := A_I A'_{I-1}/A'_I$; $D_I := A_I - 1$;
$f^0 := \dfrac{-\mu(1-\lambda\mu-\lambda)D_{I-1}}{\lambda(1+\mu)-\mu(1-\lambda\mu-\lambda)D_{I-1}}$; $g^0 := 1-f^0$; $\nu_{(0,I-1)} := -1$; $\nu_{(*,I)} := -1$;
{Output $\{\nu_{(a,i)}\}_{(a,i)\in\widetilde{\mathcal{I}}}$}

Figure 10. Fast algorithm FA for calculation of an optimal index policy for the job completions problem under the time-average criterion.

We conjecture that the second-order indices are as follows ($K \geq 1$):

$$
\begin{aligned}
\gamma_{(1,0)} &= \frac{1}{\mu} - 1, \\
\gamma_{(0,0)} &= \frac{2}{\mu} - 1 - \frac{\mu - (\mu+\lambda)\zeta}{\mu^2 + \mu\eta\zeta}, \\
\gamma_{(1,1)} &= -\left(2 + \frac{\lambda}{\mu}\right) + \frac{1}{\mu}\left(2 + \frac{\lambda}{\mu}\right), \\
\gamma_{(0,1)} &= -\left(2 + \frac{\lambda}{\mu}\right) + \frac{1}{2\mu}\left(5 + 3\frac{\lambda}{\mu}\right) + \frac{2\lambda - 1}{2\mu(2\zeta+1)}\left(1 + \frac{\lambda}{\mu} + \frac{\lambda^2}{\mu\eta}\right) + \frac{\lambda^2}{2\mu^2\eta(2\zeta+1)}, \\
\gamma_{(1,K+1)} &= \gamma_{(1,K)} + \frac{1}{\mu}\left\{1 + \frac{\lambda}{\mu} + \frac{\lambda^2}{\mu\eta}\left[1 + \frac{\zeta}{\eta} + \cdots + \left(\frac{\zeta}{\eta}\right)^{K-1}\right]\right\}, \\
\gamma_{(0,K+1)} &= \gamma_{(0,K)} + \frac{1}{\mu}\left\{1 + \frac{\lambda}{\mu} + \frac{\lambda^2}{\mu\eta}\left[1 + \frac{\zeta}{\eta} + \cdots + \left(\frac{\zeta}{\eta}\right)^{K-1}\right]\right\} + \frac{\lambda^2(\eta+\lambda)}{\mu^2\eta(2\zeta+1)}\left(\frac{\zeta}{\eta}\right)^K, \\
\gamma_{(*,I)} &= \frac{I}{\mu} - 1 + \frac{\lambda}{\mu\eta}\left[(I-1) + (I-2)\frac{\zeta}{\eta} + \cdots + \left(\frac{\zeta}{\eta}\right)^{I-2}\right].
\end{aligned}
$$

The last expressions can be simplified. If $\lambda \neq \mu$, then

$$
\begin{aligned}
\gamma_{(1,K+1)} &= \gamma_{(1,K)} + \frac{1}{\mu-\lambda} - \frac{1}{\mu-\lambda}\frac{\lambda^2}{\mu^2}\left(\frac{\zeta}{\eta}\right)^K, \\
\gamma_{(0,K+1)} &= \gamma_{(0,K)} + \frac{1}{\mu-\lambda} + \left[\frac{\eta+\lambda}{\eta(2\zeta+1)} - \frac{1}{\mu-\lambda}\right]\frac{\lambda^2}{\mu^2}\left(\frac{\zeta}{\eta}\right)^K, \\
\gamma_{(*,I)} &= \frac{I}{\mu} - 1 + \frac{\lambda}{\mu}\left[\frac{\zeta\left(\frac{\zeta}{\eta}\right)^{I-1} - \eta + I(\mu-\lambda)}{(\mu-\lambda)^2}\right].
\end{aligned}
$$

If $\lambda = \mu$, then

$$
\begin{aligned}
\gamma_{(1,K+1)} &= \gamma_{(1,K)} + \frac{2}{\mu} + \frac{K}{\eta}, \\
\gamma_{(0,K+1)} &= \gamma_{(0,K)} + \frac{2}{\mu} + \frac{K}{\eta} + \frac{\eta+\lambda}{\eta(2\zeta+1)}, \\
\gamma_{(*,I)} &= \frac{I}{\mu} - 1 + \frac{I(I-1)}{2\eta}.
\end{aligned}
$$

6 Conclusions

We have presented a restless bandit approach which yielded an efficient exact algorithm for the calculation of the marginal productivity indices and optimal threshold queue lengths for an admission control problem with an action or information delay of one period. The algorithm draws on and significantly reduces the complexity of the adaptive-greedy algorithm for the calculation of restless bandit optimal index policy.

We propose such indices as building blocks in a heuristic for a harder problem of admission control and/or routing to parallel queues, where the queues can be heterogeneous in buffer lengths, departure probabilities, holding costs, discount factors, and delays. The evaluation of the heuristic is a part of the work in progress.

Our approach seems to be tractable also for the admission control problem with larger delays and, more generally, for arbitrary restless bandits with delays.

References

Altman, E. and Nain, P. (1992). Closed-loop control with delayed information. *Performance Evaluation Review*, 20(1):193–204.

Artiges, D. (1995). Optimal routing into two heterogeneous service stations with delayed information. *IEEE Transactions on Automatic Control*, 40(7):1234–1236.

Brooks, D. M. and Leondes, C. T. (1972). Markov decision processes with state-information lag. *Operations Research*, 20(4):904–907.

Hajek, B. (1984). Optimal control of two interacting service stations. *IEEE Transactions on Automatic Control*, 29:491–499.

Hordijk, A. and Koole, G. (1990). On the optimality of the generalized shortest queue policy. *Probability in the Engineering and Informational Sciences*, 4:477–487.

Houck, D. J. (1987). Comparison of policies for routing customers to parallel queueing sytems. *Operations Research*, 35:306–310.

Kalman, D. and Mena, R. (2003). The Fibonacci numbers—exposed. *Mathematics Magazine*, 76(3):167–181.

Kuri, J. and Kumar, A. (1995). Optimal control of arrivals to queues with delayed queue length information. *IEEE Transactions on Automatic Control*, 40(8):1444–1450.

Lucas, E. (1878). Théorie des fonctions numériques simplement périodiques. *American Journal of Mathematics*, 1:184–240, 289–321.

Niño-Mora, J. (2001). Restless bandits, partial conservation laws and indexability. *Advances in Applied Probability*, 33(1):76–98.

Niño-Mora, J. (2002). Dynamic allocation indices for restless projects and queueing admission control: A polyhedral approach. *Mathematical Programming, Series A*, 93(3):361–413.

Niño-Mora, J. (2006a). Marginal productivity index policies for scheduling a multiclass delay-/loss-sensitive queue. *Queueing Systems*, 54:281–312.

Niño-Mora, J. (2006b). Restless bandit marginal productivity indices, diminishing returns, and optimal control of make-to-order/make-to-stock $M/G/1$ queues. *Mathematics of Operations Research*, 31(1):50–84.

Niño-Mora, J. (2007a). Characterization and computation of restless bandit marginal productivity indices. In *Proceedings of the 2nd International Conference on Performance Evaluation Methodologies and Tools*. ICST, Brussels, Belgium.

Niño-Mora, J. (2007b). Dynamic priority allocation via restless bandit marginal productivity indices. *TOP*, 15(2):161–198.

Niño-Mora, J. (2007c). Marginal productivity index policies for admission control and routing to parallel multi-server loss queues with reneging. *Lecture Notes in Computer Science*, 4465:138–149.

Niño-Mora, J. (2007d). Marginal productivity index policies for scheduling multiclass delay/loss-sensitive traffic with delayed state observation. In *Proceedings of the 3rd EuroNGI Conference on Next Generation Internet Networks - Design and Engineering for Heterogeneity*, pages 209–217.

Whittle, P. (1988). Restless bandits: Activity allocation in a changing world. *A Celebration of Applied Probability, J. Gani (Ed.), Journal of Applied Probability*, 25A:287–298.

Winston, W. (1977). Optimality of the shortest line discipline. *Journal of Applied Probability*, 14:181–189.

A Admission Control Problem with Delay: Analysis of Marginal Works

In this section we set out to obtain closed formulae for marginal works in the admission control problem with delay and present results in terms of action-differences in total work that will facilitate the proof of their positivity.

A.1 Preliminaries

Recall the definition of the expected total β-discounted work (briefly, *total work*) in (11). The following is the work balance equation for a fixed active set $\mathcal{S}$:

$$g^{\mathcal{S}}_{(a,i)} = \begin{cases} W_{(a,i)} + \beta \sum_{j \in \mathcal{I}} p^a_{ij} g^{\mathcal{S}}_{(1,j)} & \text{if } (a,i) \in \mathcal{S} \\ W_{(a,i)} + \beta \sum_{j \in \mathcal{I}} p^a_{ij} g^{\mathcal{S}}_{(0,j)} & \text{if } (a,i) \notin \mathcal{S} \end{cases} \tag{26}$$

The following two implications of the work balance equation were found useful in the problem analysis. First, we give a characterization of total works $g^{\mathcal{S}}_{(a,i)}$'s in terms of their action-differences $\Delta_1 g^{\mathcal{S}}_{(1,i)}$'s and state-differences $\Delta_2 g^{\mathcal{S}}_{(a,i)}$'s.

Lemma A1. *For a fixed active set $\mathcal{S}$,*

$$(1-\beta) g^{\mathcal{S}}_{(1,i)} = \lambda - \beta\mu\Delta_2 g^{\mathcal{S}}_{(1,i)} \qquad \text{if } (1,i) \in \mathcal{S} \tag{27}$$

$$(1-\beta) g^{\mathcal{S}}_{(1,i)} = \lambda - \beta\mu\Delta_2 g^{\mathcal{S}}_{(0,i)} - \beta\Delta_1 g^{\mathcal{S}}_{(1,i)} \qquad \text{if } (1,i) \notin \mathcal{S} \tag{28}$$

$$(1-\beta) g^{\mathcal{S}}_{(0,i)} = \begin{cases} \lambda - \beta\mu\Delta_2 g^{\mathcal{S}}_{(1,I)} + \beta\Delta_1 g^{\mathcal{S}}_{(1,I)} & \text{if } i = I \\ \beta\zeta\Delta_2 g^{\mathcal{S}}_{(1,i+1)} - \beta\eta\Delta_2 g^{\mathcal{S}}_{(1,i)} + \beta\Delta_1 g^{\mathcal{S}}_{(1,i)} & \text{otherwise} \end{cases} \qquad \text{if } (0,i) \in \mathcal{S} \tag{29}$$

$$(1-\beta) g^{\mathcal{S}}_{(0,i)} = \begin{cases} \lambda - \beta\mu\Delta_2 g^{\mathcal{S}}_{(0,I)} & \text{if } i = I \\ \beta\zeta\Delta_2 g^{\mathcal{S}}_{(0,i+1)} - \beta\eta\Delta_2 g^{\mathcal{S}}_{(0,i)} & \text{otherwise} \end{cases} \qquad \text{if } (0,i) \notin \mathcal{S} \tag{30}$$

Proof. Suppose first that $(1,i) \in \mathcal{S}$. By adding $-\beta g^{\mathcal{S}}_{(1,i)}$ at both sides of identity (26) for $(1,i)$, we obtain

$$(1-\beta) g^{\mathcal{S}}_{(1,i)} = W_{(1,i)} - \beta \sum_{j \in \mathcal{I}} p^1_{ij} \left(g^{\mathcal{S}}_{(1,i)} - g^{\mathcal{S}}_{(1,j)} \right),$$

which simplifies to (27) after plugging the definition of $W_{(1,i)}$ in (10) and that of p^1_{ij} in (5)-(6), and finally using (23).

The remaining identities are obtained analogously by adding $-\beta g^{\mathcal{S}}_{(a,i)}$ at both sides of identity (26), then plugging the definition of $W_{(a,i)}$ in (10) and that of p^a_{ij} in (5)-(6), and finally using (22)-(23). □

The following lemma characterizes action-differences $\Delta_1 g^{\mathcal{S}}_{(1,i)}$'s in terms of state-differences $\Delta_2 g^{\mathcal{S}}_{(a,i)}$'s.

Lemma A2. *For a fixed active set $\mathcal{S}$ and any state $0 \le i \le I-1$,*

$$\Delta_1 g^{\mathcal{S}}_{(1,i)} = \lambda - \beta\zeta\Delta_2 g^{\mathcal{S}}_{(1,i+1)} - \beta(\mu-\eta)\Delta_2 g^{\mathcal{S}}_{(1,i)} \qquad \text{if } (0,i),(1,i) \in \mathcal{S} \tag{31}$$

$$\Delta_1 g^{\mathcal{S}}_{(1,i)} = \lambda - \beta\zeta\Delta_2 g^{\mathcal{S}}_{(0,i+1)} - \beta(\mu-\eta)\Delta_2 g^{\mathcal{S}}_{(0,i)} \qquad \text{if } (0,i),(1,i) \notin \mathcal{S} \tag{32}$$

$$(1+\beta)\Delta_1 g^{\mathcal{S}}_{(1,i)} = \lambda - \beta\zeta\Delta_2 g^{\mathcal{S}}_{(1,i+1)} - \beta\mu\Delta_2 g^{\mathcal{S}}_{(0,i)} + \beta\eta\Delta_2 g^{\mathcal{S}}_{(1,i)} \qquad \text{if } (0,i) \in \mathcal{S}, (1,i) \notin \mathcal{S} \tag{33}$$

$$(1-\beta)\Delta_1 g^{\mathcal{S}}_{(1,i)} = \lambda - \beta\zeta\Delta_2 g^{\mathcal{S}}_{(0,i+1)} - \beta\mu\Delta_2 g^{\mathcal{S}}_{(1,i)} + \beta\eta\Delta_2 g^{\mathcal{S}}_{(0,i)} \qquad \text{if } (0,i) \notin \mathcal{S}, (1,i) \in \mathcal{S} \tag{34}$$

$$\Delta_1 g^{\mathcal{S}}_{(1,I)} = 0. \tag{35}$$

Proof. If $(0,i),(1,i) \in \mathcal{S}$, identity (31) is obtained by subtracting (29) from (27), and using (22). The remaining identities are obtained analogously. □

Recall the definition of marginal works $w^{\mathcal{S}}_{(a,i)}$ in (16). In order to obtain a characterization of marginal works in terms of $\Delta_1 g_{(1,i)}$'s, we specialize the work balance equation in (26) for policies $\langle 0,\mathcal{S}\rangle$ and $\langle 1,\mathcal{S}\rangle$.

Lemma A3. *For a fixed active set $\mathcal{S}$,*

$$g^{\langle 1,\mathcal{S}\rangle}_{(a,i)} = W_{(a,i)} + \beta\sum_{j\in\mathcal{I}} p^a_{ij} g^{\mathcal{S}}_{(1,j)}$$

$$g^{\langle 0,\mathcal{S}\rangle}_{(a,i)} = W_{(a,i)} + \beta\sum_{j\in\mathcal{I}} p^a_{ij} g^{\mathcal{S}}_{(0,j)}$$

Now we are ready to express marginal works $w^{\mathcal{S}}_{(a,i)}$ in terms of action-differences $\Delta_1 g^{\mathcal{S}}_{(1,i)}$'s.

Lemma A4. *For a fixed active set $\mathcal{S}$,*

$$w^{\mathcal{S}}_{(1,i)} = \begin{cases} \beta\Delta_1 g^{\mathcal{S}}_{(1,0)} & \text{if } i=0 \\ \beta\mu\Delta_1 g^{\mathcal{S}}_{(1,i-1)} + \beta(1-\mu)\Delta_1 g^{\mathcal{S}}_{(1,i)} & \text{otherwise} \end{cases} \tag{36}$$

$$w^{\mathcal{S}}_{(0,i)} = \begin{cases} \beta(1-\zeta)\Delta_1 g^{\mathcal{S}}_{(1,0)} + \beta\zeta\Delta_1 g^{\mathcal{S}}_{(1,1)} & \text{if } i=0 \\ \beta\eta\Delta_1 g^{\mathcal{S}}_{(1,i-1)} + \beta\varepsilon\Delta_1 g^{\mathcal{S}}_{(1,i)} + \beta\zeta\Delta_1 g^{\mathcal{S}}_{(1,i+1)} & \text{otherwise} \\ w^{\mathcal{S}}_{(1,I)} & \text{if } i=I \end{cases} \tag{37}$$

Proof. From plugging the identities in Lemma A3 into the definition of $w^{\mathcal{S}}_{(a,i)}$, we obtain

$$w^{\mathcal{S}}_{(a,i)} = \beta\sum_{j\in\mathcal{I}} p^a_{ij}\Delta_1 g^{\mathcal{S}}_{(1,j)}.$$

Then, using the definition of p^a_{ij} in (5)-(6) gives the result. □

The last lemma shows that marginal work is equal to the *expected next-period β-discounted* increment in total work if starting from state $(1, i)$ instead of $(0, i)$, i.e.,

$$w^{\mathcal{S}}_{(a,i)} = \mathbb{E}^a_i \left[\Delta_1 g^{\mathcal{S}}\right], \tag{38}$$

where the random variable $\Delta_1 g^{\mathcal{S}}$ has value $\Delta_1 g^{\mathcal{S}}_{(1,j)}$ with probability p^a_{ij}. This suggests a way for establishing positivity of marginal works needed in Definition 1(i) by establishing positivity of $\Delta_1 g^{\mathcal{S}}_{(1,i)}$'s for all states $0 \le i \le I-1$ (recall that by (35), $\Delta_1 g^{\mathcal{S}}_{(1,i)} = 0$ for state I).

Before calculating the action-differences needed above, we prepare notation and state an auxiliary result. These quantities will appear in the action-differences recursion developed in the following subsection. For $j \ge 0$ we define

$$A_0 := 0, \qquad A_{j+1} := \frac{\beta\zeta}{1-\beta+\beta\zeta+\beta\eta(1-A_j)}, \qquad B := \frac{\beta\mu}{1-\beta+\beta\mu}. \tag{39}$$

and

$$A'_j := \beta\zeta + \beta(\mu-\eta)A_j, \qquad Z_{j+1} := A_{j+1}\frac{A'_j}{A'_{j+1}}, \qquad B' := \beta\zeta B + \beta(\mu-\eta). \tag{40}$$

Lemma A5.

(i) $0 < B < \beta$;

(ii) $0 < A_{j+1} < \beta$ *and* $A_j \le A_{j+1}$ *for all* $j \ge 0$;

(iii) $0 < Z_{j+1} < \beta$ *for all* $j \ge 0$.

Proof. (i) The positivity is straightforward from the definition of B in (39), because $\beta\mu > 0$ and $1-\beta > 0$ by the model parameter assumptions given in (7). On the other hand, the same assumptions imply $B < \beta$. □

(ii) We proceed by induction. Since $\beta\eta(1-A_0) \ge 0$, we have $A_1 \le \frac{\beta\zeta}{1-\beta+\beta\zeta} < \beta$, where the last inequality is due to the model parameter assumptions. Hence, assuming inductively $\beta\eta(1-A_j) \ge 0$, we have $A_{j+1} \le \frac{\beta\zeta}{1-\beta+\beta\zeta} < \beta$. The positivity of A_{j+1} follows from $\beta\zeta > 0$, $1-\beta > 0$ and $\beta\eta(1-A_j) \ge 0$.
Similarly by induction we prove the monotonicity. As the first step, the above implies $0 = A_0 < A_1$. Hence we have $\beta\eta(1-A_0) \ge \beta\eta(1-A_1)$, which implies $A_1 \le A_2$. Inductively, assuming $A_{j-1} \le A_j$ analogously implies $A_j \le A_{j+1}$. □

(iii) The model parameter assumptions given in (7) imply that $A'_j > 0$, since $\mu > \eta$ for all $j \ge 0$. Therefore, the definition of Z_{j+1} in (40) implies $Z_{j+1} > 0$. On the other hand, using (40) we can write

$$Z_{j+1} = \frac{\beta\zeta + \beta(\mu-\eta)A_j}{\frac{\beta\zeta}{A_{j+1}} + \beta(\mu-\eta)} < \frac{\beta\zeta + \beta(\mu-\eta)\beta}{\frac{\beta\zeta}{\beta} + \beta(\mu-\eta)} = \beta,$$

where the inequality is due to $A_j, A_{j+1} < \beta$ by (ii). □

A.2 Calculation of Action-Differences in Total Work

Since Lemma A4 characterizes marginal works $w^{\mathcal{S}}_{(a,i)}$'s as weighted averages of action-differences $\Delta_1 g^{\mathcal{S}}_{(1,i)}$'s, in this subsection we focus on the calculation of the latter. The ultimate goal of proving positivity of marginal works under $\mathcal{F}$ in order to establish condition Definition 1(i), will be reached by proving positivity of action-differences in subsection A.3 and subsection A.4 for active sets $\widetilde{\mathcal{I}}_{K,K}$ and $\widetilde{\mathcal{I}}_{K,K+1}$, respectively.

In the following we show that all the relevant state-differences can be obtained by recursion from two *pivot state-differences* associated to the two thresholds K_0, K_1 under any policy $\widetilde{\mathcal{I}}_{K_0,K_1}$. Denote by $K_0' := \min\{K_0, I\}$ and note that the relevant state-differences needed in the subsequent analysis are $\Delta_2 g^{\mathcal{S}}_{(0,i)}$'s for $1 \le i \le K_0' - 1$, and $\Delta_2 g^{\mathcal{S}}_{(1,i)}$ for $K_1 + 1 \le i \le I$.

Lemma A6. *For a fixed active set $\mathcal{S} = \widetilde{\mathcal{I}}_{K_0,K_1}$,*

$$\Delta_2 g^{\mathcal{S}}_{(0,i)} = \Delta_2 g^{\mathcal{S}}_{(0,K_0')} \prod_{j=i}^{K_0'-1} A_j, \qquad \text{for } 1 \le i \le K_0' - 1, \tag{41}$$

$$\Delta_2 g^{\mathcal{S}}_{(1,i)} = \Delta_2 g^{\mathcal{S}}_{(1,K_1)} B^{i-K_1}, \qquad \text{for } K_1 + 1 \le i \le I. \tag{42}$$

Proof. For $1 \le i \le K_0' - 1$, augmented states $(0,i), (0,i-1) \notin \mathcal{S}$, so taking the difference of (30) for i and for $i-1$ gives

$$(1 - \beta + \beta\eta + \beta\zeta)\Delta_2 g^{\mathcal{S}}_{(0,i)} = \beta\eta\Delta_2 g^{\mathcal{S}}_{(0,i-1)} + \beta\zeta\Delta_2 g^{\mathcal{S}}_{(0,i+1)}.$$

Expressed for $i = 1$ and divided by $1 - \beta + \beta\eta + \beta\zeta$, we have $\Delta_2 g^{\mathcal{S}}_{(0,1)} = \Delta_2 g^{\mathcal{S}}_{(0,2)} A_1$, since, by definition,

$$\Delta_2 g^{\mathcal{S}}_{(0,0)} = 0 \qquad \text{and} \qquad A_1 = \frac{\beta\zeta}{1 - \beta + \beta\zeta + \beta\eta}.$$

Inductively, if $\Delta_2 g^{\mathcal{S}}_{(0,i-1)} = \Delta_2 g^{\mathcal{S}}_{(0,i)} A_{i-1}$, then

$$(1 - \beta + \beta\eta + \beta\zeta(1 - A_{i-1}))\Delta_2 g^{\mathcal{S}}_{(0,i)} = \beta\zeta\Delta_2 g^{\mathcal{S}}_{(0,i+1)}.$$

which is the same as $\Delta_2 g^{\mathcal{S}}_{(0,i)} = \Delta_2 g^{\mathcal{S}}_{(0,i+1)} A_i$ for all $1 \le i \le K_0' - 1$. This recursion gives (41).

Similarly for $K_1 + 1 \le i \le I$, augmented states $(1,i), (1,i-1) \in \mathcal{S}$, so taking the difference of (27) for i and for $i-1$ gives $\Delta_2 g^{\mathcal{S}}_{(1,i)} = \Delta_2 g^{\mathcal{S}}_{(1,i-1)} B$. This recursion gives (42). □

Further we identify a recursion to calculate action-differences $\Delta_1 g^{\mathcal{S}}_{(1,i)}$'s in terms of the two pivot state-differences $\Delta_2 g^{\mathcal{S}}_{(0,K_0')}$ and $\Delta_2 g^{\mathcal{S}}_{(1,K_1)}$. Thus, this is a simplification of Lemma A2.

Proposition A1. *For a fixed active set $\mathcal{S} = \widetilde{\mathcal{I}}_{K_0,K_1}$,*

$$\Delta_1 g^{\mathcal{S}}_{(1,K_0'-1)} = \lambda - A'_{K_0'-1}\Delta_2 g^{\mathcal{S}}_{(0,K_0')}, \qquad \text{if } 1 \le K_0', \tag{43}$$

$$\Delta_1 g^{\mathcal{S}}_{(1,i)} = \lambda(1 - Z_{i+1}) + \Delta_1 g^{\mathcal{S}}_{(1,i+1)} Z_{i+1}, \qquad \text{for } 0 \le i \le K_0' - 2, \tag{44}$$

$$\Delta_1 g^{\mathcal{S}}_{(1,K_1)} = \lambda - B'\Delta_2 g^{\mathcal{S}}_{(1,K_1)}, \qquad \text{if } K_1 \le I - 1, \tag{45}$$

$$\Delta_1 g^{\mathcal{S}}_{(1,i)} = \lambda(1 - B) + \Delta_1 g^{\mathcal{S}}_{(1,i-1)} B, \qquad \text{for } K_1 + 1 \le i \le I - 1. \tag{46}$$

Proof. As a consequence of plugging (41) into (32) we have

$$\Delta_1 g^{\mathcal{S}}_{(1,i)} = \lambda - A'_i \Delta_2 g^{\mathcal{S}}_{(0,K'_0)} \prod_{j=i+1}^{K'_0-1} A_j, \qquad \text{for } 0 \le i \le K'_0 - 1,$$

This identity expressed for $i = K'_0 - 1$ gives (43), and expressed for i and $i+1$ implies (44).

Similarly, by plugging (42) into (31) we have

$$\Delta_1 g^{\mathcal{S}}_{(1,i)} = \lambda - B' \Delta_2 g^{\mathcal{S}}_{(1,K_1)} B^{i-K_1}, \qquad \text{for } K_1 \le i \le I - 1.$$

This identity expressed for $i = K_1$ gives (45), and expressed for i and $i-1$ implies (46). □

The above results help significantly simplify the subsequent analysis, which we present in separate subsections for optimal active sets $\widetilde{\mathcal{I}}_{K,K}$ and $\widetilde{\mathcal{I}}_{K,K+1}$. In each subsection we first present expressions for the pivot state-differences in a lemma and then establish positivity of $\Delta_1 g^{\mathcal{S}}_{(1,i)} > 0$ for all $0 \le i \le I-1$.

A.3 Positivity of Action-Differences in Total Work under Active Set $\widetilde{\mathcal{I}}_{K,K}$

Lemma A7. *Under active set $\mathcal{S} = \widetilde{\mathcal{I}}_{K,K}$,*

(i) if $K = 0$, then

$$\Delta_2 g^{\mathcal{S}}_{(1,0)} = 0. \tag{47}$$

(ii) if $1 \le K \le I-1$, then

$$\Delta_2 g^{\mathcal{S}}_{(0,K)} = \frac{\beta\lambda\,(1+B')}{\frac{A'_K}{A_K} + \beta A'_{K-1} B' + \beta\zeta\beta\mu(1 - B A_{K-1})}, \tag{48}$$

$$\Delta_2 g^{\mathcal{S}}_{(1,K)} = \frac{1 + A'_{K-1}}{1+B'} \Delta_2 g^{\mathcal{S}}_{(0,K)}. \tag{49}$$

(iii) if $K = I$, then

$$\Delta_2 g^{\mathcal{S}}_{(0,I)} = \frac{\lambda(1+\beta\mu)}{\frac{A'_I}{A_I} + \beta\mu A'_{I-1}}. \tag{50}$$

(iv) if $K = I+1$, then

$$\Delta_2 g^{\mathcal{S}}_{(0,I)} = \frac{\lambda A_I}{A'_I}. \tag{51}$$

Proof. (i) The identity is by definition. □

(ii) Taking the difference of (27) for $i = K$ and (28) for $i = K-1$, and plugging (32) for $i = K-1$, gives

$$(1 - \beta + \beta\mu)\Delta_2 g^{\mathcal{S}}_{(1,K)} = \beta\lambda + \beta\left[\mu - \beta(\mu - \eta)\right]\Delta_2 g^{\mathcal{S}}_{(0,K-1)} - \beta^2\zeta\Delta_2 g^{\mathcal{S}}_{(0,K)}. \tag{52}$$

Similarly, taking the difference of (29) for $i = K$ and (30) for $i = K-1$, and plugging (31) for $i = K$, gives

$$(1-\beta+\beta\zeta)\Delta_2 g^{\mathcal{S}}_{(0,K)} = \beta\lambda + (1-\beta)\beta\zeta\Delta_2 g^{\mathcal{S}}_{(1,K+1)} - \beta\left[\eta+\beta(\mu-\eta)\right]\Delta_2 g^{\mathcal{S}}_{(1,K)} + \beta\eta\Delta_2 g^{\mathcal{S}}_{(0,K-1)}. \tag{53}$$

Using (41) for $i = K-1$ and (42) for $i = K+1$, and solving the above system of two equations yields the results. □

(iii) The proof goes along the same lines as in the previous case, yet the latter identity becomes

$$(1-\beta+\beta\zeta)\Delta_2 g^{\mathcal{S}}_{(0,I)} = \lambda - \beta\mu\Delta_2 g^{\mathcal{S}}_{(1,I)} + \beta\eta\Delta_2 g^{\mathcal{S}}_{(0,I-1)}. \tag{54}$$

□

(iv) Taking the difference of (30) for $i = I$ and for $i = I-1$, and plugging (41) for $i = I-1$, yields the result. □

Proposition A2. *Under active set* $\mathcal{S} = \widetilde{\mathcal{I}}_{K,K}$ *with* $0 \le K \le I+1$, *action-differences* $\Delta_1 g^{\mathcal{S}}_{(1,i)} > 0$ *for all* $0 \le i \le I-1$ *and* $\Delta_1 g^{\mathcal{S}}_{(1,I)} = 0$.

Proof. We will divide the proof into three steps, in which we prove the following:

(i) if $1 \le K' := \min\{K, I\}$, then action-difference $\Delta_1 g^{\mathcal{S}}_{(1,K'-1)} > 0$;
(ii) if $K \le I-1$, then action-difference $\Delta_1 g^{\mathcal{S}}_{(1,K)} > 0$;
(iii) action-differences $\Delta_1 g^{\mathcal{S}}_{(1,i)} > 0$ for all $0 \le i \le I-1$ and $\Delta_1 g^{\mathcal{S}}_{(1,I)} = 0$.

(i) Suppose first that $1 \le K \le I-1$. Identity (43) gives

$$\Delta_1 g^{\mathcal{S}}_{(1,K-1)} = \lambda - A'_{K-1}\Delta_2 g^{\mathcal{S}}_{(0,K)}.$$

In order to show $\Delta_1 g^{\mathcal{S}}_{(1,K-1)} > 0$, using (48) we need to have

$$\lambda > \frac{\beta\lambda\,(1+B')\,A'_{K-1}}{\frac{A'_K}{A_K} + \beta A'_{K-1}B' + \beta\zeta\beta\mu(1-BA_{K-1})}.$$

Since the denominator is positive due to Lemma A5, this is equivalent to

$$\beta A'_{K-1} < \frac{A'_K}{A_K} + \beta\zeta\beta\mu(1-BA_{K-1}).$$

This is true, because Lemma A5(iii) with $\beta < 1$ implies $\beta A'_{K-1} < \frac{A'_K}{A_K}$, and Lemma A5(i)-(ii) implies $\beta\zeta\beta\mu(1-BA_{K-1}) > 0$.
For $K = I$ and $K = I+1$, we have $\Delta_1 g^{\mathcal{S}}_{(1,I-1)} = \lambda - A'_{I-1}\Delta_2 g^{\mathcal{S}}_{(0,I)}$ as above. In order to show $\Delta_1 g^{\mathcal{S}}_{(1,I-1)} > 0$, using (50) for $K = I$ we need to have

$$\lambda > \frac{\lambda(1+\beta\mu)A'_{I-1}}{\frac{A'_I}{A_I} + \beta\mu A'_{I-1}}, \qquad \text{which is equivalent to} \qquad 1 > \frac{A'_{I-1}}{\frac{A'_I}{A_I}},$$

which is true by Lemma A5(iii).
Finally, using (51) for $K = I + 1$ we need to have

$$\lambda > \frac{\lambda A_I A'_{I-1}}{A'_I}, \qquad \text{which is equivalent to} \qquad 1 > \frac{A_I A'_{I-1}}{A'_I},$$

which is again true by Lemma A5(iii). □

(ii) Similarly, for $1 \le K \le I - 1$ identity (45) gives

$$\Delta_1 g^{\mathcal{S}}_{(1,K)} = \lambda - B' \Delta_2 g^{\mathcal{S}}_{(1,K)}.$$

In order to show $\Delta_1 g^{\mathcal{S}}_{(1,K)} > 0$, using (49) we need to have

$$\lambda > \frac{\beta\lambda\left(1 + A'_{K-1}\right) B'}{\frac{A'_K}{A_K} + \beta A'_{K-1} B' + \beta\zeta\beta\mu(1 - BA_{K-1})},$$

which is equivalent to

$$\beta B' < \frac{A'_K}{A_K} + \beta\zeta\beta\mu(1 - BA_{K-1}).$$

This is true, because $\beta B' < \frac{A'_K}{A_K}$ using the definitions of B' and A'_K in (40) and properties from Lemma A5.
Finally, for $K = 0$, plugging (47) into (31) gives $\Delta_1 g^{\mathcal{S}}_{(1,0)} = \lambda > 0$. □

(iii) We will show that positivity of action-difference $\Delta_1 g^{\mathcal{S}}_{(1,K'-1)}$ implies $\Delta_1 g^{\mathcal{S}}_{(1,i)} > 0$ for all $0 \le i \le K' - 1$, and positivity of action-difference $\Delta_1 g^{\mathcal{S}}_{(1,K)}$ implies $\Delta_1 g^{\mathcal{S}}_{(1,i)} > 0$ for all $K \le i \le I - 1$.
Recursion (44) shows that $\Delta_1 g^{\mathcal{S}}_{(1,i)}$ is a weighted average of $\lambda > 0$ and $\Delta_1 g^{\mathcal{S}}_{(1,i+1)}$ for all $0 \le i \le K' - 2$ (the weights are between 0 and 1 due to Lemma A5). Since state-difference $\Delta_1 g^{\mathcal{S}}_{(1,K'-1)} > 0$ by (i), by induction we obtain $\Delta_1 g^{\mathcal{S}}_{(1,i)} > 0$ for all $0 \le i \le K' - 1$.
Similarly, recursion (46) shows that $\Delta_1 g^{\mathcal{S}}_{(1,i)}$ is a weighted average of $\lambda > 0$ and $\Delta_1 g^{\mathcal{S}}_{(1,i-1)}$ for all $K + 1 \le i \le I - 1$ (the weights are between 0 and 1 due to Lemma A5). Since state-difference $\Delta_1 g^{\mathcal{S}}_{(1,K)} > 0$ by (ii), by induction we obtain $\Delta_1 g^{\mathcal{S}}_{(1,i)} > 0$ for all $K \le i \le I - 1$.
In summary, we have shown that $\Delta_1 g^{\mathcal{S}}_{(1,i)} > 0$ for all $0 \le i \le I - 1$. Finally, $\Delta_1 g^{\mathcal{S}}_{(1,I)} = 0$ by (35). □

A.4 Positivity of Action-Differences in Total Work under Active Set $\widetilde{\mathcal{I}}_{K,K+1}$

Lemma A8. *Under active set $\mathcal{S} = \widetilde{\mathcal{I}}_{K,K+1}$,*

(i) if $1 \le K \le I - 1$, then

$$\Delta_2 g^{\mathcal{S}}_{(0,K)} = \frac{\beta\lambda\left[1 - \beta(1 - \zeta - \mu)\right]}{\left[1 - \beta^2(1 - \zeta - \mu)\right]\frac{A'_K}{A_K} + \beta\eta\beta(1 - \mu)\left[1 - \beta + \beta\mu(1 - A_{K-1})\right]}, \tag{55}$$

$$\Delta_2 g^{\mathcal{S}}_{(1,K+1)} = \frac{\left[1 - \beta(1 - \zeta - \mu)\right] + \beta\eta\left[\beta - A_{K-1} - \beta\mu(1 - A_{K-1})\right]}{1 - \beta(1 - \zeta - \mu)}\Delta_2 g^{\mathcal{S}}_{(0,K)}, \tag{56}$$

$$\Delta_1 g^{\mathcal{S}}_{(1,K)} = \lambda - \frac{A'_K}{A_K}\Delta_2 g^{\mathcal{S}}_{(0,K)}. \tag{57}$$

(ii) if $K = 0$, then

$$\Delta_1 g^{\mathcal{S}}_{(1,0)} = \lambda \frac{1-\beta+\beta\mu}{1-\beta^2(1-\zeta-\mu)+\beta\mu}, \tag{58}$$

$$\Delta_1 g^{\mathcal{S}}_{(1,1)} = \lambda \frac{1-\beta^2(1-\zeta-\mu)+\beta\mu-\beta B'}{1-\beta^2(1-\zeta-\mu)+\beta\mu}. \tag{59}$$

Proof. (i) Taking the difference of (27) for $i = K+1$ and (28) for $i = K$, gives

$$\Delta_2 g^{\mathcal{S}}_{(1,K+1)} = B\Delta_2 g^{\mathcal{S}}_{(0,K)} + \frac{B}{\mu}\Delta_1 g^{\mathcal{S}}_{(1,K)}. \tag{60}$$

Taking the difference of (28) for $i = K$ and for $i = K-1$, employing the identity (24) together with (41) for $i = K-1$ and the expression of A_K in (39) gives

$$\Delta_2 g^{\mathcal{S}}_{(1,K)} = \left[\beta - \beta\mu(1-A_{K-1})\right]\Delta_2 g^{\mathcal{S}}_{(0,K)}. \tag{61}$$

Taking the difference of (29) for $i = K$ and (30) for $i = K-1$, and plugging again (41) for $i = K-1$ and (39), gives

$$\frac{\beta\zeta}{A_K}\Delta_2 g^{\mathcal{S}}_{(0,K)} = \beta\Delta_1 g^{\mathcal{S}}_{(1,K)} - \beta\eta\Delta_2 g^{\mathcal{S}}_{(1,K)} + \beta\eta\Delta_2 g^{\mathcal{S}}_{(0,K)} + \beta\zeta\Delta_2 g^{\mathcal{S}}_{(1,K+1)}. \tag{62}$$

By (33), the right-hand side of the above identity is equal to $\lambda - \Delta_1 g^{\mathcal{S}}_{(1,K)} - \beta\mu\Delta_2 g^{\mathcal{S}}_{(0,K)}$, yielding (57). Using (57), (60) can be reformulated as

$$\Delta_2 g^{\mathcal{S}}_{(1,K+1)} = \frac{B\lambda}{\mu} - \frac{B}{\mu}\left[\frac{A'_K}{A_K} - \mu\right]\Delta_2 g^{\mathcal{S}}_{(0,K)}. \tag{63}$$

Further, (42), (60) and (61) can be used to reformulate (62) as (56). Finally, (56) and (63) after some algebra yield (55). □

(ii) (60) holds as before and simplifies to

$$\Delta_2 g^{\mathcal{S}}_{(1,1)} = \frac{B}{\mu}\Delta_1 g^{\mathcal{S}}_{(1,0)}. \tag{64}$$

By (33),

$$(1+\beta)\Delta_1 g^{\mathcal{S}}_{(1,0)} = \lambda - \beta\zeta\Delta_2 g^{\mathcal{S}}_{(1,1)}. \tag{65}$$

Solving and rearranging yields (58).
Further, (45) together with (64) and (58) give (59). □

Proposition A3. *Under active set $\mathcal{S} = \widetilde{\mathcal{I}}_{K,K+1}$ with $0 \le K \le I-1$, action-differences $\Delta_1 g^{\mathcal{S}}_{(1,i)} > 0$ for all $0 \le i \le I-1$ and $\Delta_1 g^{\mathcal{S}}_{(1,I)} = 0$.*

Proof. We will divide the proof into four steps, in which we prove the following:

(i) if $1 \le K$, then action-difference $\Delta_1 g^{\mathcal{S}}_{(1,K-1)} > 0$;

(ii) action-difference $\Delta_1 g^{\mathcal{S}}_{(1,K)} > 0$;
(iii) if $K \le I-2$, then action-difference $\Delta_1 g^{\mathcal{S}}_{(1,K+1)} > 0$;
(iv) action-differences $\Delta_1 g^{\mathcal{S}}_{(1,i)} > 0$ for all $0 \le i \le I-1$ and $\Delta_1 g^{\mathcal{S}}_{(1,I)} = 0$.

(i) Suppose first that $1 \le K$. Identity (43) gives

$$\Delta_1 g^{\mathcal{S}}_{(1,K-1)} = \lambda - A'_{K-1} \Delta_2 g^{\mathcal{S}}_{(0,K)}$$

and identity (57) gives

$$\Delta_1 g^{\mathcal{S}}_{(1,K)} = \lambda - \frac{A'_K}{A_K} \Delta_2 g^{\mathcal{S}}_{(0,K)}.$$

Because of Lemma A5(iii), we have $\Delta_1 g^{\mathcal{S}}_{(1,K-1)} > \Delta_1 g^{\mathcal{S}}_{(1,K)}$. The positivity of the latter is proved in (ii).

(ii) In order to show $\Delta_1 g^{\mathcal{S}}_{(1,K)} > 0$ for $1 \le K \le I-1$, using (55) we need to have

$$\lambda > \frac{\beta\lambda \left[1 - \beta(1-\zeta-\mu)\right] \frac{A'_K}{A_K}}{\left[1-\beta^2(1-\zeta-\mu)\right] \frac{A'_K}{A_K} + \beta\eta\beta(1-\mu)\left[1-\beta+\beta\mu(1-A_{K-1})\right]}.$$

Since the denominator is positive due to Lemma A5, this is equivalent to

$$\beta \frac{A'_K}{A_K} < \frac{A'_K}{A_K} + \beta\eta\beta(1-\mu)\left[1-\beta+\beta\mu(1-A_{K-1})\right].$$

This is true, because of the properties in Lemma A5(iii).

For $K=0$, positivity of action-difference $\Delta_1 g^{\mathcal{S}}_{(1,0)}$ can be seen directly in (58).

(iii) Similarly, if $K \le I-2$ identity (45) gives

$$\Delta_1 g^{\mathcal{S}}_{(1,K+1)} = \lambda - B' \Delta_2 g^{\mathcal{S}}_{(1,K+1)}.$$

In order to show $\Delta_1 g^{\mathcal{S}}_{(1,K+1)} > 0$ for $1 \le K$, using (56) where we have plugged (55), we need to have

$$\lambda > \frac{B'\beta\lambda \left\{\left[1-\beta(1-\zeta-\mu)\right] + \beta\eta\left[\beta - A_{K-1} - \beta\mu(1-A_{K-1})\right]\right\}}{\left[1-\beta^2(1-\zeta-\mu)\right] \frac{A'_K}{A_K} + \beta\eta\beta(1-\mu)\left[1-\beta+\beta\mu(1-A_{K-1})\right]},$$

which is equivalent to

$$\begin{aligned}&\left[1-\beta^2(1-\zeta-\mu)\right] \frac{A'_K}{A_K} + \beta\eta\beta(1-\mu)\left[1-\beta+\beta\mu(1-A_{K-1})\right] \\ &> B'\beta \left\{\left[1-\beta(1-\zeta-\mu)\right] + \beta\eta\left[1 - A_{K-1} - 1 + \beta - \beta\mu(1-A_{K-1})\right]\right\}.\end{aligned}$$

This can be further reformulated as

$$\begin{aligned}&\beta^2(\zeta+\mu)\frac{A'_K}{A_K} + \left(1-\beta^2\right) \frac{A'_K}{A_K} + \beta\eta\beta(1-\mu)\left[1-\beta+\beta\mu(1-A_{K-1})\right] \\ &> B'\beta \left\{\left[1-\beta(1-\zeta-\mu)\right] + \beta\eta(1-A_{K-1})\right\} - B'\beta\left\{\beta\eta\left[1-\beta+\beta\mu(1-A_{K-1})\right]\right\}.\end{aligned}$$

This is true, if the first terms of both sides (divided by $\beta > 0$) satisfy

$$\beta(\zeta+\mu)\frac{A'_K}{A_K} > B'\left\{[1-\beta(1-\zeta-\mu)]+\beta\eta(1-A_{K-1})\right\},$$

because the remaining two terms on the left-hand side are non-negative, and the last term on the right-hand side is non-positive. We further reformulate the last inequality as

$$\left(\frac{A'_K}{A_K}-B'\right)[1-\beta(1-\zeta-\mu)+\beta\eta(1-A_{K-1})] > [1-\beta+\beta\eta(1-A_{K-1})]\,\frac{A'_K}{A_K}.$$

Now, definitions in (39)-(40) imply the following identities:

$$\begin{aligned}\frac{A'_K}{A_K}-B' &= \frac{\beta\zeta}{A_K}-\beta\zeta B\\ 1-\beta(1-\zeta-\mu)+\beta\eta(1-A_{K-1}) &= \frac{A'_K}{A_K}+\beta\eta\\ 1-\beta+\beta\eta(1-A_{K-1}) &= \frac{\beta\zeta}{A_K}-\beta\zeta.\end{aligned}$$

The above inequality is therefore equivalent to

$$\left(\frac{\beta\zeta}{A_K}-\beta\zeta B\right)\left(\frac{A'_K}{A_K}+\beta\eta\right) > \left(\frac{\beta\zeta}{A_K}-\beta\zeta\right)\frac{A'_K}{A_K},$$

which is true because $B < 1$ and $\beta\eta \geq 0$.

For $K = 0$, positivity of action-difference $\Delta_1 g^{\mathcal{S}}_{(1,1)}$ given in (59) is straightforward after substituting for B' and using Lemma A5(i). □

(iv) As in the proof of Proposition A2(iii), one can show that positivity of action-difference $\Delta_1 g^{\mathcal{S}}_{(1,K-1)}$ implies $\Delta_1 g^{\mathcal{S}}_{(1,i)} > 0$ for all $0 \leq i \leq K-1$, and positivity of action-difference $\Delta_1 g^{\mathcal{S}}_{(1,K+1)}$ implies $\Delta_1 g^{\mathcal{S}}_{(1,i)} > 0$ for all $K+1 \leq i \leq I-1$.

Therefore, $\Delta_1 g^{\mathcal{S}}_{(1,i)} > 0$ for all $0 \leq i \leq I-1$. Finally, $\Delta_1 g^{\mathcal{S}}_{(1,I)} = 0$ by (35). □

B Admission Control Problem with Delay: Analysis of the Marginal Rewards

General case.

Analogously to section A, in this section we set out to obtain closed formulae for marginal rewards in the admission control problem with delay. The proofs are similar to the formers and therefore they are omitted.

B.1 Preliminaries

Next we state analogies of Lemma A1, Lemma A2, and Lemma A4 for reward measures in the admission control problem with delay. First we give a characterization of total rewards $f^{\mathcal{S}}_{(a,i)}$'s in terms of their action-differences $\Delta_1 f^{\mathcal{S}}_{(1,i)}$'s and state-differences $\Delta_2 f^{\mathcal{S}}_{(a,i)}$'s.

Lemma B1. *For a fixed active set $\mathcal{S}$,*

$$(1-\beta)f^{\mathcal{S}}_{(1,i)} = R_{(1,i)} - \beta\mu\Delta_2 f^{\mathcal{S}}_{(1,i)} \qquad \text{if } (1,i)\in\mathcal{S} \tag{66}$$

$$(1-\beta)f^{\mathcal{S}}_{(1,i)} = R_{(1,i)} - \beta\mu\Delta_2 f^{\mathcal{S}}_{(0,i)} - \beta\Delta_1 f^{\mathcal{S}}_{(1,i)} \qquad \text{if } (1,i)\notin\mathcal{S} \tag{67}$$

$$(1-\beta)f^{\mathcal{S}}_{(0,i)} = \begin{cases} R_{(0,I)} - \beta\mu\Delta_2 f^{\mathcal{S}}_{(1,I)} + \beta\Delta_1 f^{\mathcal{S}}_{(1,I)} & \text{if } i = I \\ R_{(0,i)} + \beta\zeta\Delta_2 f^{\mathcal{S}}_{(1,i+1)} - \beta\eta\Delta_2 f^{\mathcal{S}}_{(1,i)} + \beta\Delta_1 f^{\mathcal{S}}_{(1,i)} & \text{otherwise} \end{cases} \qquad \text{if } (0,i)\in\mathcal{S} \tag{68}$$

$$(1-\beta)f^{\mathcal{S}}_{(0,i)} = \begin{cases} R_{(0,I)} - \beta\mu\Delta_2 f^{\mathcal{S}}_{(0,I)} & \text{if } i = I \\ R_{(0,i)} + \beta\zeta\Delta_2 f^{\mathcal{S}}_{(0,i+1)} - \beta\eta\Delta_2 f^{\mathcal{S}}_{(0,i)} & \text{otherwise} \end{cases} \qquad \text{if } (0,i)\notin\mathcal{S} \tag{69}$$

The following lemma characterizes action-differences $\Delta_1 f^{\mathcal{S}}_{(1,i)}$'s in terms of state-differences $\Delta_2 f^{\mathcal{S}}_{(a,i)}$'s.

Lemma B2. *For a fixed active set $\mathcal{S}$ and any state $0 \le i \le I-1$,*

$$\Delta_1 f^{\mathcal{S}}_{(1,i)} = \Delta_1 R_{(1,i)} - \beta\zeta\Delta_2 f^{\mathcal{S}}_{(1,i+1)} - \beta(\mu-\eta)\Delta_2 f^{\mathcal{S}}_{(1,i)} \qquad \text{if } (0,i),(1,i)\in\mathcal{S} \tag{70}$$

$$\Delta_1 f^{\mathcal{S}}_{(1,i)} = \Delta_1 R_{(1,i)} - \beta\zeta\Delta_2 f^{\mathcal{S}}_{(0,i+1)} - \beta(\mu-\eta)\Delta_2 f^{\mathcal{S}}_{(0,i)} \qquad \text{if } (0,i),(1,i)\notin\mathcal{S} \tag{71}$$

$$(1+\beta)\Delta_1 f^{\mathcal{S}}_{(1,i)} = \Delta_1 R_{(1,i)} - \beta\zeta\Delta_2 f^{\mathcal{S}}_{(1,i+1)} - \beta\mu\Delta_2 f^{\mathcal{S}}_{(0,i)} + \beta\eta\Delta_2 f^{\mathcal{S}}_{(1,i)} \qquad \text{if } (0,i)\in\mathcal{S}, (1,i)\notin\mathcal{S} \tag{72}$$

$$(1-\beta)\Delta_1 f^{\mathcal{S}}_{(1,i)} = \Delta_1 R_{(1,i)} - \beta\zeta\Delta_2 f^{\mathcal{S}}_{(0,i+1)} - \beta\mu\Delta_2 f^{\mathcal{S}}_{(1,i)} + \beta\eta\Delta_2 f^{\mathcal{S}}_{(0,i)} \qquad \text{if } (0,i)\notin\mathcal{S}, (1,i)\in\mathcal{S} \tag{73}$$

$$\Delta_1 f^{\mathcal{S}}_{(1,I)} = 0. \tag{74}$$

Now we are ready to express marginal rewards $r^{\mathcal{S}}_{(a,i)}$'s in terms of action-differences $\Delta_1 f^{\mathcal{S}}_{(1,i)}$'s.

Lemma B3. *For a fixed active set $\mathcal{S}$,*

$$r^{\mathcal{S}}_{(1,i)} = \begin{cases} \beta\Delta_1 f^{\mathcal{S}}_{(1,0)} & \text{if } i = 0 \\ \beta\mu\Delta_1 f^{\mathcal{S}}_{(1,i-1)} + \beta(1-\mu)\Delta_1 f^{\mathcal{S}}_{(1,i)} & \text{otherwise} \end{cases} \tag{75}$$

$$r^{\mathcal{S}}_{(0,i)} = \begin{cases} \beta(1-\zeta)\Delta_1 f^{\mathcal{S}}_{(1,0)} + \beta\zeta\Delta_1 f^{\mathcal{S}}_{(1,1)} & \text{if } i = 0 \\ \beta\eta\Delta_1 f^{\mathcal{S}}_{(1,i-1)} + \beta\varepsilon\Delta_1 f^{\mathcal{S}}_{(1,i)} + \beta\zeta\Delta_1 f^{\mathcal{S}}_{(1,i+1)} & \text{otherwise} \\ r^{\mathcal{S}}_{(1,I)} & \text{if } i = I \end{cases} \tag{76}$$

Let, for $j \ge i \ge 0$,

$$C_{i,i} := 0, \qquad C_{i,j+1} := \left[C_{i,j} - \frac{\Delta_2 R_{(1,j+1)}}{\beta\mu}\right] B, \qquad C''_{i,j} := \Delta_1 R_{(1,j)} + \beta\zeta C_{i,j+1}, \tag{77}$$

$$D_0 := 0, \qquad D_{j+1} := \left[\beta\eta D_j - \Delta_2 R_{(1,j+1)}\right] \frac{A_{j+1}}{\beta\zeta}, \qquad D'_j := \Delta_1 R_{(1,j)} + \beta(\mu-\eta) D_j. \tag{78}$$

B.2 Calculation of Action-Differences in Total Reward

Since Lemma B3 characterizes marginal rewards $r^{\mathcal{S}}_{(a,i)}$'s in terms of action-differences $\Delta_1 f^{\mathcal{S}}_{(1,i)}$'s, in this subsection we focus on the calculation of the latter. In the following we show that all the relevant state-differences can be obtained by recursion from two pivot state-differences associated to the two thresholds K_0, K_1 under any policy $\widetilde{\mathcal{I}}_{K_0,K_1}$.

Lemma B4. *For a fixed active set $\mathcal{S} = \widetilde{\mathcal{I}}_{K_0,K_1}$,*

$$\Delta_2 f^{\mathcal{S}}_{(0,i)} = \Delta_2 f^{\mathcal{S}}_{(0,i+1)} A_i - D_i, \qquad \text{for } 1 \le i \le K'_0 - 1, \tag{79}$$

$$\Delta_2 f^{\mathcal{S}}_{(1,i)} = \Delta_2 f^{\mathcal{S}}_{(1,K_1)} B^{i-K_1} - C_{K_1,i}, \qquad \text{for } K_1 + 1 \le i \le I. \tag{80}$$

Further we identify a recursion to calculate action-differences $\Delta_1 f^{\mathcal{S}}_{(1,i)}$'s in terms of the two pivot state-differences $\Delta_2 f^{\mathcal{S}}_{(0,K'_0)}$ and $\Delta_2 f^{\mathcal{S}}_{(1,K_1)}$. Thus, this is a simplification of Lemma B2.

Proposition B1. *For a fixed active set $\mathcal{S} = \widetilde{\mathcal{I}}_{K_0,K_1}$,*

$$\Delta_1 f^{\mathcal{S}}_{(1,K'_0-1)} = D'_{K'_0-1} - \Delta_2 f^{\mathcal{S}}_{(0,K'_0)} A'_{K'_0-1}, \qquad \text{if } 1 \le K'_0, \tag{81}$$

$$\Delta_1 f^{\mathcal{S}}_{(1,i)} = D'_i(1 - Z_{i+1}) + \left[\Delta_1 f^{\mathcal{S}}_{(1,i+1)} + \beta\mu D_i - \Delta_2 R_{(1,i+1)}\right] Z_{i+1}, \qquad \text{for } 0 \le i \le K'_0 - 2, \tag{82}$$

$$\Delta_1 f^{\mathcal{S}}_{(1,K_1)} = C''_{K_1,K_1} - \Delta_2 f^{\mathcal{S}}_{(1,K_1)} B', \qquad \text{if } K_1 \le I - 1, \tag{83}$$

$$\Delta_1 f^{\mathcal{S}}_{(1,i)} = \left[C''_{K_1,i} + \beta(\mu - \eta) C_{K_1,i}\right] - \left[C''_{K_1,i-1} + \beta(\mu - \eta) C_{K_1,i-1}\right] B + \Delta_1 f^{\mathcal{S}}_{(1,i-1)} B, \qquad \text{for } K_1 + 1 \le i \le I - 1. \tag{84}$$

The above results help significantly simplify the subsequent analysis, which we present in the next subsection for optimal active set $\widetilde{\mathcal{I}}_{K,K}$, and where we identify closed-form expressions for pivot state-differences in total reward. The results under the active set $\widetilde{\mathcal{I}}_{K,K+1}$ are not necessary, since they are not implemented in the algorithm FA.

B.3 Pivot State-Differences under Active Set $\widetilde{\mathcal{I}}_{K,K}$

Lemma B5. *Under active set $\mathcal{S} = \widetilde{\mathcal{I}}_{K,K}$*

(i) if $K = 0$, then

$$\Delta_2 f^{\mathcal{S}}_{(1,0)} = 0. \tag{85}$$

(ii) if $1 \le K \le I - 1$,

$$\Delta_2 f^{\mathcal{S}}_{(0,K)} = -\frac{\frac{\beta\zeta}{A_K} D_K - \Delta_1 R_{(1,K)} + (1-\beta) C''_{K,K+1} - (\beta D'_{K-1} + \Delta_2 R_{(1,K)}) B' + \beta\mu(\beta\zeta D_{K-1} B + C''_{K,K+1} - \Delta_1 R_{(1,K-1)})}{\frac{A'_K}{A_K} + \beta A'_{K-1} B' + \beta\zeta\beta\mu(1 - B A_{K-1})}, \tag{86}$$

$$\Delta_2 f^{\mathcal{S}}_{(1,K)} = -\frac{\frac{\beta\zeta}{A_K} D_K - \Delta_1 R_{(1,K)} + (1-\beta) D'_{K-1} - (\beta C''_{K,K+1} + \Delta_2 R_{(1,K)}) A'_{K-1} + \beta\mu(\beta\zeta D_{K-1} + (C''_{K,K+1} - \Delta_1 R_{(1,K-1)}) A_{K-1})}{\frac{A'_K}{A_K} + \beta A'_{K-1} B' + \beta\zeta\beta\mu(1 - B A_{K-1})}. \tag{87}$$

(iii) if $K = I$,

$$\Delta_2 f^{\mathcal{S}}_{(0,I)} = -\frac{\frac{\beta\zeta}{A_I} D_I - \beta\mu D'_{I-1}}{\frac{A'_I}{A_I} + \beta\mu A'_{I-1}}. \tag{88}$$

(iv) *if* $K = I + 1$,

$$\Delta_2 f^{\mathcal{S}}_{(0,I)} = -\frac{\beta\zeta D_I}{A'_I}. \tag{89}$$