OPTIMAL ADMISSION AND ROUTING WITH CONGESTION-SENSITIVE CUSTOMER CLASSES

AYSE ASLAN 🆔

Department of Operations, University of Groningen, Groningen 9747 AD, The Netherlands
E-mail: ayse.aslan@rug.nl

This paper considers optimal admission and routing control in multi-class service systems in which customers can either receive quality regular service which is subject to congestion or can receive congestion-free but less desirable service at an alternative service station, which we call the self-service station. We formulate the problem within the Markov decision process framework and focus on characterizing the structure of dynamic optimal policies which maximize the expected long-run rewards. For this, value function and sample path arguments are used. The congestion sensitivity of customers is modeled with class-independent holding costs at the regular service station. The results show how the admission rewards of customer classes affect their priorities at the regular and self-service stations. We explore that the priority for regular service may not only depend on regular service admission rewards of classes but also on the difference between regular and self-service admission rewards. We show that optimal policies have monotonicity properties, regarding the optimal decisions of individual customer classes such that they divide the state space into three connected regions per class.

Keywords: admission control, congestion, Markov decision processes, revenue management, routing control

1. INTRODUCTION

This paper focuses on the challenge of service providers to cope with dynamic and varying customer demands with their limited resources. The service providers have to find effective control mechanisms to manage their revenues, or customer satisfaction, to make the best use of their service capacities. Most commonly, in the literature these service systems are modeled as multi-class queueing systems and *admission* or/and *dynamic routing* controls are studied, within the queueing systems formulations, for contributing to the management of dynamic demands from distinguished customers.

Admission control of multi-class queueing systems for revenue management is intensively studied. Miller [21] proved the optimality of the *trunk reservation policy* for a multi-class loss queueing system in which the rewards that customer classes pay for being admitted to the system are orderable and the service rates of customers are not dependent

---

774

on their classes. In this policy, acceptance decisions on individual customer classes have threshold structures, with respect to the number of customers in the system; if customer class i is accepted when there are n customers in the system, then class i should also be accepted when the system is less crowded. Moreover, the so-called *trunks* reserved per customer class depend on the admission rewards; if customer class i offers reward r_i which is greater than r_j, the admission reward that class j offers, then at any congestion level that class j is accepted, for sure class i must be also admitted to the system. After Miller's result, the optimality of trunk reservation policies in various other multi-class single-station queueing models are shown [8,9,13,14]. Feinberg and Yang [9] focused on the trunk reservation optimality for an $M/M/c/N$ queue. Later, Fan-Orzechowski and Feinberg [8] extended this, by proving the optimality of randomized trunk reservation policies for an $M/M/c/N$ model with constraints on the blocking probabilities of customer classes. The admission control studies which consider class-dependent service rates have focused on providing heuristic policies. The most common approaches of these studies include Linear Programming (LP) techniques [18] and asymptotic analyses [13].

In many service systems, waiting times of customers for being served are important service quality indicators, which can affect the revenues, or customer satisfaction levels, that service systems can achieve. For instance, waiting times can affect customer behavior in choosing among different service providers, and thus for a specific service provider, waiting times can be a determinant of the demand intensity of their system. The effects of system congestion on customer behavior in a queueing system are considered earliest by Naor [22], for an $M/M/1$ queue. Knudsen [15] and Stidham [27] considered an $M/M/k$ and a $GI/M/1$ queues, respectively, by extending results by Naor [22]. These studies do not consider that customers can be of different types, for the brevity of their discussion in explaining how the potentials of systems to obtain admission rewards are lost when customers behave greedily based on congestion levels.

Koçağa and Ward [16] considered congestion-related costs through the abandonment of customers in their single-class multi-server model for controlling the admission decisions of arriving customers. Atar and Lev-Ari [2] studied the admission control in a single-server model with retrials. In their model, holding costs are used as means to incorporate congestion sensitivity of customers. We note that congestion sensitivity of customers is a recognized issue in call centers. Ata and Peng [1] studied the callback option to mitigate congestion in call centers. In this study, arriving customers are routed to an offline queue to be called back later when they accept the callback offer; otherwise, they are routed to the online queue in which customers incur congestion-related waiting time costs. Feinberg and Yang [10] considered congestion effects through class-dependent holding costs in the admission control problem of a multi-class queue model. Feinberg and Yang [10], being inspired by Miller [21], used a continuous-time Markov decision process formulation and relative bias functions in their policy iteration algorithm for obtaining the optimal policy of an $M/M/c/N$ queue with class and congestion-dependent admission rewards. Authors of this study focused on extending trunk reservation properties to their model.

Similarly, we also consider a service system in which admission rewards of customer classes are dependent on the congestion levels they experience in the system. However, differently, we assume that congestion-sensitive customers have an alternative service station, in which they can commence their services immediately without waiting. In this paper, we have two stations which work in parallel in the service system. Our model is inspired from systems, in which customers can either get quality *regular* service at a station supplied with professionals, say the *regular station*, or can serve themselves at the *self-service* station, in an unsupervised fashion. There are several examples of such service systems in practice. For instance, in many service stations, thanks to the availability of online tools, self-help

desks (or self-check out points) are available for customers. For a more concrete example, let us consider the education sector. These days, online learning is assisting the teacher-led instruction in many schools with flexible educational models such as personalized learning. In these schools, learners may opt for learning by themselves, in a self-service fashion, instead of waiting for teacher assistance.

In these types of service systems, it is most likely that the regular service station will be costlier to operate than the self-service, such that the system is more likely to have a lower capacity at the regular station than at the self-service station. One can imagine that in such systems, customers would prefer the regular station for getting quality professional-assisted services. However, since the preference for the regular station will be prevalent for many customers and also since the capacity of the regular station would not be high due to costs, the congestion is likely to become an issue at the regular station. As a result, the preference of customers for regular service might decrease in favor of self-service, as the regular service station gets crowded.

For incorporating the congestion effects for the regular station, our model includes a finite buffer for this station such that there are holding costs incurred from the customers present at the station. With the holding costs, it is assumed that both the time spent waiting for service at the buffer and being served degrade the rewards obtained from customer admissions. Considering the nature of self-service, in which each self-server can only serve a single customer at a time, the self-service station is modeled as a loss network. In some service systems, there might also be a waiting room for using self-servers (e.g., self-checkouts in supermarkets). In this study, we assume that waiting for service is more relevant for the regular service station and thus only consider a waiting room for this station. With the introduction of the self-service station, the routing decisions can also be controlled. In this paper, for every customer class we dynamically, as a function of the state of the system, decide on whether we are going to *admit* the class into the system or not, and if we are accepting, then to which of the stations we should *route* for the maximization of long-run expected rewards. The problem of admission and routing control in such a service system is illustrated in Figure 1.

Dynamic routing control is also intensively studied. Compared with the dynamic admission control literature, congestion-related costs such as holding costs are much commonly considered in the control of routing arriving customers. In some cases, admission and routing controls are simultaneously considered. Bertsimas and Chryssikou [5] provided approximate LPs to extract heuristic admission and routing policies. Chong *et al.* [6] studied both routing and admission control for a two-class two-station system with class priorities. For systems with many service stations and/or customer classes, the use of asymptotic analysis is common in the literature to provide efficient routing policies. Bassamboo *et al.* [4] studied the optimal admission and routing control in the stochastic fluid approximation for a multi-class service system with multiple service pools. Dai and Tezcan [7] presented an asymptotically optimal routing policy under a heavy traffic regime for a parallel server system. Atar *et al.* [3] studied the routing control problem in the diffusion models of multi-class many-server queueing systems. Ward and Armony [28] considered routing control under the heavy traffic regime for multi-class systems with heterogeneous servers. For single-class systems with parallel servers with holding costs, there are many results on the optimal dynamic routing policies in the literature. For these systems, commonly the route-to-least-workload policy is explored [11,12,29].

The slow server problem, which considers a supporting slower server that can be switched on or off depending on the number of customers waiting in the common buffer of the system, is also relevant to the problem described in this paper, if we consider that self-servers are slower than regular ones [19,23]. If we only have a single customer class with no
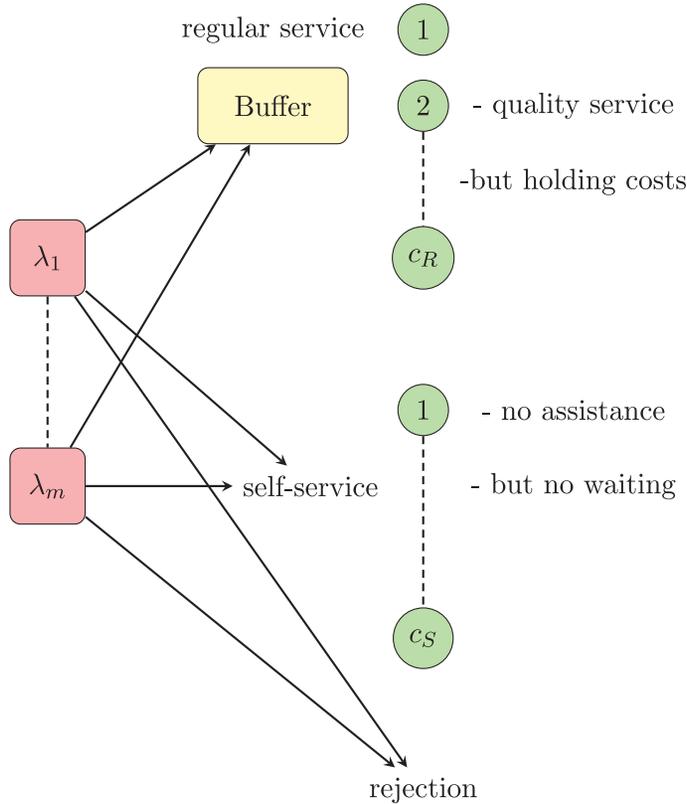
FIGURE **1.** Illustrating the admission and routing control problem in the service system with the regular and self-service stations, for customer classes $1, 2, \ldots, m$ with arrival rates $\lambda_1, \lambda_2, \ldots, \lambda_m$.

rejection option, then the routing decision to slower self-servers can be seen as the switching on decision of the slow server problem. Lin and Kumar [19] considered the *threshold* type policy which switches the slow server on when the queue length exceeds a threshold. Nobel and Tijms [23] additionally considered setup costs involving the switching the slow server on and off and focused on a two-level hysteric switching rule which switches the slow server on if the queue length exceeds an upper threshold and switches it off if the queue length becomes lower than a lower threshold. There are also studies on the slow server problem which considered the waiting option for getting service at the fast server. Rubinovitch [25] considered a model with a slow and a fast server with a buffer in front of the fast server for letting customers to greedily choose between the immediate service at the slow server or queueing up for the busy fast server.

We formulate the admission and routing control problem in this system as a Markov decision process and focus on characterizing the properties of the optimal policy. We make use of a discrete-time Markov decision process formulation (following [20,26]) to investigate the structure of optimal policies. For this, value function and sample path arguments are used intensively. The paper investigates inter-class and intra-class properties of the optimal stationary policies. In the inter-class properties, the focus is on the priorities of customer classes for receiving service at both stations. In the intra-class properties, we focus on characterizing monotonicity properties of the optimal policies with respect

to the number of customers present at stations, or say the remaining capacity levels at stations.

The remainder of the paper is organized as follows. Section 2 presents the Markov decision process formulation. Section 3 presents the properties of optimal policies. Lastly, we conclude the paper with Section 4.

2. MODEL FORMULATION

We consider that a set of customer classes $i \in I = \{1, \ldots, m\}$ with distinguished characteristics arrive to a service system with two parallel service stations which we refer to as *regular* and *self-service* stations. Each class $i \in I$ arrives according to a Poisson process with rate λ_i. The arriving customers can be routed to one of the two service stations, or can be rejected, dynamically. It is assumed that all servers in stations are exponential and service times are independent of customer classes. Let μ^R and μ^S denote the service rates, and c^R and c^S denote the number of servers at regular and self-service stations, respectively. The self-service station is considered as an $M/M/c^S/c^S$ loss network and the regular station is considered to have a buffer with capacity $B = N - c^R$ throughout the paper. Let $n^R \in S^1 = \{0, 1, \ldots, N\}$ and $n^S \in S^2 = \{0, 1, \ldots, c^S\}$ denote the state, the number of customers present, at regular and self-service stations, respectively. We consider that the non-negative reward that customer class i pays for regular service R_i^R is not less than the non-negative reward that the class would pay for self-service R_i^S. It is assumed that customers pay the admission rewards upon entry. Let $\mu^R(n^R)$ and $\mu^S(n^S)$ denote effective service rates, when there are n^R and n^S customers present at the regular and self-service stations, respectively. Although regular service is preferred by customers, it might involve holding costs. Let $H(n^R)$ denote the holding cost rate that the system incurs for having n^R customers at the regular station. In this study, we restrict to linear holding costs; we let h denote the unit time holding cost rate per customer at the regular station, so $H(n^R) = hn^R$.

We can consider $S = S^1 \times S^2$ as the state space of the system which observes the number of customers present at both stations. On this state space, we formulate a Markov decision process. Let us focus on stationary policies and let Π denote the set of such policies. Interevent times are exponentially distributed and their rates are always bounded above by $\Lambda := \sum_i \lambda_i + \mu^R(N) + \mu^S(c^S)$. Thus, we can apply uniformization (see [20,26]) and consider an equivalent decision process in discrete time. We let, without loss of generality, the uniformization constant Λ be equal to 1. Let $v_{T,\alpha}^\pi(s)$ denote the finite T-horizon α-discounted total expected reward under policy $\pi \in \Pi$ for the process starting at state $s \in S$. We define

$$v_{T,\alpha}^\pi(s) := \mathbb{E}\left[\sum_{t=0}^{T-1} \alpha^t r(S_t^\pi, A_t^\pi) | S_0^\pi = s\right], \tag{1}$$

where S_t^π denotes the state at the beginning of tth period, A_t^π denotes the action picked by policy π for it and $r(S_t^\pi, A_t^\pi)$ denotes the corresponding net reward collected resulting from this at state S_t^π (the sum of admission rewards resulting from the action picked minus the holding costs incurred during tth period). $v_{T,\alpha}^\pi(s)$ is well-defined for each initial state s and T, as the reward obtained at any state is bounded above by $\sum_i \lambda_i R_i^R$. We let $v_{T,\alpha}(s) = \sup_{\pi \in \Pi} v_{T,\alpha}^\pi(s)$ denote the optimal finite T-horizon α-discounted total expected reward obtained from the process which starts at state s.

We next define the infinite-horizon α-discounted total expected reward under policy $\pi \in \Pi$, for the process starting at state s

$$v_\alpha^\pi(s) := \lim_{T \to \infty} v_{T,\alpha}^\pi(s) = \mathbb{E}\left[\sum_{t=0}^\infty \alpha^t r(S_t^\pi, A_t^\pi)|S_0^\pi = s\right] \tag{2}$$

and

$$v^\pi(s) := \lim_{T \to \infty} \frac{1}{T} v_{T,1}^\pi(s), \tag{3}$$

for its long-run average reward counterpart. Note that for any policy π, long-run average reward $v^\pi(s)$ is independent of initial state s, as the Markov chain induced by any stationary policy π on our finite state space S is unichain. Next, we denote optimal expected rewards; let $v_\alpha(s) = \sup_{\pi \in \Pi} v_\alpha^\pi(s)$ denote the α-discounted total expected reward obtained by an optimal policy for the process which starts at state s and let $v = \sup_{\pi \in \Pi} v^\pi(s)$ denote the optimal long-run average reward.

By Theorem 6.2.6 of Puterman [24], we write the optimality equations $v_\alpha(s) = T_\alpha v_\alpha(s)$ for $s = (n^R, n^S)$, where

$$T_\alpha v_\alpha(n^R, n^S) = R_\alpha v_\alpha(n^R, n^S) + \alpha \mu^R(n^R) v_\alpha((n^R - 1)^+, n^S) + \alpha \mu^S(n^S) v_\alpha(n^R, (n^S - 1)^+)$$
$$+ \alpha\{\mu^R(N) - \mu^R(n^R) + \mu^R(c^S) - \mu^S(n^S)\} v_\alpha(n^R, n^S). \tag{4}$$

Here, we let $R_\alpha v_\alpha(n^R, n^S)$ be

$$- H(n^R) + \sum_i \lambda_i \max\{\mathbb{1}_{\{n^R < N\}}(R_i^R + \alpha v_\alpha(n^R + 1, n^S)),$$
$$\mathbb{1}_{\{n^S < c^S\}}(R_i^S + \alpha v_\alpha(n^R, n^S + 1)), \alpha v_\alpha(n^R, n^S)\}. \tag{5}$$

The three choices in $R_\alpha v_\alpha(n^R, n^S)(s)$, for each customer class i at state s, correspond to their routing to the regular and self-service stations and to their rejection, respectively.

By letting $v_{0,\alpha} = 0$, we write the optimality equations for the finite-horizon counterpart as

$$v_{T+1,\alpha}(s) = T_\alpha v_{T,\alpha}(s). \tag{6}$$

For the long-run average reward case, we let $y(s)$ denote the relative value function such that

$$y(s) - y(s') = \lim_{T \to \infty} v_{T,1}(s) - v_{T,1}(s') \tag{7}$$

and then, we write the optimality equation

$$v + y(s) = T_1 y(s). \tag{8}$$

This equation has a feasible solution for our finite state and action space unichain Markov decision process with bounded rewards by Theorem 8.4.3 of Puterman [24].

3. CHARACTERIZATION OF THE OPTIMAL POLICY

We characterize the properties of the optimal policy in this section.

In some single-station multi-class networks with static admission rewards and no holding costs, in which customer rewards can be ordered strictly, trunk reservation policies are

optimal (see [21]). In these policies, acceptance decisions of customer classes are directly related to their admission rewards. On contrary, in our system with two stations, optimal acceptance and routing decisions of different classes are affected by not only regular station admission rewards R_i^Rs, but also by the involved holding costs $H(n^R)$ and by the self-service admission rewards R_i^Ss.

3.1. Basic Properties

Let us begin by exploring some basic properties of the value functions.

LEMMA 3.1: *For any $\alpha \in [0,1)$, we have that (i) $v_\alpha(n^R, n^S) - v_\alpha(n^R + 1, n^S) \geq 0$, $\forall n^R < N$, $\forall n^S$ and (ii) $v_\alpha(n^R, n^S) - v_\alpha(n^R, n^S + 1) \geq 0, \forall n^R, \forall n^S < c^S$. Similarly, for the long-run average case, if v and y are the solutions of the optimality equation (8), then the above statements will hold with v_α replaced by y.*

PROOF: We use sample path arguments to show these results. Let us consider property (i) for v_α only, for a specific $\alpha \in [0,1)$. For y, or for property (ii), a similar proof will follow. Consider a process which starts at state $(n^R + 1, n^S)$ and follows an optimal policy π^*, let us call this Process 1. On the other hand, consider another process which starts at (n^R, n^S) and uses a (potentially) suboptimal policy π which imitates π^*, let us call this Process 2.

We suppose that these two processes are defined on the same probability space and thus move in parallel, observe the same arrival and service completion transitions simultaneously, whenever it is possible. However, some events can not be observed in both processes. Firstly, Process 1 can observe a service completion which Process 2 can not. Secondly, when Process 1 reaches a state in which there is no capacity left at both stations, the follower Process 2 would be at a state where there is still one more spot in the regular station. So, in this situation, if an arrival occurs, Process 1 has to reject this arrival, although Process 2 can admit the arrival to the regular station. We call these events *coupling* events such that after their occurrences, both processes transition into the same state, and behave identically thereafter.

Let δ be a random variable denoting the difference in the (net) reward obtained by Process 2 from that of Process 1, until one of the coupling events occurs. We have $\mathbb{E}(\delta) = v_\alpha^\pi(n^R, n^S) - v_\alpha(n^R + 1, n^S)$. Since $\mathbb{E}(\delta) = v_\alpha^\pi(n^R, n^S) - v_\alpha(n^R + 1, n^S) \leq v_\alpha(n^R, n^S) - v_\alpha(n^R + 1, n^S)$, it is sufficient that we show $\mathbb{E}(\delta) \geq 0$ to have $v_\alpha(n^R, n^S) - v_\alpha(n^R + 1, n^S) \geq 0$. The admission events occurring before coupling provide the same rewards to both processes (R_i^Rs and R_i^Ss for each class i). We need to also consider the holding costs for the congestion at the regular station. Since until coupling occurs Process 1 always experiences a regular station which is at least as crowded as the one in Process 2, the net rewards collected by Process 2 would be at least as large as the net rewards which could be obtained by Process 1 for any sample path. This implies that $\delta \geq 0$ pathwise, thus $\mathbb{E}(\delta) \geq 0$. ∎

Lemma 3.1 tells us that the service system benefits from more *idle* servers, or say spots in both of the stations, under the α-discounted total expected or long-run average reward optimalities. This is already intuitive.

We can additionally show that having idle servers or spots at the regular station is more beneficial than having them at the self-service station. This is possible because $R_i^R \geq R_i^S$ for any class i.

LEMMA 3.2: *For any $\alpha \in [0,1)$, we have that $v_\alpha(n^R, n^S + 1) - v_\alpha(n^R + 1, n^S) \geq 0$, $\forall n^R <$
N, $\forall n^S < c^S$. Similarly, for the long-run average case, if v and y are the solutions of the
optimality equation (8), then the above statement will hold with v_α replaced by y.*

PROOF: We again use sample path arguments. Let us consider this for v_α, for a specific $\alpha \in$
$[0,1)$. For y, a similar proof will follow. Consider a process which starts at state $(n^R + 1, n^S)$
and follows an optimal policy π^*, let us call this Process 1. On the other hand, consider
another process which starts at $(n^R, n^S + 1)$ and uses a (potentially) suboptimal policy
π by imitating π^*, let us call this Process 2. We again make the assumption that both
processes are defined on the same probability space; they observe the same arrival and
service completion events, whenever possible.

Let δ be a random variable denoting the difference in the (net) reward obtained by Pro-
cess 2 from that of Process 1, until a coupling event occurs. We have $\mathbb{E}(\delta) = v_\alpha^\pi(n^R, n^S +$
$1) - v_\alpha(n^R + 1, n^S)$. Again it is sufficient to show that $\mathbb{E}(\delta) \geq 0$ to have that $v_\alpha(n^R, n^S +$
$1) - v_\alpha(n^R + 1, n^S) \geq 0$ as $v_\alpha^\pi(n^R, n^S + 1) - v_\alpha(n^R + 1, n^S) \leq v_\alpha(n^R, n^S + 1) - v_\alpha(n^R +$
$1, n^S)$.

Events occurring before coupling provide the same admission rewards to both processes.
Also note that until coupling occurs Process 1 always experiences a regular station which
is at least as crowded as the one in Process 2, thus incurring larger holding costs. The first
coupling event could be the service completion event at the regular station for Process 2
and the service completion event at the self-service station for Process 1. In this event, there
are no admission rewards realized in both processes. The second possible coupling event can
occur when Process 1 is at a state in which there are no spots left to admit a customer to
the regular station, and on the contrary, Process 2 has still one spot left. In this situation,
Process 1 can admit an arriving customer to the self-service, but Process 2 can obtain a
larger reward by admitting the customer to the regular station instead as $R_i^R \geq R_i^S$ for each
class i and couple with Process 1 as a result. Thirdly, we can consider a situation in which
Process 2 can no longer admit customers to the self-service, but Process 1 still can, with its
single remaining idle server. If π^* chooses to admit an arriving customer to the self-service
in such a situation, then Process 2 can admit the same customer to the regular station and
can obtain a larger admission reward. So, whichever coupling event occurs first, we have
that $\delta \geq 0$ pathwise. Thus, we will have $\mathbb{E}(\delta) \geq 0$. ∎

3.2. Inter-Class Properties

Firstly, from the optimality equations (4)–(8), we can directly infer the following result.

PROPOSITION 3.3: *In an α-discounted $(\alpha \in [0,1))$ total expected reward optimal policy
(finite- or infinite-horizon formulations), or in a long-run average reward optimal policy, if
class i is routed to the regular (self-service) station at state (n^R, n^S) with reward R_i^R (R_i^S),
then any class j with reward $R_j^R \geq R_i^R$ $(R_j^S \geq R_i^S)$ and $R_j^S \leq R_i^S$ $(R_j^R \leq R_i^R)$ is also routed
to the regular service (self-service) at state (n^R, n^S) in the optimal policy.*

3.2.1. Highest priority for service In multi-class single-station systems with state-
independent rewards, we can usually talk about a customer class with the highest priority
to receive service. For instance, if we only have the regular station with $h = 0$, by using the
result by Miller [21], we can say that any class j with $R_j^R = \max_i R_i^R$ is among the customer
classes with highest service priority (for receiving service whenever there is capacity at the
service station).

For the case that $h = 0$ in our network with two stations, we have the following value function properties which indicate that priority for regular service is related to regular service rewards R_i^Rs and also on the difference between regular and self-service admission rewards.

LEMMA 3.4: *If $h = 0$, for any $\alpha \in [0, 1)$, we have that $v_\alpha(n^R, n^S) - v_\alpha(n^R + 1, n^S) \leq \max_i R_i^R$, $\forall n^R < N$, $\forall n^S$. Similarly, for the long-run average case, if v and y are the solutions of the optimality equation ([8](#)), then the above statement will hold with v_α replaced by y.*

PROOF: Let us show this for v_α, for a specific $\alpha \in [0, 1)$, with sample path arguments. Consider a process which starts at state (n^R, n^S) and follows an optimal policy π^*, let us call this Process 1. On the other hand, consider another process which starts at $(n^R + 1, n^S)$ and follows a (potentially) suboptimal policy π by imitating π^*, let us call this Process 2. We again make the assumption that both processes are defined on the same probability space.

Events occurring before coupling provide the same admission rewards to both processes. Firstly, Process 2 can observe a service completion which Process 1 can not. After this event, both processes will couple, without changing the rewards obtained, in any of them. Secondly, when Process 2 reaches a state in which there is no capacity left at both stations, Process 1 would be at a state where there is still one more spot in the regular station. When an arrival from class i occurs in this situation, Process 1 can obtain R_i^R while Process 2 will obtain no rewards, before they couple. This disadvantage of Process 2 can be at most $\max_i R_i^R$. That is why in any of these coupling cases, we will have that the (net) reward obtained by Process 2 will be at most $\max_i R_i^R$ less than of Process 1 pathwise, then we also have $v_\alpha(n^R, n^S) - v_\alpha^\pi(n^R + 1, n^S) \leq \max_i R_i^R$. Since $v_\alpha(n^R, n^S) - v_\alpha(n^R + 1, n^S) \leq v_\alpha(n^R, n^S) - v_\alpha^\pi(n^R + 1, n^S)$, this is sufficient to confirm the lemma. ∎

LEMMA 3.5: *If $h = 0$, for any $\alpha \in [0, 1)$, we have that $v_\alpha(n^R, n^S + 1) - v_\alpha(n^R + 1, n^S) \leq \max_i R_i^R - R_i^S$, $\forall n^R < N$, $\forall n^S < c^S$. Similarly, for the long-run average case, if v and y are the solutions of the optimality equation ([8](#)), then the above statement will hold with v_α replaced by y.*

PROOF: Let us show this for v_α, for a specific $\alpha \in [0, 1)$, by using sample path arguments. Consider a process which starts at state $(n^R, n^S + 1)$ and follows an optimal policy π^*, let us call this Process 1. On the other hand, consider another process which starts at $(n^R + 1, n^S)$ and follows a (potentially) suboptimal policy π by imitating π^*, let us call this Process 2. We assume that both processes are constructed in the same probability space.

Again events occurring before coupling provide the same admission rewards to both processes. The first coupling event could be the service completion event at the regular station for Process 2 and the service completion event at the self-service station for Process 1. In this event, there are no admission rewards realized in both processes. The second possible coupling event can occur when Process 2 is at a state in which there are no spots left to admit a customer to the regular station, and on the contrary, Process 1 has still one spot left. At this event, Process 2 can admit the arriving customer to the self-service, but Process 1 can obtain a larger reward by admitting the customer to the regular station. In this situation, considering arrivals from all possible customer classes, in any arrival coupling event, the advantage of the reward obtained by Process 1 over Process 2 can be at most $\max_i R_i^R - R_i^S$. Thirdly, we can consider a situation in which Process 1 can no longer admit customers to the self-service, but Process 2 has still one free server at the station. In this case,

when π^* admits an arriving customer to the regular service, we let Process 2 to admit the customer to the self-service station for coupling. So, in any of these coupling cases, we will have that the (net) reward obtained by Process 2 will be at most $\max_i R_i^R - R_i^S$ less than of Process 1 pathwise, then we also have $v_\alpha(n^R, n^S + 1) - v_\alpha^\pi(n^R + 1, n^S) \leq \max_i R_i^R - R_i^S$. Since $v_\alpha(n^R, n^S + 1) - v_\alpha(n^R + 1, n^S) \leq v_\alpha(n^R, n^S + 1) - v_\alpha^\pi(n^R + 1, n^S)$, this is sufficient to confirm the lemma. ∎

With these two lemmas, we can show the following.

PROPOSITION 3.6: *When the regular station is free of holding costs ($h = 0$), if for any customer class j we have that $R_j^R = \max_i R_i^R$ and $R_j^S = \min_i R_i^S$, then this class is routed to the regular station at any state $s = (n^R, n^S)$ whenever $n^R < N$, in an α-discounted ($\alpha \in [0, 1)$) total expected or long-run average reward optimal policy.*

PROOF: We show this for α-discounted ($\alpha \in [0, 1)$) total expected reward optimality. Class j is, preferably, routed to the regular station, than to be rejected, at any state $s = (n^R, n^S), n^R < N$ when $R_j^R + \alpha v_\alpha(n^R + 1, n^S) > \alpha v_\alpha(n^R, n^S)$. With Lemma 3.4, this holds for class j always. Moreover, this class is, preferably, routed to the regular, than to the self-service station, at any state $s = (n^R, n^S), n^R < N, n^S < c^S$ when $R_j^R + \alpha v_\alpha(n^R + 1, n^S) > R_j^S + \alpha v_\alpha(n^R, n^S + 1)$. With Lemma 3.5, this holds for class j always. ∎

For the self-service station, which is modeled as a loss network and in which there are no holding costs involved, we have the following value function property.

LEMMA 3.7: *For any $\alpha \in [0, 1)$, we have that $v_\alpha(n^R, n^S) - v_\alpha(n^R, n^S + 1) \leq \max_i R_i^S$, $\forall n^S < c^S$, $\forall n^R$. Similarly, for the long-run average case, if v and y are the solutions of the optimality equation (8), then the above statement will hold with v_α replaced by y.*

PROOF: We again use sample path arguments. Let us show this for v_α, for a specific $\alpha \in [0, 1)$. Consider a process which starts at state (n^R, n^S) and follows an optimal policy π^*, let us call this Process 1. On the other hand, consider another process which starts in $(n^R, n^S + 1)$ and uses a (potentially) suboptimal policy π which imitates π^*, let us call this Process 2. We again make the assumption that both processes are constructed in the same probability space. Again events occurring before coupling provide the same admission rewards to both processes. Also, before coupling, both processes incur the same holding costs.

Firstly, Process 2 can observe a service completion which Process 1 can not. After this event, both processes will couple, without changing the rewards obtained, in any of them. Secondly, when Process 2 reaches a state in which there is no capacity left at both stations, Process 1 would be at a state where there is still one more spot in the self-service station. When an arrival from class i occurs in this situation, Process 1 can obtain R_i^S more rewards than Process 1, before they couple. This disadvantage of Process 2 can be at most $\max_i R_i^S$. That is why in any of these coupling cases, we will have that the (net) reward obtained by Process 2 will be at most $\max_i R_i^S$ less than of Process 1 pathwise, then we also have $v_\alpha(n^R, n^S) - v_\alpha^\pi(n^R, n^S + 1) \leq \max_i R_i^S$. Since $v_\alpha(n^R, n^S) - v_\alpha(n^R, n^S + 1) \leq v_\alpha(n^R, n^S) - v_\alpha^\pi(n^R, n^S + 1)$, this is sufficient to confirm the lemma. ∎

PROPOSITION 3.8: *Under an α-discounted ($\alpha \in [0, 1)$) total expected or long-run average reward optimal policy, at any state $s = (n^R, n^S)$ with $n^S < c^S$, any class j with $R_j^S = \max_i R_i^S$ is not rejected.*

PROOF: We show this for α-discounted ($\alpha \in [0,1)$) total expected reward optimality. Class j is, preferably, routed to the self-service station, than to be rejected, at any state $s = (n^R, n^S), n^R < N, n^S < c^S$ when $R_j^S + \alpha v_\alpha(n^R, n^S + 1) > \alpha v_\alpha(n^R, n^S)$. With Lemma 3.7, this holds for class j always. ∎

3.3. Intra-Class Properties

In this section, we show the monotonicity properties of the optimal policy, regarding the optimal admission and routing decisions of individual customer classes as functions of the state of the system, or say the free capacities of the stations. Value function arguments are used for this.

The following lemma is useful to infer the monotonicity properties of the optimal policy. In naming these value function properties, we adapt the terminology by Koole [17].

LEMMA 3.9: *For any $\alpha \in [0,1)$, we have that*

(i) (Convexity in n^S) $v_\alpha(n^R, n^S) - v_\alpha(n^R, n^S + 1) \leq v_\alpha(n^R, n^S + 1) - v_\alpha(n^R, n^S + 2)$, $\forall n^S \leq c^S - 2, \forall n^R$

(ii) (Supermodularity) $v_\alpha(n^R, n^S) - v_\alpha(n^R, n^S + 1) \leq v_\alpha(n^R + 1, n^S) - v_\alpha(n^R + 1, n^S + 1), \forall n^S \leq c^S - 1, \forall n^R \leq N - 1$

(iii) (Superconvexity-1) $v_\alpha(n^R + 1, n^S) - v_\alpha(n^R + 1, n^S + 1) \leq v_\alpha(n^R, n^S + 1) - v_\alpha(n^R, n^S + 2), \forall n^S \leq c^S - 2, \forall n^R \leq N - 1$

(iv) (Superconvexity-2) $v_\alpha(n^R, n^S + 1) - v_\alpha(n^R + 1, n^S + 1) \leq v_\alpha(n^R + 1, n^S) - v_\alpha(n^R + 2, n^S), \forall n^S \leq c^S - 1, \forall n^R \leq N - 2$

(v) (Convexity in n^R) $v_\alpha(n^R, n^S) - v_\alpha(n^R + 1, n^S) \leq v_\alpha(n^R + 1, n^S) - v_\alpha(n^R + 2, n^S), \forall n^R \leq N - 2, \forall n^S$

Similarly, for the long-run average case, if v and y are the solutions of the optimality equation (8), then the above statements will hold with v_α replaced by y.

The proof of the lemma can be found in the Appendix.

When we consider an α-discounted total expected or the long-run average reward optimality, these properties suggest the following. Property (i) tells that more capacity at the self-service station means that more customers can be admitted to the station. Property (ii) can be interpreted as more capacity at the self-service (regular) station means that more customers can be admitted to the regular (self-service) station. Property (iii) shows that more capacity at the self-service station has more potential to increase admissions to the self-service than to the regular service station. Likewise, property (iv) shows the counterpart of this for the regular service station. Lastly, property (v) indicates that more capacity at the regular service station means that more customers can be admitted to the station.

The following propositions imply that α-discounted total expected or long-run average reward optimal policies have monotonicity properties.

PROPOSITION 3.10: *Under an α-discounted ($\alpha \in [0,1)$) total expected or long-run average reward optimal policy, if any class j is routed to the self-service station at any state $s = (n^R, n^S), n^S < c^S$, then class j is routed to the self-service station also at state $(n^R, n^S - 1)$.*

PROOF: We show this for α-discounted total expected reward optimality. Class j is, preferably, routed to the self-service station, than being rejected, at any state $s = (n^R, n^S), n^R < N, n^S < c^S$ when $R_j^S + \alpha v_\alpha(n^R, n^S + 1) > \alpha v_\alpha(n^R, n^S)$. With property (i) of Lemma 3.9, we

have $\alpha v_\alpha(n^R, n^S) - \alpha v_\alpha(n^R, n^S + 1) \geq \alpha v_\alpha(n^R, n^S - 1) - \alpha v_\alpha(n^R, n^S)$, which shows that $R_j^S > \alpha v_\alpha(n^R, n^S - 1) - \alpha v_\alpha(n^R, n^S)$. This confirms that class j will not be rejected from the self-service station at state $(n^R, n^S - 1)$. Moreover, class j is routed to the self-service station, than to the regular service station at state $s = (n^R, n^S), n^R < N, n^S < c^S$ when $\alpha v_\alpha(n^R, n^S + 1) - \alpha v_\alpha(n^R + 1, n^S) > R_j^R - R_j^S$. With property (iii) of Lemma 3.9, $\alpha v_\alpha(n^R, n^S) - \alpha v_\alpha(n^R + 1, n^S - 1) \geq \alpha v_\alpha(n^R, n^S + 1) - \alpha v_\alpha(n^R + 1, n^S)$ (note that $n^S - 1 \leq c^S - 2$). So, we have $\alpha v_\alpha(n^R, n^S) - \alpha v_\alpha(n^R + 1, n^S - 1) > R_j^R - R_j^S$, which confirms the advantage of routing the class j to the self-service over the regular service at state $(n^R, n^S - 1)$. ∎

PROPOSITION 3.11: *Under an α-discounted ($\alpha \in [0, 1)$) total expected or long-run average reward optimal policy, if any class j is routed to the regular service station at any state $s = (n^R, n^S)$, $n^R < N$, then class j is routed to the regular service station also at state $(n^R - 1, n^S)$.*

PROOF: We show this for α-discounted total expected reward optimality. We can prove this with the help of properties (iv) and (v) of Lemma 3.9, in a similar fashion that we describe in the proof of Proposition 3.10. ∎

In our model, we also control admission decisions such that we have the option to reject customers. The following property presents the monotonicity of rejection decisions of customer classes.

PROPOSITION 3.12: *Under an α-discounted ($\alpha \in [0, 1)$) total expected or long-run average reward optimal policy, if any class j is rejected at any state $s = (n^R, n^S)$, then class j is rejected also at states $(n^R + 1, n^S)$ and $(n^R, n^S + 1)$.*

PROOF: We show this for α-discounted total expected reward optimality. Let us show this for any state $s = (n^R, n^S)$ with $n^R < N, n^S < c^S$ (in which both of the routing options are feasible). At state $s = (n^R, n^S), n^R < N, n^S < c^S$ class j is rejected when $R_i^R + \alpha v_\alpha(n^R + 1, n^S) < \alpha v_\alpha(n^R, n^S)$ and $R_i^S + \alpha v_\alpha(n^R, n^S + 1) < \alpha v_\alpha(n^R, n^S)$. First, let us show that class j is rejected also at state $(n^R + 1, n^S)$. By property (v) of Lemma 3.9, $\alpha v_\alpha(n^R, n^S) - \alpha v_\alpha(n^R + 1, n^S) \leq \alpha v_\alpha(n^R + 1, n^S) - \alpha v_\alpha(n^R + 2, n^S)$. So, we have $R_j^R < \alpha v_\alpha(n^R + 1, n^S) - \alpha v_\alpha(n^R + 2, n^S)$, which shows that class j is rejected from the regular service station at state $(n^R + 1, n^S)$. By property (ii) of Lemma 3.9, $\alpha v_\alpha(n^R, n^S) - \alpha v_\alpha(n^R, n^S + 1) \leq \alpha v_\alpha(n^R + 1, n^S) - \alpha v_\alpha(n^R + 1, n^S + 1)$. With this, we have that $R_j^S < \alpha v_\alpha(n^R + 1, n^S) - \alpha v_\alpha(n^R + 1, n^S + 1)$, which confirms that class j is rejected also from the self-service station at state $(n^R + 1, n^S)$. For showing that class j is rejected from the self-service and regular service stations also at state $(n^R, n^S + 1)$, we can use properties (i) and (ii) of Lemma 3.9 in a similar fashion. ∎

These monotonicity properties imply that the discounted total expected or long-run average reward optimal policies divide the state space into three connected regions for any customer class such that in each region either the class is routed to the regular service or self-service stations, or it is rejected. For illustrating this, we present Figure 2 for a system with two classes. In obtaining this figure, we let $\lambda_1 = \lambda_2 = 3$, $\mu^R = 1$, $\mu_S = 0.5$, $h = 2$, $c^R = 3$, $c^S = 10$, $B = 5$ and consider long-run average reward optimal policies.

Policies in Figure 2(a) and (b) correspond to the scenario that we have $R_1^R = 15$ and $R_1^S = 2$ for the first customer class and $R_2^R = 13$ and $R_1^S = 3$ for the second customer class. Figure 2(a) and (b) presents the optimal admission and routing decisions of the first and
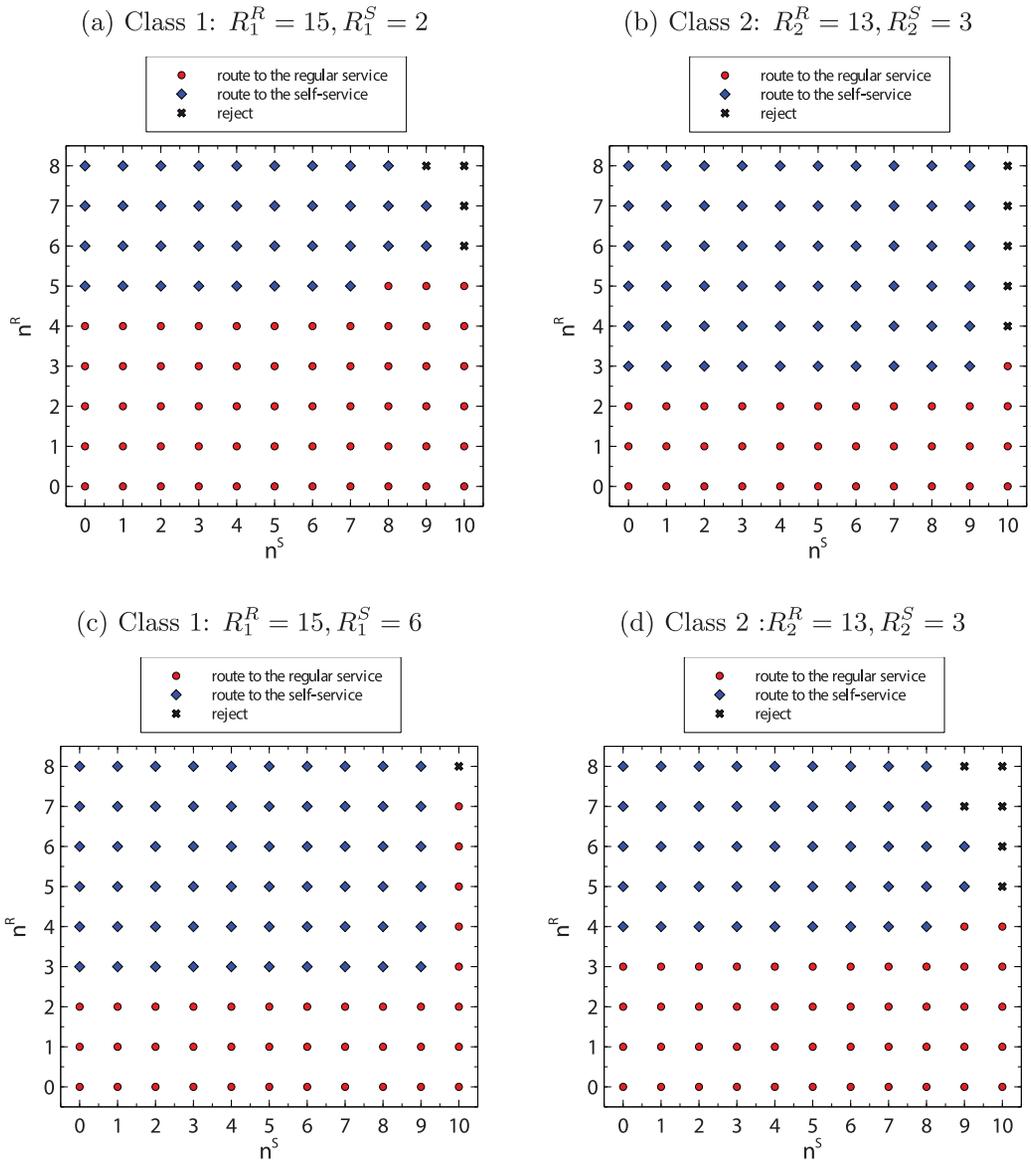
FIGURE 2. Long-run average reward optimal policies of a two class system with $\lambda_1 = \lambda_2 = 3$, $\mu^R = 1$, $\mu_S = 0.5$, $h = 2$, $c^R = 3$, $c^S = 10$, $B = 5$ under two different admission reward scenarios. (a) Class 1: $R_1^R = 15$, $R_1^S = 2$. (b) Class 2: $R_2^R = 13$, $R_2^S = 3$. (c) Class 1: $R_1^R = 15$, $R_1^S = 6$. (d) Class 2: $R_2^R = 13$, $R_2^S = 3$.

second customer classes, respectively, for this scenario. We observe that the first class is accepted to the regular station more than the second class. Note that the regular service admission reward of the first class is higher than of the second class ($R_1^R > R_2^R$) and also the reward difference $R_1^R - R_1^S$ is higher than $R_2^R - R_2^S$. For the systems with $h = 0$, we have the result in Proposition 3.6 which confirms the regular service priority of the class who has the highest regular service admission reward and the lowest self-service admission

reward. For the system in this figure with $h = 2$, we are able to observe the regular service priority of such a customer class.

In order to illustrate how the optimal policies change with respect to the admission rewards of customers, we then look at Figure 2(c) and (d). Differently from the setting in Figure 2(a) and (b), we have $R_1^S = 6$, instead of $R_1^S = 2$. Figure 2(c) and (d) presents the optimal admission and routing decisions of the first and second customer classes, respectively, for this different scenario. In this scenario, we still have that the first class has the highest regular service admission reward. However, this time the difference between the regular service and self-service admission rewards is higher for the second customer class. We can observe from the figure how this increase in the self-service admission reward of the first customer class reduces the priority of the first customer class for the regular service and at the same time increases the regular service priority of the second class.

4. CONCLUSIONS

This paper focuses on the optimal dynamic admission and routing control problem for the revenue management in a specific service system setting. In this setting, we imagine that customers arriving are not identical with respect to their rewards for service and they are sensitive to congestion. For modeling these differences, we consider customer classes. We get inspiration from service systems in which customers have the option to receive congestion-free services, for instance by self-serving their own demands through self-help desks, instead of opting for regular service, which, as we imagine, is provided by professionals. We argue that in such systems, the regular service can entail congestion-related costs due to high demands by customers and/or low staff levels. For studying the optimal admission and routing control of such systems, a queueing model with two parallel stations with multi-servers is devised. The congestion-free self-service station is represented with a loss network. For the regular service station, a finite buffer where customers incur holding costs is used. The focus of this paper is on the discounted total expected and long-run average reward optimal policies. We use Markov decision process formulations to characterize the structure of optimal policies and show the well-structuredness of optimal policies, through value function and sample path arguments.

References

1. Ata, B. & Peng, X. (2020). An optimal callback policy for general arrival processes: a pathwise analysis. *Operations Research* 68: 327–347.
2. Atar, R. & Lev-Ari, A. (2018). Optimizing buffer size for the retrial queue: two state space collapse results in heavy traffic. *Queueing Systems* 90: 225–255.
3. Atar, R., Mandelbaum, A., & Shaikhet, G. (2009). Simplified control problems for multiclass many-server queueing systems. *Mathematics of Operations Research* 34: 795–812.
4. Bassamboo, A., Harrison, J.M., & Zeevi, A. (2005). Dynamic routing and admission control in high-volume service systems: asymptotic analysis via multi-scale fluid limits. *Queueing Systems* 51: 249–285.
5. Bertsimas, D. & Chryssikou, T. (1999). Bounds and policies for dynamic routing in loss networks. *Operations Research* 47: 379–394.
6. Chong, K.C., Henderson, S.G., & Lewis, M.E. (2018). Two-class routing with admission control and strict priorities. *Probability in the Engineering and Informational Sciences* 32: 163–178.
7. Dai, J.G. & Tezcan, T. (2008). Optimal control of parallel server systems with many servers in heavy traffic. *Queueing Systems* 59: 95–134.
8. Fan-Orzechowski, X. & Feinberg, E.A. (2006). Optimality of randomized trunk reservation for a problem with a single constraint. *Advances in Applied Probability* 38: 199–220.
9. Feinberg, E.A. & Reiman, M.I. (1994). Optimality of randomized trunk reservation. *Probability in the Engineering and Informational Sciences* 8: 463–489.

10. Feinberg, E.A. & Yang, F. (2011). Optimality of trunk reservation for an M/M/K/N queue with several customer types and holding costs. *Probability in the Engineering and Informational Sciences* 25: 537–560.
11. Frostig, E. & Levikson, B. (1999). Optimal routing of customers to two parallel heterogeneous servers: the case of IHR service times. *Operations Research* 47: 438–444.
12. Hordijk, A. & Koole, G. (1992). On the assignment of customers to parallel queues. *Probability in the Engineering and Informational Sciences* 6: 495–511.
13. Hunt, P.J. & Laws, C.N. (1997). Optimization via trunk reservation in single resource loss systems under heavy traffic. *Annals of Applied Probability* 7: 1058–1079.
14. Key, P. (1990). Optimal control and trunk reservation in loss networks. *Probability in the Engineering and Informational Sciences* 4: 203–242.
15. Knudsen, N.C. (1972). Individual and social optimization in a multiserver queue with a general cost-benefit structure. *Econometrica* 40: 515–528.
16. Koçağa, Y.L. & Ward, A.R. (2010). Admission control for a multi-server queue with abandonment. *Queueing Systems* 65: 275–323.
17. Koole, G. (2006). Monotonicity in Markov reward and decision chains: theory and applications. *Foundation and Trends in Stochastic Systems* 1: 1–76.
18. Levi, R. & Radovanovic, A. (2010). Provably near-optimal LP-based policies for revenue management in systems with reusable resources. *Operations Research* 58: 503–507.
19. Lin, W. & Kumar, P. (1984). Optimal control of a queueing system with two heterogeneous servers. *IEEE Transactions on Automatic Control* 29: 696–703.
20. Lippman, S.A. (1975). Applying a new device in the optimization of exponential queuing systems. *Operations Research* 23: 687–710.
21. Miller, B.L. (1969). A queueing reward system with several customer classes. *Management Science* 16: 234–245.
22. Naor, P. (1969). The regulation of queue size by levying tolls. *Econometrica* 37: 15–24.
23. Nobel, R.D. & Tijms, H.C. (2000). Optimal control of a queueing system with heterogeneous servers and setup costs. *IEEE Transactions on Automatic Control* 45: 780–784.
24. Puterman, M.L. (1994). *Markov decision processes: discrete stochastic dynamic programming.* Hoboken, NJ: John Wiley and Sons.
25. Rubinovitch, M. (1985). The slow server problem: a queue with stalling. *Journal of Applied Probability* 22: 879–892.
26. Serfozo, R.F. (1979). Technical note: An equivalence between continuous and discrete time Markov decision processes. *Operations Research* 27: 616–620.
27. Stidham, S. (1978). Socially and individually optimal control of arrivals to a GI/M/1 queue. *Management Science* 24: 1598–1610.
28. Ward, A.R. & Armony, M. (2013). Blind fair routing in large-scale service systems with heterogeneous customers and servers. *Operations Research* 61: 228–243.
29. Winston, W. (1977). Optimality of the shortest line discipline. *Journal of Applied Probability* 14: 181–189.

APPENDIX

A.1. Proof of Lemma 3.9

We show this lemma on the α-discounted finite-horizon formulation $v_{T,\alpha}$, by using induction on T (the number of remaining time periods). We can then reason that the lemma will hold for v_α or y.

Let us start by showing the following.

LEMMA A.1: *For any $T \geq 1$ and $\alpha \in [0, 1)$, we have that*

(i) $v_{T,\alpha}(n^R, n^S) - v_{T,\alpha}(n^R, n^S + 1) \leq v_{T+1,\alpha}(n^R, n^S) - v_{T+1,\alpha}(n^R, n^S + 1)$

(ii) $v_{T,\alpha}(n^R, n^S) - v_{T,\alpha}(n^R + 1, n^S) \leq v_{T+1,\alpha}(n^R, n^S) - v_{T+1,\alpha}(n^R + 1, n^S)$

PROOF: This lemma can be shown using sample path arguments. Here, we do it for the first statement. Construct two processes on the same probability space. Let Process 1 start at state

(n^R, n^S) and let Process 2 start at state $(n^R + 1, n^S)$. If these processes couple somewhere in the first T periods, then the difference in the rewards obtained by two processes will be the same if both processes have T or $T + 1$ periods remaining. If not, then Process 1 will obtain as many rewards as Process 2 in the first T periods, and in the last remaining period, there is no chance that the advantage of Process 1 over Process 2 will decrease. ∎

Now it is sufficient to show the following statements, to prove Lemma 3.9. For any $T \geq 1$ and $\alpha \in [0, 1)$, we have that

(I) $v_{T,\alpha}(n^R, n^S) - v_{T,\alpha}(n^R, n^S + 1) \leq v_{T,\alpha}(n^R, n^S + 1) - v_{T,\alpha}(n^R, n^S + 2)$, $\quad \forall n^S \leq c^S - 2$, $\forall n^R$

(II) $v_{T,\alpha}(n^R, n^S) - v_{T,\alpha}(n^R, n^S + 1) \leq v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 1, n^S + 1)$, $\forall n^S \leq c^S - 1$, $\forall n^R \leq N - 1$

(III) $v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 1, n^S + 1) \leq v_{T,\alpha}(n^R, n^S + 1) - v_{T,\alpha}(n^R, n^S + 2)$, $\forall n^S \leq c^S - 2$, $\forall n^R \leq N - 1$

(IV) $v_{T,\alpha}(n^R, n^S + 1) - v_{T,\alpha}(n^R + 1, n^S + 1) \leq v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 2, n^S)$, $\forall n^S \leq c^S - 1$, $\forall n^R \leq N - 2$

(V) $v_{T,\alpha}(n^R, n^S) - v_{T,\alpha}(n^R + 1, n^S) \leq v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 2, n^S)$, $\forall n^S$, $\forall n^R \leq N - 2$

Proving Statement (I): This is the convexity property of the value functions with respect to the number of customers in the self-service station. We show this by induction on the value functions. We can show the initial induction step by letting $v_{0,\alpha}(\ldots) = 0$. Then, assuming that (I) holds for some T and $\alpha \in [0, 1)$ and we need to show that (I) also holds for $T + 1$ and $\alpha \in [0, 1)$.

We know that

$$
\begin{aligned}
v_{T+1,\alpha}(n^R, n^S) - v_{T+1,\alpha}(n^R, n^S + 1) &= R_\alpha v_{T,\alpha}(n^R, n^S) - R_\alpha v_{T,\alpha}(n^R, n^S + 1) \\
&+ \alpha \mu^R(n^R)[v_{T,\alpha}(n^R - 1, n^S) - v_{T,\alpha}(n^R - 1, n^S + 1)] \\
&+ \alpha \mu^S(n^S)[v_{T,\alpha}(n^R, n^S - 1) - v_{T,\alpha}(n^R, n^S)] \\
&+ \alpha(\mu^S(c^S - (n^S + 1)) + \mu^R(N) - \mu^R(n^R)) \\
&\times [v_{T,\alpha}(n^R, n^S) - v_{T,\alpha}(n^R, n^S + 1)]
\end{aligned}
\tag{A.1}
$$

where $R_\alpha v_{T,\alpha}(n^R, n^S) - R_\alpha v_{T,\alpha}(n^R, n^S + 1)$ is

$$
\sum_i \lambda_i [\max[R_i^R + \alpha v_{T,\alpha}(n^R + 1, n^S), R_i^S + \alpha v_{T,\alpha}(n^R, n^S + 1), \alpha(n^R, n^S)]
$$
$$
- \max[R_i^R + \alpha v_{T,\alpha}(n^R + 1, n^S + 1), R_i^S + \alpha v_{T,\alpha}(n^R, n^S + 2), \alpha(n^R, n^S + 1)]].
\tag{A.2}
$$

Note that the holding costs $(h(n^R))$ cancel out in $R_\alpha v_{T,\alpha}(n^R, n^S) - R_\alpha v_{T,\alpha}(n^R, n^S + 1)$.

We show this statement by proving that each term in brackets in Eq. (A.1) is bounded above by $v_{T+1,\alpha}(n^R, n^S + 1) - v_{T+1,\alpha}(n^R, n^S + 2)$. Then, we can be sure that the statement will hold as the coefficients of these terms sum up to 1.

Let us first look into the three parts of (A.1) that do not relate to the rewards.

(a) $v_{T,\alpha}(n^R - 1, n^S) - v_{T,\alpha}(n^R - 1, n^S + 1) \leq v_{T,\alpha}(n^R, n^S) - v_{T,\alpha}(n^R, n^S + 1)$ by the induction hypothesis of statement (II), and $v_{T,\alpha}(n^R, n^S) - v_{T,\alpha}(n^R, n^S + 1) \leq v_{T,\alpha}(n^R, n^S + 1) - v_{T,\alpha}(n^R, n^S + 2)$, by the induction hypothesis of statement (I) and we know that $v_{T,\alpha}(n^R, n^S + 1) - v_{T,\alpha}(n^R, n^S + 2) \leq v_{T+1,\alpha}(n^R, n^S + 1) - v_{T+1,\alpha}(n^R, n^S + 2)$ by Lemma A.1.

(b) $v_{T,\alpha}(n^R, n^S - 1) - v_{T,\alpha}(n^R, n^S) \le v_{T,\alpha}(n^R, n^S) - v_{T,\alpha}(n^R, n^S + 1) \le v_{T,\alpha}(n^R, n^S + 1) - v_{T,\alpha}(n^R, n^S + 2)$, by the induction hypothesis of statement (I) and $v_{T,\alpha}(n^R, n^S + 1) - v_{T,\alpha}(n^R, n^S + 2) \le v_{T+1,\alpha}(n^R, n^S + 1) - v_{T+1,\alpha}(n^R, n^S + 2)$ by Lemma A.1.

(c) $v_{T,\alpha}(n^R, n^S) - v_{T,\alpha}(n^R, n^S + 1) \le v_{T,\alpha}(n^R, n^S + 1) - v_{T,\alpha}(n^R, n^S + 2)$, by the induction hypothesis of statement (I) and $v_{T,\alpha}(n^R, n^S + 1) - v_{T,\alpha}(n^R, n^S + 2) \le v_{T+1,\alpha}(n^R, n^S + 1) - v_{T+1,\alpha}(n^R, n^S + 2)$ by Lemma A.1.

Now, we check the reward differences. $R_\alpha v_{T,\alpha}(n^R, n^S) - R_\alpha v_{T,\alpha}(n^R, n^S + 1)$ is affected by the admission and routing choices made. We look into all possible combinations of these decisions. We show that, for any arbitrary customer class i, $R_\alpha^i v_{T,\alpha}(n^R, n^S) - R_\alpha^i v_{T,\alpha}(n^R, n^S + 1) = \max[R_i^R + \alpha v_{T,\alpha}(n^R + 1, n^S), R_i^S + \alpha v_{T,\alpha}(n^R, n^S + 1), \alpha(n^R, n^S)] - \max[R_i^R + \alpha v_{T,\alpha}(n^R + 1, n^S + 1), R_i^S + \alpha v_{T,\alpha}(n^R, n^S + 2), \alpha(n^R, n^S + 1)]$ is bounded above by $v_{T+1,\alpha}(n^R, n^S + 1) - v_{T+1,\alpha}(n^R, n^S + 2)$.

(a) Consider that it is optimal to route class i to the regular station at both (n^R, n^S) and $(n^R, n^S + 1)$ states. Then, we know that $R_\alpha^i v_{T,\alpha}(n^R, n^S) - R_\alpha^i v_{T,\alpha}(n^R, n^S + 1) = \alpha[v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 1, n^S + 1)]$. By using the induction hypothesis of statement (III), $\alpha[v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 1, n^S + 1)] \le \alpha[v_{T,\alpha}(n^R, n^S + 1) - v_{T,\alpha}(n^R, n^S + 2)]$. We know that with Lemma A.1 that $\alpha[v_{T,\alpha}(n^R, n^S + 1) - v_{T,\alpha}(n^R, n^S + 2)] \le v_{T+1,\alpha}(n^R, n^S + 1) - v_{T,\alpha}(n^R, n^S + 2)$.

(b) Consider that it is optimal to route class i to the regular station at (n^R, n^S) but to the self-service station at $(n^R, n^S + 1)$. We know that, as it is not optimal to route class i to the regular station at $(n^R, n^S + 1)$, $R_\alpha^i v_{T,\alpha}(n^R, n^S) - R_\alpha^i v_{T,\alpha}(n^R, n^S + 1) \le \alpha[v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 1, n^S + 1)]$. For the rest, we can follow (d).

(c) Similarly, for the case that it is optimal to route class i to the regular station at (n^R, n^S) but to reject at $(n^R, n^S + 1)$, we can infer that $R_\alpha^i v_{T,\alpha}(n^R, n^S) - R_\alpha^i v_{T,\alpha}(n^R, n^S + 1) \le \alpha[v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 1, n^S + 1)]$ and follow the lines in (d).

(d) Consider that it is optimal to route class i to the self-service station at both (n^R, n^S) and $(n^R, n^S + 1)$ states. Thus, $R_\alpha^i v_{T,\alpha}(n^R, n^S) - R_\alpha^i v_{T,\alpha}(n^R, n^S + 1) = \alpha[v_{T,\alpha}(n^R, n^S + 1) - v_{T,\alpha}(n^R, n^S + 2)]$. With Lemma A.1, we can show that $\alpha[v_{T,\alpha}(n^R, n^S + 1) - v_{T,\alpha}(n^R, n^S + 2)] \le v_{T+1,\alpha}(n^R, n^S + 1) - v_{T,\alpha}(n^R, n^S + 2)$.

(e) Consider that it is optimal to route class i to the self-service station at (n^R, n^S) but to the regular station at $(n^R, n^S + 1)$. As it is not optimal to route class i to the self-service station at $(n^R, n^S + 1)$, $R_\alpha^i v_{T,\alpha}(n^R, n^S) - R_\alpha^i v_{T,\alpha}(n^R, n^S + 1) \le \alpha[v_{T,\alpha}(n^R, n^S + 1) - v_{T,\alpha}(n^R, n^S + 2)]$. Then, we can follow (h) to show the rest.

(f) Consider that it is optimal to route class i to the self-service station at (n^R, n^S) but to reject at $(n^R, n^S + 1)$. We can follow the arguments in (i).

(g) Consider that it is optimal to reject class i at both (n^R, n^S) and $(n^R, n^S + 1)$ states. Then, $R_\alpha^i v_{T,\alpha}(n^R, n^S) - R_\alpha^i v_{T,\alpha}(n^R, n^S + 1) = \alpha[v_{T,\alpha}(n^R, n^S) - v_{T,\alpha}(n^R, n^S + 1)]$. With the induction hypothesis of statement (I), $\alpha[v_{T,\alpha}(n^R, n^S) - v_{T,\alpha}(n^R, n^S + 1)] \le \alpha[v_{T,\alpha}(n^R, n^S + 1) - v_{T,\alpha}(n^R, n^S + 2)]$. Then, Lemma A.1 completes the proof.

The case that class i is rejected at state (n^R, n^S) but routed to the self-service station at state $(n^R, n^S + 1)$ is impossible due to the induction hypothesis of statement (I). Likewise, it is also impossible that class i is rejected at state (n^R, n^S) but routed to the regular station at state $(n^R, n^S + 1)$, due to the induction hypothesis of statement (II).

Proving Statement (II): This is the supermodularity property of the value functions. We show this with induction on T by sample path arguments. Consider four processes on the same probability space such that all have $T + 1$ periods remaining. Let Process 1 and Process 4 start at

(n^R, n^S) and $(n^R + 1, n^S + 1)$, respectively, and assume that both processes use an optimal policy π^*. On the other hand, let Process 2 and Process 3 start at states $(n^R, n^S + 1)$ and $(n^R + 1, n^S)$, respectively, and let them follow (potentially) suboptimal policies. However, we let that these policies will only deviate from the optimal policy π^* during the first time period.

We let R_k and R_k^* be the random variables denoting the (net) rewards obtained by the policies that Process $k \in \{1, 2, 3, 4\}$ follows and the rewards that could be obtained if Process k was following an optimal policy instead, respectively. In order to show that $\mathbb{E}(R_1^*) - \mathbb{E}(R_2^*) \leq \mathbb{E}(R_3^*) - \mathbb{E}(R_4^*)$, it is sufficient to show $\mathbb{E}(R_1^*) - \mathbb{E}(R_2) \leq \mathbb{E}(R_3) - \mathbb{E}(R_4^*)$ as $\mathbb{E}(R_1^*) - \mathbb{E}(R_2^*) \leq \mathbb{E}(R_1^*) - \mathbb{E}(R_2) \leq \mathbb{E}(R_3) - \mathbb{E}(R_4^*) \leq \mathbb{E}(R_3^*) - \mathbb{E}(R_4^*)$.

We now condition on the possible events that might occur in the first time period, by using the fact that after this event, we have T periods left in the horizon. The first event partitions the state space. By using the law of total expectation, it suffices to show that $\mathbb{E}(R_1^* - R_2 | A_n) \leq \mathbb{E}(R_3 - R_4^* | A_n)$ for any transition event A_n. We skip writing the holding costs incurred at the states (during the first time period) as they cancel out in $\mathbb{E}(R_1^* - R_2 | A_n)$ and $\mathbb{E}(R_3 - R_4^* | A_n)$ irrespective of the transition events (A_ns).

First focus on arrival events. Let class i be an arbitrary customer class whose arrival we observe in the first time period. Let A_1 denote this arrival event. As we do not know the decisions that the optimal policy π^* will take after observing this event at states (n^R, n^S) and $(n^R + 1, n^S + 1)$, below we consider all possible scenarios for the decisions that the optimal policy can take at these states.

Scenario 1 for A_1: Assume that the optimal policy routes the arrived customer of class i to the regular station at both (n^R, n^S) and $(n^R + 1, n^S + 1)$ states. Let us consider that Process 2 and Process 3 also route this class to the regular station. Then, $\mathbb{E}(R_1^* - R_2 | A_1) = (R_i^R + \alpha v_{T,\alpha}(n^R + 1, n^S) - R_i^R - \alpha v_{T,\alpha}(n^R + 1, n^S + 1)) = \alpha[v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 1, n^S + 1)]$. By the induction hypothesis of statement (II), we know that $\alpha[v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 1, n^S + 1)] \leq \alpha[v_{T,\alpha}(n^R + 2, n^S) - v_{T,\alpha}(n^R + 2, n^S + 1)]$, where $\alpha[v_{T,\alpha}(n^R + 2, n^S) - v_{T,\alpha}(n^R + 2, n^S + 1)] = (R_i^R + \alpha v_{T,\alpha}(n^R + 2, n^S) - R_i^R - \alpha v_{T,\alpha}(n^R + 2, n^S + 1)) = \mathbb{E}(R_3 - R_4^* | A_1)$.

Scenario 2 for A_1: Assume that the optimal policy routes the arrived customer of class i to the regular station at (n^R, n^S) and to the self-service at $(n^R + 1, n^S + 1)$. Let Process 2 and Process 3 mimic Process 1 and Process 4, respectively. Then, $\mathbb{E}(R_1^* - R_2 | A_1) = (R_i^R + \alpha v_{T,\alpha}(n^R + 1, n^S) - R_i^R - \alpha v_{T,\alpha}(n^R + 1, n^S + 1)) = \alpha[v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 1, n^S + 1)]$. By the induction hypothesis of statement (I), we know that $\alpha[v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 1, n^S + 1)] \leq \alpha[v_{T,\alpha}(n^R + 1, n^S + 1) - v_{T,\alpha}(n^R + 1, n^S + 2)]$, where $\alpha[v_{T,\alpha}(n^R + 1, n^S + 1) - v_{T,\alpha}(n^R + 1, n^S + 2)] = (R_i^S + \alpha v_{T,\alpha}(n^R + 1, n^S + 1) - R_i^S - \alpha v_{T,\alpha}(n^R + 1, n^S + 2)) = \mathbb{E}(R_3 - R_4^* | A_1)$.

Scenario 3 for A_1: Assume that the optimal policy routes the arrived customer of class i to the regular station at (n^R, n^S) and rejects at $(n^R + 1, n^S + 1)$. Let Process 2 and Process 3 mimic Process 1 and Process 4, respectively. Then, $\mathbb{E}(R_1^* - R_2 | A_1) = (R_i^R + \alpha v_{T,\alpha}(n^R + 1, n^S) - R_i^R - \alpha v_{T,\alpha}(n^R + 1, n^S + 1)) = \alpha[v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 1, n^S + 1)] = \mathbb{E}(R_3 - R_4^* | A_1)$.

Scenario 4 for A_1: Assume that the optimal policy routes the arrived customer of class i to the self-service station at both (n^R, n^S) and $(n^R + 1, n^S + 1)$ states. Let us consider that also Process 2 and Process 3 route this class to the self-service station. Then, $\mathbb{E}(R_1^* - R_2 | A_1) = (R_i^S + \alpha v_{T,\alpha}(n^R, n^S + 1) - R_i^S - \alpha v_{T,\alpha}(n^R, n^S + 2)) = \alpha[v_{T,\alpha}(n^R, n^S + 1) - v_{T,\alpha}(n^R, n^S + 2)] \leq (R_i^S + \alpha v_{T,\alpha}(n^R + 1, n^S + 1) - R_i^S - \alpha v_{T,\alpha}(n^R + 1, n^S + 2)) = \mathbb{E}(R_3 - R_4^* | A_1)$, as by the induction hypothesis of statement (I).

Scenario 5 for A_1: Assume that the optimal policy routes the arrived customer of class i to the self-service station at (n^R, n^S) and to the regular station at $(n^R + 1, n^S + 1)$. Let Process 2 to mimic Process 4 and Process 3 to mimic Process 1. Then, $\mathbb{E}(R_1^* - R_2 | A_1) = (R_i^S + \alpha v_{T,\alpha}(n^R, n^S + 1) - R_i^R - \alpha v_{T,\alpha}(n^R + 1, n^S + 1)) \leq (R_i^S + \alpha v_{T,\alpha}(n^R + 1, n^S + 1) - R_i^R - \alpha v_{T,\alpha}(n^R + 2, n^S + 1))$ by the induction hypothesis of statement (II), where $(R_i^S + \alpha v_{T,\alpha}(n^R + 1, n^S + 1) - R_i^R - \alpha v_{T,\alpha}(n^R + 2, n^S + 1)) = \mathbb{E}(R_3 - R_4^* | A_1)$.

Scenario 6 for A_1: Assume that the optimal policy routes the arrived customer of class i to the self-service station at (n^R, n^S) and rejects at $(n^R + 1, n^S + 1)$. Let Process 2 to mimic Process 4 and Process 3 to mimic Process 1. Then, $\mathbb{E}(R_1^* - R_2|A_1) = (R_i^S + \alpha v_{T,\alpha}(n^R, n^S + 1) - \alpha v_{T,\alpha}(n^R, n^S + 1)) = R_i^S = \mathbb{E}(R_3 - R_4^*|A_1)$.

Scenario 7 for A_1: Assume that the optimal policy rejects class i at both (n^R, n^S) and $(n^R + 1, n^S + 1)$ states. Then, let Process 2 and Process 3 reject as well. Then, $\mathbb{E}(R_1^* - R_2|A_1) = \alpha v_{T,\alpha}(n^R, n^S) - \alpha v_{T,\alpha}(n^R, n^S + 1) \leq \alpha v_{T,\alpha}(n^R + 1, n^S) - \alpha v_{T,\alpha}(n^R + 1, n^S + 1) = \mathbb{E}(R_3 - R_4^*|A_1)$, by the induction hypothesis of statement (II).

We realize that it is impossible that an optimal policy rejects class i at (n^R, n^S) but accepts either to the self-service or regular stations at $(n^R + 1, n^S + 1)$, by the induction hypothesis.

In all decision scenarios possible after observing event A_1, we see that we can confirm $\mathbb{E}(R_1^* - R_2|A_1) \leq \mathbb{E}(R_3 - R_4^*|A_1)$.

Now, we focus on the service completion events. It is possible that these types of events occur in the first time period. The effect of service completion events are not related to the decisions of the optimal policy; however, this time we need to consider scenarios for the states of the systems, due to the fact that we have a bounded state space.

Let us first consider the service completion event from the regular service station. Let A_2 denote this event.

Scenario 1 for A_2: First, let us consider the case that $n^R > 0$ such that all processes can observe this event. Then, $\mathbb{E}(R_1^* - R_2|A_2) = \alpha v_{T,\alpha}(n^R - 1, n^S) - \alpha v_{T,\alpha}(n^R - 1, n^S + 1) \leq \alpha v_{T,\alpha}(n^R, n^S) - \alpha v_{T,\alpha}(n^R, n^S + 1)$ by the induction hypothesis of statement (II), where $\alpha v_{T,\alpha}(n^R, n^S) - \alpha v_{T,\alpha}(n^R, n^S + 1) = \mathbb{E}(R_3 - R_4^*|A_2)$.

Scenario 2 for A_2: Now consider that $n^R = 0$ such that we can observe this event only in Process 3 and Process 4. For this case, we have $\mathbb{E}(R_1^* - R_2|A_2) = \alpha v_{T,\alpha}(n^R, n^S) - \alpha v_{T,\alpha}(n^R, n^S + 1) = \mathbb{E}(R_3 - R_4^*|A_2)$.

Now, let us consider the service completion event from the self-service station. Let A_3 denote this event.

Scenario 1 for A_3: First, let us consider the case that $n^S > 0$ such that all processes can observe this event. Then, $\mathbb{E}(R_1^* - R_2|A_3) = \alpha v_{T,\alpha}(n^R, n^S - 1) - \alpha v_{T,\alpha}(n^R, n^S) \leq \alpha v_{T,\alpha}(n^R + 1, n^S - 1) - \alpha v_{T,\alpha}(n^R + 1, n^S)$, by the induction hypothesis of statement (II), where $\alpha v_{T,\alpha}(n^R + 1, n^S - 1) - \alpha v_{T,\alpha}(n^R + 1, n^S) = \mathbb{E}(R_3 - R_4^*|A_3)$.

Scenario 2 for A_3: Now consider that $n^S = 0$ such that only Process 2 and Process 4 can observe this event. For this case, we have $\mathbb{E}(R_1^* - R_2|A_3) = \alpha v_{T,\alpha}(n^R, n^S) - \alpha v_{T,\alpha}(n^R, n^S) = 0 = \alpha v_{T,\alpha}(n^R + 1, n^S) - \alpha v_{T,\alpha}(n^R + 1, n^S)\mathbb{E}(R_3 - R_4^*|A_3)$.

We note that a dummy transition event might also occur due to uniformization. However, this will not change the state of any processes. So, the induction hypothesis of statement (II) is sufficient in this case.

Proving Statement (III): This is a superconvexity property of the value functions. We again show this with induction on T by sample path arguments. Consider four processes on the same probability space such that all have $T + 1$ periods remaining. Let Process 1 and Process 4 start at $(n^R + 1, n^S)$ and $(n^R, n^S + 2)$, respectively, and assume that both processes use an optimal policy π^*. On the other hand, let Process 2 and Process 3 start at $(n^R + 1, n^S + 1)$ and $(n^R, n^S + 1)$. For these processes, we do not assume that they follow an optimal policy. However, we consider that after the first time period, all policies follow an optimal policy.

We again let R_k and R_k^* be the random variables denoting the (net) rewards obtained by the policies that Process $k \in \{1, 2, 3, 4\}$ follows and the rewards that could be obtained if Process k was following an optimal policy instead, respectively. In order to show that $\mathbb{E}(R_1^*) - \mathbb{E}(R_2^*) \leq \mathbb{E}(R_3^*) - \mathbb{E}(R_4^*)$, it is sufficient to show $\mathbb{E}(R_1^*) - \mathbb{E}(R_2) \leq \mathbb{E}(R_3) - \mathbb{E}(R_4^*)$ as $\mathbb{E}(R_1^*) - \mathbb{E}(R_2^*) \leq \mathbb{E}(R_1^*) - \mathbb{E}(R_2) \leq \mathbb{E}(R_3) - \mathbb{E}(R_4^*) \leq \mathbb{E}(R_3^*) - \mathbb{E}(R_4^*)$.

We now condition on the possible events that might occur in the first time period, by using the fact that after this event, we have T periods left in the horizon. The first event partitions the state space. By using the law of total expectation, it suffices to show that $\mathbb{E}(R_1^* - R_2|A_n) \leq \mathbb{E}(R_3 - R_4^*|A_n)$ for any transition event A_n. We skip writing the holding costs incurred at the states

(during the first time period) as they cancel out in $\mathbb{E}(R_1^* - R_2|A_n)$ and $\mathbb{E}(R_3 - R_4^*|A_n)$ irrespective of the transition events (A_ns).

First focus on arrival events. Let class i be an arbitrary customer class whose arrival we observe in the first time period. Let A_1 denote this arrival event. As we do not know the decisions that the optimal policy π^* will take after observing this event at states $(n^R + 1, n^S)$ and $(n^R, n^S + 2)$, below we consider all possible scenarios for the decisions that the optimal policy can take at these states.

Scenario 1 for A_1: Assume that the optimal policy routes the arrived customer of class i to the regular station at both $(n^R + 1, n^S)$ and $(n^R, n^S + 2)$ states. Let us consider that also Process 2 and Process 3 route this class to the regular station. Then, $\mathbb{E}(R_1^* - R_2|A_1) = (R_i^R + \alpha v_{T,\alpha}(n^R + 2, n^S) - R_i^R - \alpha v_{T,\alpha}(n^R + 2, n^S + 1)) = \alpha[v_{T,\alpha}(n^R + 2, n^S) - v_{T,\alpha}(n^R + 2, n^S + 1)]$. By the induction hypothesis of statement (III), we know that $\alpha[v_{T,\alpha}(n^R + 2, n^S) - v_{T,\alpha}(n^R + 2, n^S + 1)] \leq \alpha[v_{T,\alpha}(n^R + 1, n^S + 1) - v_{T,\alpha}(n^R + 1, n^S + 2)]$, where $\alpha[v_{T,\alpha}(n^R + 1, n^S + 1) - v_{T,\alpha}(n^R + 1, n^S + 2)] = (R_i^R + \alpha v_{T,\alpha}(n^R + 1, n^S + 1) - R_i^R - \alpha v_{T,\alpha}(n^R + 1, n^S + 2)) = \mathbb{E}(R_3 - R_4^*|A_1)$.

Scenario 2 for A_1: Assume that the optimal policy routes the arrived customer of class i to the regular station at $(n^R + 1, n^S)$ but to the self-service station at $(n^R, n^S + 2)$. Let us consider that Process 2 and Process 3 mimic Process 1 and Process 4, respectively. Then, $\mathbb{E}(R_1^* - R_2|A_1) = \alpha[v_{T,\alpha}(n^R + 2, n^S) - v_{T,\alpha}(n^R + 2, n^S + 1)]$. By the induction hypothesis of statement (III), $\alpha[v_{T,\alpha}(n^R + 2, n^S) - v_{T,\alpha}(n^R + 2, n^S + 1)] \leq \alpha[v_{T,\alpha}(n^R + 1, n^S + 1) - v_{T,\alpha}(n^R + 1, n^S + 2)] \leq \alpha[v_{T,\alpha}(n^R, n^S + 2) - v_{T,\alpha}(n^R, n^S + 3)]$ and $\alpha[v_{T,\alpha}(n^R, n^S + 2) - v_{T,\alpha}(n^R, n^S + 3)] = \mathbb{E}(R_3 - R_4^*|A_1)$.

Scenario 3 for A_1: Assume that the optimal policy routes the arrived customer of class i to the regular station at $(n^R + 1, n^S)$ but rejects at $(n^R, n^S + 2)$. Let Process 2 mimic Process 4 and Process 3 mimic Process 1. Then, $\mathbb{E}(R_1^* - R_2|A_1) = R_i^R + \alpha[v_{T,\alpha}(n^R + 2, n^S) - v_{T,\alpha}(n^R + 1, n^S + 1)]$. By the induction hypothesis of statement (IV), $R_i^R + \alpha[v_{T,\alpha}(n^R + 2, n^S) - v_{T,\alpha}(n^R + 1, n^S + 1)] \leq R_i^R + \alpha[v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R, n^S + 1)]$. By the induction hypothesis of statement (III), $R_i^R + \alpha[v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R, n^S + 1)] \leq R_i^R + \alpha[v_{T,\alpha}(n^R + 1, n^S + 1) - v_{T,\alpha}(n^R, n^S + 2)]$, where $R_i^R + \alpha[v_{T,\alpha}(n^R + 1, n^S + 1) - v_{T,\alpha}(n^R, n^S + 2)] \leq R_i^R + \alpha[v_{T,\alpha}(n^R + 1, n^S + 1) - v_{T,\alpha}(n^R, n^S + 2)] = \mathbb{E}(R_3 - R_4^*|A_1)$.

Scenario 4 for A_1: Assume that the optimal policy routes the arrived customer of class i to the self-service station at both $(n^R + 1, n^S)$ and $(n^R, n^S + 2)$ states. Let us consider that also Process 2 and Process 3 route this class to the self-service station. Then, $\mathbb{E}(R_1^* - R_2|A_1) = \alpha[v_{T,\alpha}(n^R + 1, n^S + 1) - v_{T,\alpha}(n^R + 1, n^S + 2)]$. By the induction hypothesis of statement (III), $\alpha[v_{T,\alpha}(n^R + 1, n^S + 1) - v_{T,\alpha}(n^R + 1, n^S + 2)] \leq \alpha[v_{T,\alpha}(n^R, n^S + 2) - v_{T,\alpha}(n^R+, n^S + 3)]$, where $\alpha[v_{T,\alpha}(n^R, n^S + 2) - v_{T,\alpha}(n^R+, n^S + 3)] = \mathbb{E}(R_3 - R_4^*|A_1)$.

Scenario 5 for A_1: Assume that the optimal policy routes the arrived customer of class i to the self-service station at $(n^R + 1, n^S)$ but to the regular station at $(n^R, n^S + 2)$. Let us consider that Process 2 and Process 3 mimic Process 1 and Process 4, respectively. Then, $\mathbb{E}(R_1^* - R_2|A_1) = \alpha[v_{T,\alpha}(n^R + 1, n^S + 1) - v_{T,\alpha}(n^R + 1, n^S + 2)]$, where $\alpha[v_{T,\alpha}(n^R + 1, n^S + 1) - v_{T,\alpha}(n^R + 1, n^S + 2)] = \mathbb{E}(R_3 - R_4^*|A_1)$.

Scenario 6 for A_1: Assume that the optimal policy routes the arrived customer of class i to the self-service station at $(n^R + 1, n^S)$ but rejects at $(n^R, n^S + 2)$. Let us consider that Process 2 and Process 3 mimic Process 4 and Process 2, respectively. Then, $\mathbb{E}(R_1^* - R_2|A_1) = R_i^S + \alpha[v_{T,\alpha}(n^R + 1, n^S + 1) - v_{T,\alpha}(n^R + 1, n^S + 1)] = R_i^S$, where $R_i^S = R_i^S + \alpha[v_{T,\alpha}(n^R, n^S + 2) - v_{T,\alpha}(n^R, n^S + 2)] = \mathbb{E}(R_3 - R_4^*|A_1)$.

Scenario 7 for A_1: Assume that the optimal policy rejects the arrived customer of class i at both $(n^R + 1, n^S)$ and $(n^R, n^S + 2)$ states. Let Process 2 and Process 3 reject as well. Then, $\mathbb{E}(R_1^* - R_2|A_1) = \alpha[v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 1, n^S + 1)]$, where $\alpha[v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 1, n^S + 1)] \leq \alpha[v_{T,\alpha}(n^R, n^S + 1) - v_{T,\alpha}(n^R, n^S + 2)] = \mathbb{E}(R_3 - R_4^*|A_1)$ holds due to the induction hypothesis of statement (III).

Scenario 8 for A_1: Assume that the optimal policy rejects the arrived class i at $(n^R + 1, n^S)$ and routes to the regular station at $(n^R, n^S + 2)$. Let us consider that Process 2 and Process 3 mimic Process 1 and Process 4, respectively. Then $\mathbb{E}(R_1^* - R_2 | A_1) = \alpha[v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 1, n^S + 1)]$, by the induction hypothesis of statement (I), $\alpha[v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 1, n^S + 1)] \leq \alpha[v_{T,\alpha}(n^R + 1, n^S + 1) - v_{T,\alpha}(n^R + 1, n^S + 2)]$, where $\alpha[v_{T,\alpha}(n^R + 1, n^S + 1) - v_{T,\alpha}(n^R + 1, n^S + 2)] = \mathbb{E}(R_3 - R_4^* | A_1)$.

By the induction hypothesis, it is impossible that an optimal policy routes class i to the self-service at $(n^R, n^S + 2)$ but rejects at $(n^R + 1, n^S)$. If it is optimal to not to route class i to the self-service at $(n^R + 1, n^S)$, then $R_i^S < \alpha[v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 1, n^S + 1)]$. By the induction hypothesis of statement (III), $\alpha[v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 1, n^S + 1)] \leq \alpha[v_{T,\alpha}(n^R, n^S + 1) - v_{T,\alpha}(n^R, n^S + 2)]$ and by the induction hypothesis of statement (I), $\alpha[v_{T,\alpha}(n^R, n^S + 1) - v_{T,\alpha}(n^R, n^S + 2)] \leq \alpha[v_{T,\alpha}(n^R, n^S + 2) - v_{T,\alpha}(n^R, n^S + 3)]$. Then, it is impossible that $R_i^S + \alpha v_{T,\alpha}(n^R, n^S + 3) > \alpha v_{T,\alpha}(n^R, n^S + 2)$.

In all decision scenarios possible after observing event A_1, we see that we can confirm $\mathbb{E}(R_1^* - R_2 | A_1) \leq \mathbb{E}(R_3 - R_4^* | A_1)$.

Now let us look into the case service completion events occur in the first time period. Let us first consider the service completion event from the regular service station. Let A_2 denote this event.

Scenario 1 for A_2: First, let us consider the case that $n^R > 0$ such that all processes can observe this event. Then, $\mathbb{E}(R_1^* - R_2 | A_2) = \alpha[v_{T,\alpha}(n^R, n^S) - v_{T,\alpha}(n^R, n^S + 1)] \leq \alpha[v_{T,\alpha}(n^R - 1, n^S + 1) - v_{T,\alpha}(n^R - 1, n^S + 2)] = \mathbb{E}(R_3 - R_4^* | A_2)$, by the induction hypothesis of statement (III).

Scenario 2 for A_2: Now consider that $n^R = 0$ such that only Process 1 and Process 2 can observe this event. Then, we will have that $\mathbb{E}(R_1^* - R_2 | A_2) = \alpha[v_{T,\alpha}(n^R, n^S) - v_{T,\alpha}(n^R, n^S + 1)]$ where $\alpha[v_{T,\alpha}(n^R, n^S) - v_{T,\alpha}(n^R, n^S + 1)] \leq \alpha[v_{T,\alpha}(n^R, n^S + 1) - v_{T,\alpha}(n^R, n^S + 2)] = \mathbb{E}(R_3 - R_4^* | A_2)$, by the induction hypothesis of statement (I).

Let us now consider the service completion event from the self-service station. Let A_3 denote this event.

Scenario 1 for A_3: First, let us consider the case that $n^S > 0$ such that all processes can observe this event. Then, $\mathbb{E}(R_1^* - R_2 | A_3) = \alpha[v_{T,\alpha}(n^R + 1, n^S - 1) - v_{T,\alpha}(n^R + 1, n^S)] \leq \alpha[v_{T,\alpha}(n^R, n^S) - v_{T,\alpha}(n^R, n^S + 1)] = \mathbb{E}(R_3 - R_4^* | A_3)$, by the induction hypothesis of statement (III).

Scenario 2 for A_3: Now consider the case that $n^S = 0$. For this case, Process 1 can not observe this event. We have that $\mathbb{E}(R_1^* - R_2 | A_3) = \alpha[v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 1, n^S)] = 0 \leq \alpha[v_{T,\alpha}(n^R, n^S) - v_{T,\alpha}(n^R, n^S + 1)] = \mathbb{E}(R_3 - R_4^* | A_3)$, by Lemma 3.1.

Lastly, a dummy transition event might occur due to uniformization, which does not change the state of any processes. So, the induction hypothesis of statement (III) is sufficient in this case.

Proving Statement (IV): This is a superconvexity property of the value functions. We again show this with induction on T by sample path arguments. Consider four processes on the same probability space such that all have $T + 1$ periods remaining. Let Process 1 and Process 4 start at $(n^R, n^S + 1)$ and $(n^R + 2, n^S)$, respectively, and assume that both processes use an optimal policy π^*. On the other hand, let Process 2 and Process 3 start at $(n^R + 1, n^S + 1)$ and $(n^R + 1, n^S)$. For these processes, we do not assume that they follow an optimal policy. However, we consider that after the first time period, all policies follow an optimal policy.

We again let R_k and R_k^* be the random variables denoting the (net) rewards obtained by the policies that Process $k \in \{1, 2, 3, 4\}$ follows and the rewards that could be obtained if Process k was following an optimal policy instead, respectively. In order to show that $\mathbb{E}(R_1^*) - \mathbb{E}(R_2^*) \leq \mathbb{E}(R_3^*) - \mathbb{E}(R_4^*)$, it is sufficient to show $\mathbb{E}(R_1^*) - \mathbb{E}(R_2) \leq \mathbb{E}(R_3) - \mathbb{E}(R_4^*)$ as $\mathbb{E}(R_1^*) - \mathbb{E}(R_2^*) \leq \mathbb{E}(R_1^*) - \mathbb{E}(R_2) \leq \mathbb{E}(R_3) - \mathbb{E}(R_4^*) \leq \mathbb{E}(R_3^*) - \mathbb{E}(R_4^*)$.

We now condition on the possible events that might occur in the first time period, by using the fact that after this event, we have T periods left in the horizon. The first event partitions the state space. By using the law of total expectation, it suffices to show that $\mathbb{E}(R_1^* - R_2 | A_n) \leq$

$\mathbb{E}(R_3 - R_4^*|A_n)$ for any transition event A_n. We skip writing the holding costs incurred at the states (during the first time period) as they cancel out in $\mathbb{E}(R_1^* - R_2|A_n)$ and $\mathbb{E}(R_3 - R_4^*|A_n)$ irrespective of the transition events (A_ns).

First focus on arrival events. Let class i be an arbitrary customer class whose arrival we observe in the first time period. Let A_1 denote this arrival event. As we do not know the decisions that the optimal policy π^* will take after observing this event at states $(n^R, n^S + 1)$ and $(n^R + 2, n^S)$, below we consider all possible scenarios for the decisions that the optimal policy can take at these states.

Scenario 1 for A_1: Assume that the optimal policy routes the arrived customer of class i to the regular station at both $(n^R, n^S + 1)$ and $(n^R + 2, n^S)$ states. Let us consider that also Process 2 and Process 3 route this class to the regular station. Then, $\mathbb{E}(R_1^* - R_2|A_1) = (R_i^R + \alpha v_{T,\alpha}(n^R + 1, n^S + 1) - R_i^R - \alpha v_{T,\alpha}(n^R + 2, n^S + 1))$. By the induction hypothesis of statement (IV), we know that $(R_i^R + \alpha v_{T,\alpha}(n^R + 1, n^S + 1) - R_i^R - \alpha v_{T,\alpha}(n^R + 2, n^S + 1)) \leq (R_i^R + \alpha v_{T,\alpha}(n^R + 2, n^S) - R_i^R - \alpha v_{T,\alpha}(n^R + 3, n^S)) = \mathbb{E}(R_3 - R_4^*|A_1)$.

Scenario 2 for A_1: Assume that the optimal policy routes the arrived customer of class i to the regular station at $(n^R, n^S + 1)$ and to the self-service at $(n^R + 2, n^S)$ state. Let us consider that Process 2 and Process 3 mimic Process 1 and Process 4, respectively. Then, $\mathbb{E}(R_1^* - R_2|A_1) = (R_i^R + \alpha v_{T,\alpha}(n^R + 1, n^S + 1) - R_i^R - \alpha v_{T,\alpha}(n^R + 2, n^S + 1))$. We have that $R_i^R \geq R_i^R$, so $R_1^* - R_2^* \leq \alpha v_{T,\alpha}(n^R + 1, n^S + 1) - R_i^R - \alpha v_{T,\alpha}(n^R + 2, n^S + 1) = \mathbb{E}(R_3 - R_4^*|A_1)$.

Scenario 3 for A_1: Assume that the optimal policy routes the arrived customer of class i to the regular station at $(n^R, n^S + 1)$ and rejects at $(n^R + 2, n^S)$ state. Let us consider that Process 2 and Process 3 mimic Process 4 and Process 1, respectively. Then, $\mathbb{E}(R_1^* - R_2|A_1) = (R_i^R + \alpha v_{T,\alpha}(n^R + 1, n^S + 1) - \alpha v_{T,\alpha}(n^R + 1, n^S + 1))$. As we assume that R_i^R is non-negative always, $R_1^* - R_2^* \leq \alpha v_{T,\alpha}(n^R + 1, n^S + 1) - \alpha v_{T,\alpha}(n^R + 1, n^S + 1)$ and by the induction hypothesis of statement (IV), $\alpha v_{T,\alpha}(n^R + 1, n^S + 1) - \alpha v_{T,\alpha}(n^R + 1, n^S + 1) \leq \alpha v_{T,\alpha}(n^R + 2, n^S) - \alpha v_{T,\alpha}(n^R + 2, n^S + 1)$, where $\alpha v_{T,\alpha}(n^R + 2, n^S) - \alpha v_{T,\alpha}(n^R + 2, n^S + 1) = \mathbb{E}(R_3 - R_4^*|A_1)$.

Scenario 4 for A_1: Assume that the optimal policy routes the arrived customer of class i to the self-service station at both $(n^R, n^S + 1)$ and $(n^R + 2, n^S)$ states. Let us consider that also Process 2 and Process 3 route this class to the self-service station. Then, $\mathbb{E}(R_1^* - R_2|A_1) = (R_i^S + \alpha v_{T,\alpha}(n^R, n^S + 2) - R_i^S - \alpha v_{T,\alpha}(n^R + 1, n^S + 2))$. By the induction hypothesis of statement (IV), $\alpha v_{T,\alpha}(n^R, n^S + 2) - \alpha v_{T,\alpha}(n^R + 1, n^S + 2) \leq \alpha v_{T,\alpha}(n^R + 1, n^S + 1) - \alpha v_{T,\alpha}(n^R + 2, n^S + 1)$, where $\alpha v_{T,\alpha}(n^R + 1, n^S + 1) - \alpha v_{T,\alpha}(n^R + 2, n^S + 1) = \mathbb{E}(R_3 - R_4^*|A_1)$.

Scenario 5 for A_1: Assume that the optimal policy routes the arrived customer of class i to the self-service station at $(n^R, n^S + 1)$ and to the regular station at $(n^R + 2, n^S)$. Let us consider that Process 2 and Process 3 mimic Process 1 and Process 4, respectively. Then, $\mathbb{E}(R_1^* - R_2|A_1) = \alpha v_{T,\alpha}(n^R, n^S + 2) - \alpha v_{T,\alpha}(n^R + 1, n^S + 2))$. By the induction hypothesis of statement (IV), $\alpha v_{T,\alpha}(n^R, n^S + 2) - \alpha v_{T,\alpha}(n^R + 1, n^S + 2) \leq \alpha v_{T,\alpha}(n^R + 1, n^S + 1) - \alpha v_{T,\alpha}(n^R + 2, n^S + 1) \leq \alpha v_{T,\alpha}(n^R + 2, n^S) - \alpha v_{T,\alpha}(n^R + 3, n^S)$. As $R_i^R - r_i(n^R + 2) \geq 0$, $\alpha v_{T,\alpha}(n^R + 2, n^S) - \alpha v_{T,\alpha}(n^R + 3, n^S) = \mathbb{E}(R_3 - R_4^*|A_1)$.

Scenario 6 for A_1: Assume that the optimal policy routes the arrived customer of class i to the self-service station at $(n^R, n^S + 1)$ and rejects at $(n^R + 2, n^S)$ state. Let us consider that Process 2 and Process 3 mimic Process 1 and Process 4, respectively. Then, $\mathbb{E}(R_1^* - R_2|A_1) = \alpha v_{T,\alpha}(n^R, n^S + 2) - \alpha v_{T,\alpha}(n^R + 1, n^S + 2)$. By the induction hypothesis of statement (IV), $\alpha v_{T,\alpha}(n^R, n^S + 2) - \alpha v_{T,\alpha}(n^R + 1, n^S + 2) \leq \alpha v_{T,\alpha}(n^R + 1, n^S + 1) - \alpha v_{T,\alpha}(n^R + 2, n^S + 1)$. By the induction hypothesis of statement (II), $\alpha v_{T,\alpha}(n^R + 1, n^S + 1) - \alpha v_{T,\alpha}(n^R + 2, n^S + 1) \leq \alpha v_{T,\alpha}(n^R, n^S + 2) - \alpha v_{T,\alpha}(n^R + 1, n^S + 2)$, where $\alpha v_{T,\alpha}(n^R, n^S + 2) - \alpha v_{T,\alpha}(n^R + 1, n^S + 2) = \mathbb{E}(R_3 - R_4^*|A_1)$.

Scenario 7 for A_1: Assume that the optimal policy rejects at both states. Let Process 2 and Process 3 reject as well. Then, $\mathbb{E}(R_1^* - R_2|A_1) = \alpha v_{T,\alpha}(n^R, n^S + 1) - \alpha v_{T,\alpha}(n^R + 1, n^S + 1) \leq \alpha v_{T,\alpha}(n^R + 1, n^S) - \alpha v_{T,\alpha}(n^R + 2, n^S) = \mathbb{E}(R_3 - R_4^*|A_1)$, by the induction hypothesis of statement (IV).

Scenario 8 for A_1: Assume that the optimal policy rejects at $(n^R, n^S + 1)$ and routes to the self-service station at $(n^R + 2, n^S)$. Let us consider that Process 2 and Process 3 mimic Process 1 and Process 4, respectively. Then, $\mathbb{E}(R_1^* - R_2|A_1) = \alpha v_{T,\alpha}(n^R, n^S + 1) - \alpha v_{T,\alpha}(n^R + 1, n^S + 1)$. With the induction hypothesis of statement (II), $\alpha v_{T,\alpha}(n^R, n^S + 1) - \alpha v_{T,\alpha}(n^R + 1, n^S + 1) \leq \alpha v_{T,\alpha}(n^R, n^S + 2) - \alpha v_{T,\alpha}(n^R + 1, n^S + 2)$ and with the induction hypothesis of statement (IV), $\alpha v_{T,\alpha}(n^R, n^S + 2) - \alpha v_{T,\alpha}(n^R + 1, n^S + 2) \leq \alpha v_{T,\alpha}(n^R + 1, n^S + 1) - \alpha v_{T,\alpha}(n^R + 2, n^S + 1)$, where $\alpha v_{T,\alpha}(n^R + 1, n^S + 1) - \alpha v_{T,\alpha}(n^R + 2, n^S + 1) = \mathbb{E}(R_3 - R_4^*|A_1)$.

Scenario 9 for A_1: Assume that the optimal policy rejects at $(n^R, n^S + 1)$ and routes to the regular station at $(n^R + 2, n^S)$. Let us consider that Process 2 and Process 3 mimic Process 1 and Process 4, respectively. Then, $\mathbb{E}(R_1^* - R_2|A_1) = \alpha v_{T,\alpha}(n^R, n^S + 1) - \alpha v_{T,\alpha}(n^R + 1, n^S + 1)$. With the induction hypothesis of statement (IV), $\alpha v_{T,\alpha}(n^R, n^S + 1) - \alpha v_{T,\alpha}(n^R + 1, n^S + 1) \leq \alpha v_{T,\alpha}(n^R + 1, n^S) - \alpha v_{T,\alpha}(n^R + 2, n^S)$. With the induction hypothesis of statement (II), $\alpha v_{T,\alpha}(n^R + 1, n^S) - \alpha v_{T,\alpha}(n^R + 2, n^S) \leq \alpha v_{T,\alpha}(n^R + 2, n^S) - \alpha v_{T,\alpha}(n^R + 3, n^S)$, where $\alpha v_{T,\alpha}(n^R + 2, n^S) - \alpha v_{T,\alpha}(n^R + 3, n^S) = \mathbb{E}(R_3 - R_4^*|A_1)$.

In all decision scenarios possible after observing event A_1, we see that we can confirm $\mathbb{E}(R_1^* - R_2|A_1) \leq \mathbb{E}(R_3 - R_4^*|A_1)$.

Now let us look into the case service completion events occur in the first time period. Let us first consider the service completion event from the regular service station. Let A_2 denote this event.

Scenario 1 for A_2: First, let us consider the case that $n^R > 0$ such that all processes can observe this event. Then, $\mathbb{E}(R_1^* - R_2|A_2) = \alpha[v_{T,\alpha}(n^R - 1, n^S + 1) - v_{T,\alpha}(n^R, n^S + 1)] \leq \alpha[v_{T,\alpha}(n^R, n^S) - v_{T,\alpha}(n^R + 1, n^S)] = \mathbb{E}(R_3 - R_4^*|A_2)$, by the induction hypothesis of statement (IV).

Scenario 2 for A_2: Now, let $n^R = 0$ such that Process 1 can not observe this event. Then, $\mathbb{E}(R_1^* - R_2|A_2) = \alpha[v_{T,\alpha}(n^R, n^S + 1) - v_{T,\alpha}(n^R, n^S + 1)] = 0 \leq \alpha[v_{T,\alpha}(n^R, n^S) - v_{T,\alpha}(n^R + 1, n^S)] = \mathbb{E}(R_3 - R_4^*|A_2)$, by Lemma 3.1.

Let us now consider the service completion event from the self-service station. Let A_3 denote this event.

Scenario 1 for A_3: First, let us consider the case that $n^S > 0$ such that all processes can observe this event. Then, $\mathbb{E}(R_1^* - R_2|A_3) = \alpha[v_{T,\alpha}(n^R, n^S) - v_{T,\alpha}(n^R + 1, n^S)] \leq \alpha[v_{T,\alpha}(n^R + 1, n^S - 1) - v_{T,\alpha}(n^R + 2, n^S - 1)] = \mathbb{E}(R_3 - R_4^*|A_3)$, by the induction hypothesis of statement (IV).

Scenario 2 for A_3: Now consider that $n^S = 0$ such that only Process 1 and Process 2 can observe this event. Then, $\mathbb{E}(R_1^* - R_2|A_3) = \alpha[v_{T,\alpha}(n^R, n^S) - v_{T,\alpha}(n^R + 1, n^S)] \leq \alpha[v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 2, n^S)] = \mathbb{E}(R_3 - R_4^*|A_3)$, by the induction hypothesis of statement (V).

Lastly, a dummy transition event might occur due to uniformization, which does not change the state of any processes. So, the induction hypothesis of statement (IV) is sufficient in this case.

Proving Statement (V): This is the convexity property of the value functions with respect to the number of customers in the regular station. We show this by induction on the value functions. We can show the initial induction step by letting $v_{0,\alpha}(..) = 0$. Then, assuming that (3) holds for some T and $\alpha \in [0, 1)$ and we need to show that (3) also holds for $T + 1$ and $\alpha \in [0, 1)$.

We know that

$$
\begin{aligned}
v_{T+1,\alpha}(n^R, n^S) - v_{T+1,\alpha}(n^R + 1, n^S) = {} & R_\alpha v_{T,\alpha}(n^R, n^S) - R_\alpha v_{T,\alpha}(n^R + 1, n^S) \\
& + \alpha \mu^R(n^R)[v_{T,\alpha}(n^R - 1, n^S) - v_{T,\alpha}(n^R, n^S)] \\
& + \alpha \mu^S(n^S)[v_{T,\alpha}(n^R, n^S - 1) - v_{T,\alpha}(n^R + 1, n^S - 1)] \\
& + \alpha(\mu^S(c^S - n^S) + \mu^R(N) - \mu^R(n^R + 1)) \\
& \times [v_{T,\alpha}(n^R, n^S) - v_{T,\alpha}(n^R + 1, n^S)]
\end{aligned}
\tag{A.3}
$$

where $R_\alpha v_{T,\alpha}(n^R, n^S) - R_\alpha v_{T,\alpha}(n^R + 1, n^S)$ is

$$- h + \sum_i \lambda_i [\max[R_i^R + \alpha v_{T,\alpha}(n^R + 1, n^S), R_i^S + \alpha v_{T,\alpha}(n^R, n^S + 1), \alpha(n^R, n^S)]$$

$$- \max[R_i^R + \alpha v_{T,\alpha}(n^R + 2, n^S), R_i^S + \alpha v_{T,\alpha}(n^R + 1, n^S + 1), \alpha(n^R + 1, n^S)]]. \qquad \textbf{(A.4)}$$

We can show this statement by proving that each term in brackets in equation (A.3) is bounded above by $v_{T+1,\alpha}(n^R + 1, n^S) - v_{T+1,\alpha}(n^R + 2, n^S)$. Then, we can be sure that the statement will hold as the coefficients of these terms sum up to 1.

We first look into the three parts of (A.3) that do not relate to the rewards.

(a) $v_{T,\alpha}(n^R - 1, n^S) - v_{T,\alpha}(n^R, n^S) \le v_{T,\alpha}(n^R, n^S) - v_{T,\alpha}(n^R + 1, n^S) \le v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 2, n^S)$ by the induction hypothesis of (V). By Lemma A.1, $v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 2, n^S) \le v_{T+1,\alpha}(n^R + 1, n^S) - v_{T+1,\alpha}(n^R + 2, n^S)$.

(b) $v_{T,\alpha}(n^R, n^S - 1) - v_{T,\alpha}(n^R + 1, n^S - 1) \le v_{T,\alpha}(n^R, n^S) - v_{T,\alpha}(n^R + 1, n^S)$ by the induction hypothesis of statement (II). By the induction hypothesis of (V), $v_{T,\alpha}(n^R, n^S) - v_{T,\alpha}(n^R + 1, n^S) \le v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 2, n^S)$, and finally by Lemma A.1, $v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 2, n^S) \le v_{T+1,\alpha}(n^R + 1, n^S) - v_{T+1,\alpha}(n^R + 2, n^S)$.

(c) $v_{T,\alpha}(n^R, n^S) - v_{T,\alpha}(n^R + 1, n^S) \le v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 1, n^S)$, by the induction hypothesis of (V) and we can finally make use of Lemma A.1.

Now check the reward differences. In the notation, we drop the holding cost $-h$ in (A.3), as it does not affect the conclusions anyway. We show that, for any arbitrary customer class i, $R_\alpha^i v_{T,\alpha}(n^R, n^S) - R_\alpha^i v_{T,\alpha}(n^R + 1, n^S) := \max[R_i^R + \alpha v_{T,\alpha}(n^R + 1, n^S), R_i^S + \alpha v_{T,\alpha}(n^R, n^S + 1), \alpha(n^R, n^S)] - \max[R_i^R + \alpha v_{T,\alpha}(n^R + 2, n^S), R_i^S + \alpha v_{T,\alpha}(n^R + 1, n^S + 1), \alpha(n^R + 1, n^S)]$ is bounded above by $v_{T+1,\alpha}(n^R + 1, n^S) - v_{T+1,\alpha}(n^R + 2, n^S)$.

(a) Consider that it is optimal to route class i to the regular station at both (n^R, n^S) and $(n^R + 1, n^S)$ states. Then, we know that $R_\alpha^i v_{T,\alpha}(n^R, n^S) - R_\alpha^i v_{T,\alpha}(n^R + 1, n^S) = \alpha[v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 2, n^S)]$. We know that with Lemma A.1 that $\alpha[v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 2, n^S)] \le \alpha[v_{T+1,\alpha}(n^R + 1, n^S) - v_{T+2,\alpha}(n^R + 2, n^S)]$.

(b) Consider that it is optimal to route class i to the regular station at (n^R, n^S) but to the self-service station at $(n^R + 1, n^S)$. We know that, as it is not optimal to route class i to the regular station at $(n^R + 1, n^S)$, $R_\alpha^i v_{T,\alpha}(n^R, n^S) - R_\alpha^i v_{T,\alpha}(n^R + 1, n^S) \le \alpha[v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 2, n^S)]$. For the rest, we can follow (d).

(c) Similarly, for the case that it is optimal to route class i to the regular station at (n^R, n^S) but to reject at $(n^R + 1, n^S)$, we can infer that $R_\alpha^i v_{T,\alpha}(n^R, n^S) - R_\alpha^i v_{T,\alpha}(n^R + 1, n^S) \le \alpha[v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 2, n^S)]$ and follow the lines in (d).

(d) Consider that it is optimal to route class i to the self-service station at both (n^R, n^S) and $(n^R + 1, n^S)$ states. Then, we know that $R_\alpha^i v_{T,\alpha}(n^R, n^S) - R_\alpha^i v_{T,\alpha}(n^R + 1, n^S) = \alpha[v_{T,\alpha}(n^R, n^S + 1) - v_{T,\alpha}(n^R + 1, n^S + 1)]$. With the property (IV), we know that $\alpha[v_{T,\alpha}(n^R, n^S + 1) - v_{T,\alpha}(n^R + 1, n^S + 1)] \le \alpha[v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 2, n^S)]$.

(e) Consider that it is optimal to route class i to the self-service station at (n^R, n^S) but to the regular station at $(n^R + 1, n^S)$. We know that, as it is not optimal to route class i to the self-service station at $(n^R + 1, n^S)$, $R_\alpha^i v_{T,\alpha}(n^R, n^S) - R_\alpha^i v_{T,\alpha}(n^R + 1, n^S) \le \alpha[v_{T,\alpha}(n^R, n^S + 1) - v_{T,\alpha}(n^R + 1, n^S + 1)]$. For the rest, we can follow (g).

(f) Consider that it is optimal to route class i to the self-service station at (n^R, n^S) but to reject at $(n^R + 1, n^S)$. We know that, as it is not optimal to route class i to the self-service station at $(n^R + 1, n^S)$, $R_\alpha^i v_{T,\alpha}(n^R, n^S) - R_\alpha^i v_{T,\alpha}(n^R + 1, n^S) \le \alpha[v_{T,\alpha}(n^R, n^S + 1) - v_{T,\alpha}(n^R + 1, n^S + 1)]$. For the rest, we can follow (g).

(g) Consider that it is optimal to reject class i at both (n^R, n^S) and $(n^R + 1, n^S)$ states. Then, we know that $R_\alpha^i v_{T,\alpha}(n^R, n^S) - R_\alpha^i v_{T,\alpha}(n^R + 1, n^S) = \alpha[v_{T,\alpha}(n^R, n^S) - v_{T,\alpha}(n^R + 1, n^S)]$. With the induction hypothesis of (V), $\alpha[v_{T,\alpha}(n^R, n^S) - v_{T,\alpha}(n^R + 1, n^S)] \leq \alpha[v_{T,\alpha}(n^R + 1, n^S) - v_{T,\alpha}(n^R + 2, n^S)]$ and finally we can use Lemma A.1.

With the induction hypothesis of (V), it is impossible that class i is routed to the regular at $(n^R + 1, n^S)$ but rejected at (n^R, n^S). With the property (II), it is also impossible that class i routed to the self-service station at $(n^R + 1, n^S)$ but rejected at (n^R, n^S).