

Dynamic admission control in a call center with one shared and two dedicated service facilities

E. Lerzan Örmeci ¹

Department of Industrial Engineering
Koç University
Rumeli Feneri Yolu
34450 Sarıyer, İstanbul, Turkey

email address: lormeci@ku.edu.tr
Tel: 90(212)338-1534
Fax: 90(212)338-1548

Abstract

Calls of two classes arrive at a call center according to two independent Poisson processes. The center has two dedicated stations, one for each class, and one shared station. All three stations consist of parallel servers and no waiting room. Calls of each type demand exponential service times with different service rates and generate different rewards. Moreover, the service rates are different in the shared and dedicated stations. We assume non-preemptive service. Our objective is to derive the structure of dynamic admission policies that maximize the total expected discounted revenue over an infinite horizon as well as the long-run average revenue. We show that it is optimal to serve a customer in her dedicated station whenever it is possible. For the shared station, we derive a sufficient condition for each class under which it is always optimal to accept customers of that class to the shared station if the dedicated station is full and the shared station has available servers. Furthermore, the optimal admission policy at the shared station can be characterized as a monotonic threshold policy.

¹The research was done while the author was at EURANDOM, Eindhoven, The Netherlands

1 Introduction

In the last decades, call centers have been an effective and low cost customer service in a variety of industries such as financial services, airlines, hotels, and retail companies. Typically, call centers serve different classes of customers, each of which can be distinguished by a different profitability, volume of calls, and expectation for service. Call center operators often require different types of training to handle different classes of customers. However, because cross-training is expensive, it is necessary to manage the workforce training and call allocation carefully. In order to develop insight regarding the extent to which operators would be cross-trained, and how calls should be dynamically assigned to operators, we consider the problem of call admission in a system with no waiting room that serves two classes of customers using dedicated and shared facilities.

Consider a call center which serves two classes of customers, each of which requires different sets of skills from the operators of the center. At one extreme, the center can have two completely independent call centers, each dedicated to one of the classes, and at the other one it may consist of only one type of operators who are able to serve both classes. The former system suffers from either low utilization or low quality of service due to a loss in economies of scale, whereas the latter one has to face the high cost of cross-training many operators. Hence, we consider an intermediate model in which the call center consists of two dedicated stations, one for each class, and one shared station that can serve both classes. We assume that the system has no waiting room so that customers are assumed to be lost if all servers are busy, or if the system chooses not to serve them. This assumption is not restrictive when modelling large call centers, since such centers can answer about half of their customers immediately, and the waiting time of those delayed is measured in seconds (Gans, Koole and Mandelbaum, [5]). Moreover, the fraction of abandoning calls while waiting varies from 0 to 1-2%. Hence, the effect of delaying a customer as well as the probability of a delay is low, which allows us to ignore the waiting room completely. This study can also be viewed as a starting point for analyzing finite-buffer systems, which can model any kind of call centers as well as many telecommunication and manufacturing systems.

Our model assumes that customers of each class arrive at the system according to a Poisson process, and demand exponential service times. Each class has different arrival and service rates, and generates different rewards. Moreover, the service rate of a customer depends on the station he is being served, i.e., the service rate for each class is different in the shared and dedicated stations. We do not allow preemption as a call center would require. Our objective is to derive the structure of dynamic admission policies that maximize the total expected discounted revenue over an infinite horizon as well as the long-run average revenue.

We show that the optimal policy accepts customers of each class to their dedicated stations whenever that station has available servers. The optimal admission policy in the shared station, on the other hand, can be characterized in two ways: We prove the existence of a monotonic threshold policy, where the thresholds depend on the number of customers in all three stations. Moreover, we derive a sufficient condition for each class to guarantee that customers of that class are always admitted to the shared station whenever their dedicated station is full, and the shared station has at least one idle server. We call such a class a *preferred class*.

This paper is organized as follows: In the next section, we present a brief literature review on the related studies. Section 3 develops a Markov decision process (MDP) model for the system described above. Section 4 characterizes the structure of optimal admission rules for the dedicated stations. In sections 5 and 6, we consider the admission control in the shared station: Section 5 presents sets of

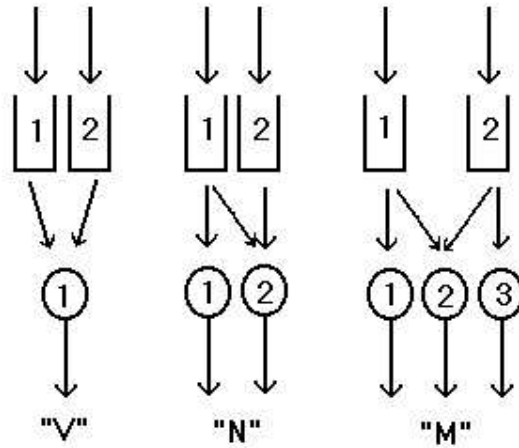


Figure 1: Some canonical designs for skills-based routing (from Garnett and Mandelbaum (2001))

sufficient conditions for each class to be preferred, while section 6 establishes the existence of an optimal monotonic threshold policy. Finally, we conclude and point out possible future research in section 7.

2 Related research

Our work is related with the literature in two different areas: call centers due to our motivation, and admission control of loss systems due to our model.

The increasing use of call centers in different industries has generated recent research on these systems. Gans et al. [5] give a comprehensive review for analytical models that support capacity management issues of call centers. Complexity of call centers provides a wide variety of different problems: Workforce management has become one of the issues, see e.g., Akşin and Harker [2] who consider a staffing problem of a call center with multiple call types and service agents that specialize in different call types. Computing performance measures of call centers, such as the fraction of lost sales due to blocking (busy signals) or renegeing, is another important issue, which is addressed by Akşin and Harker [1] among others. Our work, on the other hand, is most closely related to the problem of routing different kinds of customers to servers with different skills. There are a number of canonical designs for skill-based routing (some of which are depicted in Figure 1) introduced in Garnett and Mandelbaum [7]. In V-design, two classes of calls are served by a single pool of cross-trained agents, variants of which are analyzed by a number of authors such as Bhulai and Koole [4] and Gans and Zhou [6]. Stanford and Grassman [17] and Shumsky [16] study the N-design assuming fixed static priority policies, whereas the optimal structure of dynamic routing in such systems is considered by Xu, Richter, and Shanthikumar [19] in a context different than call centers. In this paper, we address the dynamic control in an M-design with no waiting room.

Admission control is a main focus of research on loss networks, see Chapter 4 of Ross [15] for a comprehensive review on the admission control problems of generalized stochastic knapsacks. Earlier

studies which investigate the structural properties of optimal admission policies for certain stochastic knapsacks include Miller [13] and Lippman and Ross [12]. More recently, Altman, Jiménez and Koole [3] and Koole [9] prove the existence of an optimal admission policy in a stochastic knapsack with two classes of customers, that is characterized by acceptance thresholds, while Örmeci, Burnetas, and van der Wal [14] establish the monotonicity of these thresholds under some restrictive conditions. Örmeci et al. [14] also introduce the notion of “preferred class” in stochastic knapsacks. We extend the existence of optimal thresholds and preferred class(es) to stochastic knapsacks receiving arrivals according to an overflow process rather than a Poisson process.

We also investigate the dependence of admission control in the shared station on the state of the dedicated stations. To the best of our knowledge, Ku and Jordan [10] is the only study which considers admission control in a system with multiple stations of loss systems, and the interaction between them. They show the existence of an optimal threshold policy for two stations in tandem, each with no waiting room and parallel servers.

3 Markov decision model

We assume that class- i customers arrive at the system according to a Poisson process with rate λ_i . Stations 1 and 2 are dedicated to customers of class 1 and 2, respectively, whereas station 0 can be used by either of the classes. If a class- i customer is served at station i , then she demands an exponential service with mean $1/\mu'_i$, and if she is served in station 0, then she is served for an exponential time with rate μ_i . We assume, without loss of generality, that class-1 customers require longer service times so that $\mu_2 \leq \mu_1$ and $\mu'_2 \leq \mu'_1$. The operators in station 0 will be trained to be able to serve both classes, so we assume that these servers are slower than the “specialized” servers. Therefore, we assume $\mu_j \leq \mu'_j$. Each served class- i customer brings a reward of r_i upon finishing her service, regardless of the station she is served.

The original process of the system evolves in continuous time. All the interarrival times as well as the service times are exponential. Furthermore, we interpret discounting as exponential failures, i.e., the system closes down in an exponentially distributed time with rate β (for the equivalence of the process with discounting and the process without discounting but with an exponential deadline, see e.g., Walrand [18]). These features allow us to build a discrete time equivalent of this system by using uniformization (introduced by Lippman [11]): The maximum possible rate out of any state, i.e., $\sum_{i=1,2} (\lambda_i + c_i \mu'_i) + c\mu_2 + \beta$, is finite so that we can assume, using the appropriate time scale, $\sum_{i=1,2} (\lambda_i + c_i \mu'_i) + c\mu_2 + \beta = 1$. As a result, we observe the state of the system at each instant of a potential transition, so in every exponentially distributed time with rate 1, and the system changes its state with certain probabilities to be specified below.

Now we can define the state of the system as $(\mathbf{y}; \mathbf{x}) = (y_1, y_2; x_1, x_2)$, where y_i is the number of class- i customers in station i , and x_i is the number of class- i customers in station 0. Let \mathcal{S} be the set of all feasible states, i.e., $\mathcal{S} = \{(\mathbf{y}; \mathbf{x}) : y_i \leq c_i, i = 1, 2; x_1 + x_2 \leq c\}$, so that if $\mathbf{s} = (s_1, s_2; s_3, s_4) \in \mathcal{S}$, s_1 and s_2 are the number of customers in stations 1 and 2, respectively, whereas s_3 and s_4 are the number of class-1 and class-2 customers, respectively, in station 0.

Let $u^n(\mathbf{y}; \mathbf{x})$ be the maximal expected β -discounted reward for the system starting in state $(\mathbf{y}; \mathbf{x})$ when n observation points remain in the horizon. Now we can present the optimality equations, where

we set \mathbf{e}_j as the unit vector which has a 1 at the j th coordinate, and 0 elsewhere:

$$u^{n+1}(\mathbf{y}; \mathbf{x}) = \sum_{j=1}^2 [x_j \mu_j r_j + y_j \mu'_j r_j + \lambda_j v_j^n(\mathbf{y}; \mathbf{x}) + x_j \mu_j u^n(\mathbf{y}; \mathbf{x} - \mathbf{e}_j) + y_j \mu'_j u^n(\mathbf{y} - \mathbf{e}_j; \mathbf{x})] + \gamma u^n(\mathbf{y}; \mathbf{x}), \quad (1)$$

where

$$v_j^n(\mathbf{y}; \mathbf{x}) = \max \{ u^n(\mathbf{y} + \mathbf{e}_j; \mathbf{x}), u^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_j), u^n(\mathbf{y}; \mathbf{x}) \} \quad \text{and} \quad (2)$$

$$\gamma = c\mu_2 + \sum_{i=1}^2 (c_i \mu'_i - x_i \mu_i - y_i \mu'_i), \quad (3)$$

with $u^n(\mathbf{y}; \mathbf{x} - \mathbf{e}_j) = u^n(\mathbf{y}; \mathbf{x})$ if $x_j = 0$, $u^n(\mathbf{y} - \mathbf{e}_j; \mathbf{x}) = u^n(\mathbf{y}; \mathbf{x})$ if $y_j = 0$, $u^n(\mathbf{y} + \mathbf{e}_j; \mathbf{x}) = -\infty$ when $y_j + 1 > c_j$ and $u^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_j) = -\infty$ whenever $x_1 + x_2 + 1 > c$. Therefore, if e.g., $x_1 + x_2 = c$ and $y_j = c_j$, then $v_j^n(\mathbf{y}; \mathbf{x}) = u^n(\mathbf{y}; \mathbf{x})$. We define the action $a_j^n(\mathbf{y}; \mathbf{x})$ as the optimal state to move into when a class- j customer arrives at a system in state $(\mathbf{y}; \mathbf{x})$ with n remaining transitions. If a class- j customer arrives, which happens with probability λ_j , he is either rejected, that keeps the system in the same state, i.e., $a_j^n(\mathbf{y}; \mathbf{x}) = (\mathbf{y}; \mathbf{x})$, or depending on the availability of servers he is accepted to either station j , moving the system to state $a_j^n(\mathbf{y}; \mathbf{x}) = (\mathbf{y} + \mathbf{e}_j; \mathbf{x})$, or to station 0, changing the state to $a_j^n(\mathbf{y}; \mathbf{x}) = (\mathbf{y}; \mathbf{x} + \mathbf{e}_j)$. If a class- j customer finishes his service in station j , with probability $y_j \mu'_j$, the system receives a reward of r_j and the state changes to $(\mathbf{y} - \mathbf{e}_j; \mathbf{x})$. If a class- j customer finishes his service in station 0, with probability $x_j \mu_j$, again a reward of r_j is gained and the system moves to state $(\mathbf{y}; \mathbf{x} - \mathbf{e}_j)$. The ‘‘fictitious’’ service completions, which occur with probability γ given by equation (3), affect neither the state nor the total reward of the system. Finally, if the system closes down, with probability β , the system receives no more reward.

We can observe from the optimality equations (2) that the effect of an additional class- j customer in stations j and 0, as well as of moving a customer from station j to station 0, is an important quantity. We will see in our analysis below that the effect of changing a class- j customer to a class- k at station 0 is also important. Therefore, we define the functions $D_{JK}^n(jk)(\mathbf{y}; \mathbf{x})$ as the difference in the total expected discounted rewards between two systems A and B, where system A starts in state $(\mathbf{y}; \mathbf{x})$ ‘plus’ one class- j customer at station J , and system B starts in $(\mathbf{y}; \mathbf{x})$ plus a class- k customer at station K . Of course, J can be 0 or j and K can be 0 or k . Moreover, $k = 0$ means that system B is in state $(\mathbf{y}; \mathbf{x})$, i.e., there is no additional customer. We also denote $D_{LL}^n(jk)(\mathbf{y}; \mathbf{x})$ by $D_L^n(jk)(\mathbf{y}; \mathbf{x})$, i.e., if both additional customers are in the same station, we indicate it once. We occasionally omit the arguments $(\mathbf{y}; \mathbf{x})$ and n when there is no danger of confusion. $D_{JK}(jk)$ functions that will be used in our analysis below are as follows:

$$\begin{aligned} D_j^n(j0)(\mathbf{y}; \mathbf{x}) &= u^n(\mathbf{y} + \mathbf{e}_j; \mathbf{x}) - u^n(\mathbf{y}; \mathbf{x}), \\ D_0^n(j0)(\mathbf{y}; \mathbf{x}) &= u^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_j) - u^n(\mathbf{y}; \mathbf{x}), \\ D_{j0}^n(jj)(\mathbf{y}; \mathbf{x}) &= u^n(\mathbf{y} + \mathbf{e}_j; \mathbf{x}) - u^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_j), \quad \text{and} \\ D_0^n(12)(\mathbf{y}; \mathbf{x}) &= u^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_1) - u^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_2). \end{aligned}$$

We can interpret the difference $D_j^n(j0)(\mathbf{y}; \mathbf{x})$ as the net benefit of the system due to an additional class- j customer at station J , $J = 0, j$, in state $(\mathbf{y}; \mathbf{x})$ when there are n more transitions, and $D_{j0}^n(jj)(\mathbf{y}; \mathbf{x})$ as the net benefit of the system when a class- j customer is moved from station 0 to station j in state $(\mathbf{y}; \mathbf{x} + \mathbf{e}_j)$,

whereas $D_0^n(12)(\mathbf{y}; \mathbf{x})$ is the net benefit of the system when a class-2 customer at station 0 is changed to a class-1 customer in state $(\mathbf{y}; \mathbf{x} + \mathbf{e}_2)$. Note that $D_k^n(jk) = -D_k^n(kj)$, and $D_{j0}^n(jj) = -D_{0j}^n(jj)$.

One of our major tools to show certain properties of the value functions is coupling: We couple two systems so that both will receive the same arrival stream. Moreover, the service times of customers in the two systems are coupled as follows: If the coupled customers are of the same class and being served in the same station, they depart at the same time, otherwise we use the assumptions that $\mu_1 \leq \mu_2$ or $\mu_j \leq \mu'_j$: If a class-1 customer is coupled with a class-2 customer, both in station 0, then whenever the coupled class-1 customer finishes her service, the coupled class-2 customer also completes his service with probability 1, due to $\mu_1 \leq \mu_2$. In terms of discrete time, this translates to the following: Both customers leave the system with probability μ_1 , and only the class-2 customer departs from the system with probability $\mu_2 - \mu_1$ leaving the coupled class-1 customer in the system. Thus, coupling does not allow a coupled class-1 customer to leave the system while the coupled class-2 customer is still there. Coupling additional class- j customers in stations j and 0 is similar: Both customers leave the system with probability μ_j , and only the customer in station j departs with probability $\mu'_j - \mu_j$.

We prove all our results for the objective of maximizing total expected β -discounted reward for a finite number of transitions, n , including the “fictitious” transitions due to the “fictitious” service completions. Thus, “finite” horizon problems are pseudo finite problems. They provide the powerful tool of induction to prove our results for all n . To start the induction, we need an initial value function, which satisfies the statement under consideration. Here, we present one such function which requires that the rewards of customers, who are still in the system at $n = 0$, are collected even if their services have not been finished. Of course, this makes no difference in the optimal policy for infinite horizon problems. More specifically:

$$u^0(\mathbf{y}; \mathbf{x}) = \sum_{j=1}^2 (y_j R'_j + x_j R_j) \quad \forall (\mathbf{y}; \mathbf{x}) \in \mathcal{S}, \quad (4)$$

where $R'_j = r_j \mu'_j / (\mu'_j + \beta)$ and $R_j = r_j \mu_j / (\mu_j + \beta)$ are the present values of the reward brought by a class- j customer served in station j and 0, respectively. We will refer to R'_j and R_j as the immediate reward of a class- j customer at station j and at station 0, respectively.

All the results proven for finite n are true for the limit $n \rightarrow \infty$, so the corresponding conclusions are valid when total expected β -discounted reward over an infinite horizon is maximized. Hence, we define $u(\mathbf{y}; \mathbf{x})$ as the maximal expected β -discounted reward for the system starting in state $(\mathbf{y}; \mathbf{x})$ over an infinite horizon. For $\beta > 0$, we have:

$$u(\mathbf{y}; \mathbf{x}) = \lim_{n \rightarrow \infty} u^n(\mathbf{y}; \mathbf{x}).$$

Moreover, since the state space and the action space in each state are finite and the results hold even for $\beta = 0$, we have the same conclusions for maximizing the long-run average reward. Only for $\beta = 0$, we need to consider the relative value functions and the gain in the usual MDP formulation, since $u(\mathbf{y}; \mathbf{x}) \rightarrow \infty$.

4 Optimal admission policy at station 1 and 2

This section specifies the optimal admission policy for stations 1 and 2, which needs derivation of certain bounds on $D_{JK}(jk)$ functions. The following lemma establishes these bounds, some of which will be used in the subsequent sections as well:

Lemma 1 (i) For all $(\mathbf{y} + \mathbf{e}_j; \mathbf{x}) \in \mathcal{S}$: $0 \leq D_j^n(j0)(\mathbf{y}; \mathbf{x}) \leq R'_j$ for all $n \geq 0$ and for $j = 1, 2$.

(ii) For all $(\mathbf{y}; \mathbf{x} + \mathbf{e}_1) \in \mathcal{S}$: $D_0^n(j0)(\mathbf{y}; \mathbf{x}) \leq R_j$ for all $n \geq 0$ and for $j = 1, 2$.

(iii) For all $(\mathbf{y} + \mathbf{e}_j; \mathbf{x} + \mathbf{e}_1) \in \mathcal{S}$: $0 \leq D_{j0}^n(jj)(\mathbf{y}; \mathbf{x})$ for all $n \geq 0$ and for $j = 1, 2$.

(iv) For all $(\mathbf{y}; \mathbf{x} + \mathbf{e}_1) \in \mathcal{S}$: $D_0^n(12)(\mathbf{y}; \mathbf{x}) \leq R_1 - R_2$ for all $n \geq 0$.

Proof. All proofs are by induction on the number of transitions combined with coupling. All statements are true for u^0 given by equation (4). Assume that they are true for n , and consider $n + 1$.

(i) Consider the first inequality: Let system A be in state $(\mathbf{y} + \mathbf{e}_j; \mathbf{x})$, and system B in state $(\mathbf{y}; \mathbf{x})$. We let system B follow the optimal policy. If system B accepts a class- j customer to station j , system A rejects him, so that the two systems couple. Otherwise, system A imitates all decisions of system B, keeping the difference between the two systems due to the additional customer in system A. Note that system A can imitate system B in all other decisions since they have the same number of customers in all other stations. If the additional class- j customer finishes his service in system A, which happens with probability μ'_j , then the two systems couple bringing a difference of r_j in rewards. All other service completions keep the difference between the two systems the same. Then we have:

$$\begin{aligned} D_j^{n+1}(j0)(\mathbf{y}; \mathbf{x}) &\geq \lambda_j \min \left\{ 0, \min_{\mathbf{s} + \mathbf{e}_j \in \mathcal{S}} \{D_j^n(j0)(\mathbf{s})\} \right\} + r_j \mu'_j \\ &\quad + (1 - \lambda_j - \mu'_j - \beta) \min_{\mathbf{s} + \mathbf{e}_j \in \mathcal{S}} \{D_j^n(j0)(\mathbf{s})\} \geq 0, \end{aligned}$$

where the first inequality is due to the policy of system A, uniformization and the coupling described above, and the second follows from the induction hypothesis.

Now consider the second inequality, i.e., upperbound. Let system A and B be in the same states as above, but now we let system A follow the optimal policy, and system B imitate all its decisions. Note that system B can always imitate system A since its stations have at most the same number of customers with system A. Then, the only difference between the two systems is due to the reward r_j of the additional customer in system A, gained upon her departure with probability μ'_j . Hence, we have:

$$D_j^{n+1}(j0)(\mathbf{y}; \mathbf{x}) \leq r_j \mu'_j + (1 - \mu'_j - \beta) R'_j = R'_j$$

where the inequality is due to the policy system B is following, uniformization and the induction hypothesis. Note that $R'_j(\mu'_j + \beta) = r_j \mu'_j$. Thus, both inequalities hold for all $n \geq 0$.

(ii) The proof is very similar to that of the second inequality in part (i), so it is omitted.

(iii) Let system A be in state $(\mathbf{y} + \mathbf{e}_j; \mathbf{x})$, and system B in state $(\mathbf{y}; \mathbf{x} + \mathbf{e}_j)$. We let system B follow the optimal policy. System A can imitate all decisions of system B except for a class- j arrival admitted to station j , and in this case system A accepts her to station 0, so that the two systems couple. We couple the service times of the additional class- j customers in station j , say customer d'_j , and in station 0, say customer d_j , as well as all other service and interarrival times. Then, both d'_j and d_j finish their services with probability μ_j , coupling the two systems. With probability $\mu'_j - \mu_j$, d'_j departs from system A while d_j is still receiving service, so that system B remains in its current state, while system A moves to $(\mathbf{y}; \mathbf{x})$ with an additional reward of r_j . All other service completions keep the same difference. Then

we have:

$$\begin{aligned}
D_{j0}^{n+1}(jj)(\mathbf{y}; \mathbf{x}) &\geq \lambda_j \min \left\{ 0, \min_{\mathbf{s} + \mathbf{e}_j + \mathbf{e}_3 \in \mathcal{S}} \{D_{j0}^n(jj)(\mathbf{s})\} \right\} + \mu_j \times 0 \\
&\quad + (\mu'_j - \mu_j)(r_j - D_0^n(j0)(\mathbf{y}; \mathbf{x})) \\
&\quad + (1 - \lambda_j - \mu'_j - \beta) \min_{\mathbf{s} + \mathbf{e}_j + \mathbf{e}_3 \in \mathcal{S}} \{D_{j0}^n(jj)(\mathbf{s})\} \geq 0,
\end{aligned}$$

where the first inequality is due to the policy of system A, uniformization and the coupling described above, and the second is due to the induction hypothesis, part (ii) and $r_j \geq R_j$.

(iv) Assume that system A starts in state $(\mathbf{y}; \mathbf{x} + \mathbf{e}_1)$ and system B in $(\mathbf{y}; \mathbf{x} + \mathbf{e}_2)$. Let system A follow the optimal policy and system B imitate all decisions of system A. Because both systems have the same number of customers in all stations, B can always imitate system A. We couple the additional class-1 customer in system A with the additional class-2 customer, as well as all other service and interarrival times. Therefore, both additional customers finish their services with probability μ_1 , coupling the two systems with a difference of $r_1 - r_2$ in rewards. With probability $\mu_2 - \mu_1$, only the additional class-2 customer departs, so that system B moves to $(\mathbf{y}; \mathbf{x})$ with an additional reward of r_2 , whereas system A remains in its current state. Then:

$$\begin{aligned}
D_0^{n+1}(12)(\mathbf{y}; \mathbf{x}) &\leq \mu_1(r_1 - r_2) + (\mu_2 - \mu_1)(D_0^n(10)(\mathbf{y}; \mathbf{x}) - r_2) \\
&\quad + (1 - \mu_2 - \beta) \min_{\mathbf{s} + \mathbf{e}_3 \in \mathcal{S}} \{D_0^n(12)(\mathbf{s})\} \\
&\leq \mu_1(r_1 - r_2) + (\mu_2 - \mu_1)(r_1 - r_2) + (1 - \mu_2 - \beta)(R_1 - R_2) \leq R_1 - R_2,
\end{aligned}$$

where the first inequality is due to the policy of system B, the coupling and uniformization, and the second follows from the induction hypothesis, part (ii) and some algebra. Note that $R_j(\mu_j + \beta) = r_j\mu_j$ and $R_j \leq r_j$. \square

Part (i) of Lemma 1 proves $D_j^n(j0)(\mathbf{y}; \mathbf{x}) \geq 0$ for all states and for all possible parameters, which implies that it is always better to accept a class- j customer to station j rather than rejecting him. We also have $D_{j0}^n(jj)(\mathbf{y}; \mathbf{x}) \geq 0$ by part (iii) of Lemma 1, so that the optimal policy will always prefer to accept a class- j customer to station j rather than accepting her to station 0. These two results together specify the optimal admission rules for stations 1 and 2 due to the optimality equations (2):

Theorem 1 *If $y_j < c_j$, i.e., whenever station j has available servers, then it is optimal to admit a class- j customer to station j .*

5 Existence of a preferred class in station 0

A consequence of Theorem 1 is that admission of customers to station 0 should be considered only after the dedicated station of class j becomes completely full. Then station 0 behaves like a stochastic knapsack receiving an arrival stream equivalent to the overflow process of stations 1 and 2. Altman et al. [3] and Örmeci et al. [14] analyze similar stochastic knapsacks when arrivals occur according to a Poisson process. Örmeci et al. [14] define a customer class as “preferred”, in the sense that its customers are always admitted to the system if there are free servers, regardless of the congestion level. We adopt their terminology for station 0 so that class j is preferred if it is always optimal to admit

class- j customers to station 0 whenever station 0 has at least one available server while $y_j = c_j$. In this section, we derive sufficient conditions for each class to be preferred. In determining preferred class(es), two different criteria can be considered, immediate rewards, R_j , and average rewards, $r_j\mu_j$. That is, we would expect that whenever the corresponding reward of class j is higher than that of the other class, class j is preferred. We will discuss the validity of both criteria later in conjunction with the related results.

From optimality equations (2), it is clear that having $D_0^n(j0)(\mathbf{y}; \mathbf{x}) \geq 0$ for all $(\mathbf{y}; \mathbf{x} + \mathbf{e}_1) \in \mathcal{S}$ guarantees that it is always better to serve a class- j customer at station 0 rather than rejecting him. Our first result present the sufficient conditions for $D_0^n(20)(\mathbf{y}; \mathbf{x}) \geq 0$ for all $(\mathbf{y}; \mathbf{x} + \mathbf{e}_1) \in \mathcal{S}$:

Lemma 2 *If $\lambda_1 R_1 \leq (\lambda_1 + \mu_2 + \beta)R_2$, then for all $(\mathbf{y}; \mathbf{x} + \mathbf{e}_1) \in \mathcal{S}$: $0 \leq D_0^n(20)(\mathbf{y}; \mathbf{x})$ for all $n \geq 0$.*

Proof. Let $\lambda_1 R_1 \leq (\lambda_1 + \mu_2 + \beta)R_2$. We use induction to prove the result. The function u_0 defined by equation (4) satisfies the inequality. So assume that the statement is also true for period n , and consider period $n + 1$. Let system A be in state $(\mathbf{y}; \mathbf{x} + \mathbf{e}_2)$ and system B in $(\mathbf{y}; \mathbf{x})$ in period $n + 1$. If an arrival occurs while $y_j < c_j$, both systems accept the new customer to station j , keeping the difference between the two systems the same. Hence, consider an arrival when $y_j = c_j$: System A denies service at station 0 to the new customers in period $n + 1$, while system B takes the optimal actions. If system B also rejects the incoming customer, both systems remain in their current states, preserving the extra class-2 customer. Acceptance of a class-1 customer to system B leads the two systems to two different states $(\mathbf{y}; \mathbf{x} + \mathbf{e}_2)$ and $(\mathbf{y}; \mathbf{x} + \mathbf{e}_1)$. If a class-2 customer is admitted to system B, then the two systems couple with no difference in reward. With the departure of the additional class-2 customer in system A, the systems enter the same state with a return of r_2 for system A, whereas all other service completions keep the extra class-2 customer in system A. Then:

$$\begin{aligned} D_0^{n+1}(20)(\mathbf{y}; \mathbf{x}) &\geq \lambda_1 \min \left\{ \min_{\mathbf{s} + \mathbf{e}_3 \in \mathcal{S}} \{D_0^n(20)(\mathbf{s})\}, D_0^n(21)(\mathbf{y}; \mathbf{x}) \right\} \\ &\quad + \lambda_2 \min \left\{ \min_{\mathbf{s} + \mathbf{e}_3 \in \mathcal{S}} \{D_0^n(20)(\mathbf{s})\}, 0 \right\} \\ &\quad + \mu_2 r_2 + \left(\sum_{j=1,2} c_j \mu'_j + (c-1)\mu_2 \right) \min_{\mathbf{s} + \mathbf{e}_3 \in \mathcal{S}} \{D_0^n(20)(\mathbf{s})\} \\ &\geq \lambda_1 \min\{0, R_2 - R_1\} + r_2 \mu_2 \end{aligned}$$

where the first inequality is due to the coupling, and the second follows from the induction hypothesis, and part (iv) of Lemma 1. If $R_2 \geq R_1$, then the statement is proven. Otherwise, we have:

$$D_0^{n+1}(20)(\mathbf{y}; \mathbf{x}) \geq \lambda_1(R_2 - R_1) + (\mu_2 + \beta)R_2 = (\lambda_1 + \mu_2 + \beta)R_2 - \lambda_1 R_1 \geq 0,$$

where the last inequality is due to the assumption of the theorem. \square

This lemma shows that under the specified condition it is always better to accept a class-2 customer to station 0 rather than rejecting her. We know from Theorem 1 that station 0 should admit only the overflow of stations 1 and 2. Hence, we can conclude the following:

Theorem 2 *If $y_2 = c_2$ and $\lambda_1 R_1 \leq (\lambda_1 + \mu_2 + \beta)R_2$, then it is optimal to accept a class-2 customer to station 0 whenever it has available server(s), i.e., class 2 is preferred.*

Theorem 2 shows that when the rewards satisfy the given condition, it is optimal to accept class-2 customers to station 0 regardless of the precise state of station 0 or of station 1: It is enough to have at least one available server at station 0 while station 2 is completely busy. Furthermore, if R_2 , the immediate reward of class 2 (fast class), is sufficiently high compared to that of class 1, then class 2 is preferred in station 0. Obviously, class-1 (slow class) customers would not be necessarily preferred in station 0 even if their immediate reward is higher, i.e., $R_1 > R_2$, as we may have $\mu_1 \ll \mu_2$. Hence, the criterion of immediate rewards favors class 2 by emphasizing its superiority of “being fast”.

We establish a similar condition for class-1 customers to be always accepted to station 0. For this, we need to derive lower bounds on the effect of changing a class-1 customer to a class-2 customer in station 0 together with the effect of an additional class-1 customer.

Lemma 3 *If $\lambda_2 r_2 \mu_2 \leq (\lambda_2 + \mu_2 + \beta) r_1 \mu_1$, then for all $(\mathbf{y}; \mathbf{x} + \mathbf{e}_1) \in \mathcal{S}$ and for all n :*

$$(i) D_0^n(10)(\mathbf{y}; \mathbf{x}) \geq 0.$$

$$(ii) D_0^n(12)(\mathbf{y}; \mathbf{x}) \geq \frac{r_1 \mu_1 - r_2 \mu_2}{\mu_2 + \beta}.$$

Proof. We use induction on the number of transitions, n . Both statements are satisfied for u^0 defined by (4). Assume that both are true for n .

(i) Assume that system A is in state $(\mathbf{y}; \mathbf{x} + \mathbf{e}_1)$ and system B is in $(\mathbf{y}; \mathbf{x})$ in period $n + 1$. If an arrival occurs while $y_j < c_j$, both systems accept the new customer to station j , keeping the difference between the two systems the same. So we consider an arrival when $y_j = c_j$: We let system A reject all customers and system B follow the optimal policy. If upon an arrival system B also rejects the incoming customer, both systems remain in their current states. Acceptance of a class-1 customer to system B leads both systems to enter the same state. If a class-2 customer is admitted to system B, then the systems move to two different states $(\mathbf{y}; \mathbf{x} + \mathbf{e}_1)$ and $(\mathbf{y}; \mathbf{x} + \mathbf{e}_2)$. With the departure of the additional class-1 customer in system A, the systems enter the same state with an additional reward of r_1 for system A, whereas all other service completions keep the difference between the two systems the same. Then:

$$\begin{aligned} D_0^{n+1}(10)(\mathbf{y}; \mathbf{x}) &\geq \lambda_1 \min \left\{ \min_{\mathbf{s} + \mathbf{e}_3 \in \mathcal{S}} \{D_0^n(10)(\mathbf{s})\}, 0 \right\} \\ &\quad + \lambda_2 \min \left\{ \min_{\mathbf{s} + \mathbf{e}_3 \in \mathcal{S}} \{D_0^n(10)(\mathbf{s})\}, D_0^n(12)(\mathbf{y}; \mathbf{x}) \right\} \\ &\quad + \mu_1 r_1 + \left(\sum_{j=1,2} c_j \mu_j' + (c-1)\mu_2 \right) \min_{\mathbf{s} + \mathbf{e}_3 \in \mathcal{S}} \{D_0^n(10)(\mathbf{s})\} \\ &\geq \lambda_2 \min \left\{ 0, \frac{r_1 \mu_1 - r_2 \mu_2}{\mu_2 + \beta} \right\} + r_1 \mu_1 \end{aligned}$$

where the first inequality is due to coupling, and the second follows from the induction hypotheses for parts (i) and (ii). If $r_1 \mu_1 \geq r_2 \mu_2$, the statement is proven; otherwise:

$$D_0^{n+1}(10)(\mathbf{y}; \mathbf{x}) \geq r_1 \mu_1 \frac{\lambda_2 + \mu_2 + \beta}{\mu_2 + \beta} - \frac{\lambda_2 r_2 \mu_2}{\mu_2 + \beta} \geq 0,$$

where the last inequality is due to the assumption of the theorem. Thus, the first statement is true for all $x \in \mathcal{S}$ and for all $n \geq 0$.

(ii) Now let system A be in state $(\mathbf{y}; \mathbf{x} + \mathbf{e}_1)$ and system B in $(\mathbf{y}; \mathbf{x} + \mathbf{e}_2)$ in period $n + 1$. System B takes the optimal actions and system A imitates all the actions of system B in this period. We, as in part (iv) of Lemma 1, couple the additional class-2 customer, say customer d_2 , in system B with the additional class-1 customer, say customer d_1 in system A, as well as all other service and interarrival times. Then, if d_1 and d_2 finish their services, which happens with probability μ_1 , the systems couple with a difference in reward, $r_1 - r_2$. The departure of d_2 alone, with probability $\mu_2 - \mu_1$, takes the systems to two different states, $(\mathbf{y}; \mathbf{x} + \mathbf{e}_1)$ and $(\mathbf{y}; \mathbf{x})$ with a difference of $-r_2$. Whenever there is any other transition, both systems continue to have their additional customers. Thus:

$$\begin{aligned} D_0^{n+1}(12)(\mathbf{y}; \mathbf{x}) &\geq \mu_1(r_1 - r_2) + (\mu_2 - \mu_1)(D_0^n(10)(\mathbf{y}; \mathbf{x}) - r_2) \\ &\quad + \left(\sum_{j=1,2} (\lambda_j + c_j \mu_j) + (c-1)\mu_2 \right) \min_{\mathbf{s} + \mathbf{e}_3 \in \mathcal{S}} \{D_0^n(12)(\mathbf{s})\} \\ &\geq r_1\mu_1 - r_2\mu_2 + (1 - \mu_2 - \beta) \frac{r_1\mu_1 - r_2\mu_2}{\mu_2 + \beta} = \frac{r_1\mu_1 - r_2\mu_2}{\mu_2 + \beta} \end{aligned}$$

where the first inequality is due to the coupling and the second follows by uniformization and the induction hypotheses for parts (i) and (ii). This proves the second part of the lemma. \square

This lemma together with Theorem 1 specifies the sufficient conditions for class 1 to be preferred in station 0:

Theorem 3 *If $y_1 = c_1$ and $\lambda_2 r_2 \mu_2 \leq (\lambda_2 + \mu_2 + \beta) r_1 \mu_1$, then it is optimal to accept a class-1 customer to station 0 whenever it has available server(s), i.e., class 1 is preferred.*

Theorem 3 shows that when the average reward, or reward rate, of class 1 (slow-service class) is higher than that of class 2, class 1 is preferred. In a queueing environment with two classes of customers, giving priority to the class with higher average reward would determine a preferred class, as shown in Harrison [8], but in a loss system it favors slower customers: When $r_2 \mu_2 \leq r_1 \mu_1$, class-1 customers generate reward at higher rate and keep the server busy for longer period than class-2 customers do. Hence, class 1 is obviously preferred. Class 2 (fast-service class), on the other hand, is not necessarily preferred if $r_1 \mu_1 < r_2 \mu_2$: Consider a firm which has to choose one from two customers, one of whom brings \$1,000 profit each month for 12 months and the other with a profit of \$1,200 per month for only 3 months. The possibility that the firm will have no customer after 3 months works in favor of the longer duration customer. Hence, class-2 customers with higher average rewards can be rejected if $\mu_1 \ll \mu_2$. The difference between loss systems and queueing systems is due to the fact that a queueing system is not concerned with the ‘‘occupation’’ time of a server since all customers can wait, i.e., no loss of work.

Theorems 2 and 3 present the sufficient conditions for class 2 and class 1, respectively, to be preferred, i.e., to be always admitted to station 0, when stations 2 and 1 are full and station 0 has at least one available server. These conditions are exactly the same as those in Örmeci et al. [14], where the stochastic knapsack receives two Poisson arrivals with rates λ_1 and λ_2 . Hence changing the arrival stream to an overflow process has not changed the admission policy in terms of ‘‘preferred’’ class. Finally, we note that our results are not complete in the sense that there are parameters for which none of the classes is proved to be preferred in station 0. Örmeci et al. [14] include a detailed discussion on the existence of preferred class.

Remark 1 *If the parameter values are such that:*

$$\frac{\lambda_2 + \mu_2 + \beta}{\lambda_2} < \frac{r_2 \mu_2}{r_1 \mu_1} < \frac{\lambda_1 (\mu_2 + \beta)}{(\lambda_1 + \mu_2 + \beta) (\mu_1 + \beta)},$$

then our results cannot guarantee the existence of a preferred class, or in other words, we cannot state that it is optimal to always accept customers of at least one of the classes to station 0.

6 An optimal threshold policy for station 0

Intuitively, we expect that it should be more difficult to accept customers to station 0 when there are many customers already in the system. As station 0 has more customers, the resources shared by both classes decrease, which in turn decreases the marginal benefit of the system from an additional customer in station 0. The first part of our next lemma shows that the benefit of an additional class- j customer in station 0 is decreasing in the number of class- k customers in station 0. As the number of customers in a dedicated station k increases, the probability of overflow in that station also increases. Hence, when station k has more customers, the system becomes more reluctant to admit class- j customers to station 0 since the resources in station 0 may be needed by future class- k customers soon. The second part of Lemma 4 proves that the benefit of an additional class- j customer in station 0 decreases in the number of customers in station k .

Lemma 4 *For all $n \geq 0$ and for $j = 1, 2$ and $k \neq j$:*

(i) *For all $(\mathbf{y}; \mathbf{x} + \mathbf{e}_k + \mathbf{e}_j) \in \mathcal{S}$:*

$$u^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_k + \mathbf{e}_j) - u^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_k) \leq u^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_j) - u^n(\mathbf{y}; \mathbf{x}). \quad (5)$$

(ii) *For all $(\mathbf{y} + \mathbf{e}_k; \mathbf{x} + \mathbf{e}_j) \in \mathcal{S}$:*

$$u^n(\mathbf{y} + \mathbf{e}_k; \mathbf{x} + \mathbf{e}_j) - u^n(\mathbf{y} + \mathbf{e}_k; \mathbf{x}) \leq u^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_j) - u^n(\mathbf{y}; \mathbf{x}). \quad (6)$$

Proof. We prove the statements by induction on the number of remaining arrivals. u^0 given by (4) satisfies both inequalities, so assume that they also hold for n .

(i) We first observe that this inequality implies a seemingly more strong inequality almost immediately. We iterate on $(\mathbf{x} + \mathbf{e}_k)$ in inequality (5), so that:

$$u^n(\mathbf{y}; \mathbf{x} + b_k \mathbf{e}_k + \mathbf{e}_j) - u^n(\mathbf{y}; \mathbf{x} + b_k \mathbf{e}_k) \leq u^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_j) - u^n(\mathbf{y}; \mathbf{x}),$$

for all $b_k \geq 0$ with $(\mathbf{y}; \mathbf{x} + b_k \mathbf{e}_k + \mathbf{e}_j) \in \mathcal{S}$, which can be rewritten as follows:

$$u^n(\mathbf{y}; \mathbf{x} + b_k \mathbf{e}_k + \mathbf{e}_j) - u^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_j) \leq u^n(\mathbf{y}; \mathbf{x} + b_k \mathbf{e}_k) - u^n(\mathbf{y}; \mathbf{x}). \quad (7)$$

Now we iterate on $(\mathbf{y}; \mathbf{x} + \mathbf{e}_j)$ in (7) to obtain:

$$u^n(\mathbf{y}; \mathbf{x} + b_k \mathbf{e}_k + b_j \mathbf{e}_j) - u^n(\mathbf{y}; \mathbf{x} + b_j \mathbf{e}_j) \leq u^n(\mathbf{y}; \mathbf{x} + b_k \mathbf{e}_k) - u^n(\mathbf{y}; \mathbf{x}), \quad (8)$$

for all $b_j \geq 0$ and $b_k \geq 0$ with $(\mathbf{y}; \mathbf{x} + b_k \mathbf{e}_k + b_j \mathbf{e}_j) \in \mathcal{S}$.

Now we want to show that v_i^n 's also satisfy this monotonicity. Hence we define:

$$\delta_i^n = v_i^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_k + \mathbf{e}_j) - v_i^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_k) - v_i^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_j) + v_i^n(\mathbf{y}; \mathbf{x}).$$

Let systems A, B, C and D be systems starting from states $(\mathbf{y}; \mathbf{x} + \mathbf{e}_k + \mathbf{e}_j)$, $(\mathbf{y}; \mathbf{x} + \mathbf{e}_k)$, $(\mathbf{y}; \mathbf{x} + \mathbf{e}_j)$, and $(\mathbf{y}; \mathbf{x})$, respectively, in period n . Since this inequality is symmetric in j and k , it is enough to consider only one of these indices, say $i = k$: If a class- k arrival occurs when $y_k < c_k$, then all four systems accept the customer to station k so that v_k^n 's satisfy the inequality by the induction hypothesis immediately. Hence, consider a class- k arrival when the system is in a state with $y_k = c_k$: Systems A and D take the optimal actions with a_A and a_D being the optimal number of customers accepted to station 0 in system A and D, respectively. Then, we let system B accept a_A customers and system C admit a_D customers to station 0. Notice that systems B and C can always behave as described since system D already has one more customer in station 0 than these two systems, i.e., they both have at least one free server. We have:

$$\begin{aligned} \delta_j^n &\leq u^n(\mathbf{y}; \mathbf{x} + (1 + a_A)\mathbf{e}_k + \mathbf{e}_j) - u^n(\mathbf{y}; \mathbf{x} + a_D\mathbf{e}_k + \mathbf{e}_j) \\ &\quad - u^n(\mathbf{y}; \mathbf{x} + (1 + a_A)\mathbf{e}_k) + u^n(\mathbf{y}; \mathbf{x} + a_D\mathbf{e}_k) \leq 0, \end{aligned}$$

where the first inequality follows from the optimality of v^n 's, and the second inequality holds due to inequality (8) by taking $(\mathbf{y}; \mathbf{x} + a_D\mathbf{e}_k)$ equal to $(\mathbf{y}; \mathbf{x})$, $b_k = 1 + a_A - a_D$ and $b_j = 1$. Notice that for all possible values of a_A and a_D , $b_k \geq 0$ so that we can use inequality (8).

We couple all these systems so that, except for the additional customers, they all behave in the same way. Moreover, we couple the additional class- k customers in systems A and B, and the additional class- j customers in systems A and C. Then if one of the additional customers departs from one system, the one in the coupled system also departs. Now, consider u^{n+1} 's:

$$\begin{aligned} &u^{n+1}(\mathbf{y}; \mathbf{x} + \mathbf{e}_k + \mathbf{e}_j) - u^{n+1}(\mathbf{y}; \mathbf{x} + \mathbf{e}_k) - u^{n+1}(\mathbf{y}; \mathbf{x} + \mathbf{e}_j) + u^{n+1}(\mathbf{y}; \mathbf{x}) \\ &= \sum_{i=j,k} \left(\lambda_i [v_i^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_k + \mathbf{e}_j) - v_i^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_k) - v_i^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_j) + v_i^n(\mathbf{y}; \mathbf{x})] \right. \\ &+ x_i \mu_i [u^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_k + \mathbf{e}_j - \mathbf{e}_i) - u^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_k - \mathbf{e}_i) - u^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_j - \mathbf{e}_i) + u^n(\mathbf{y}; \mathbf{x} - \mathbf{e}_i)] \\ &+ y_i \mu_i' [u^n(\mathbf{y} - \mathbf{e}_i; \mathbf{x} + \mathbf{e}_k + \mathbf{e}_j) - u^n(\mathbf{y} - \mathbf{e}_i; \mathbf{x} + \mathbf{e}_k) - u^n(\mathbf{y} - \mathbf{e}_i; \mathbf{x} + \mathbf{e}_j) + u^n(\mathbf{y} - \mathbf{e}_i; \mathbf{x})] \\ &+ \mu_k [u^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_j) - u^n(\mathbf{y}; \mathbf{x}) - u^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_j) + u^n(\mathbf{y}; \mathbf{x})] \\ &+ \mu_j [u^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_k) - u^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_k) - u^n(\mathbf{y}; \mathbf{x}) + u^n(\mathbf{y}; \mathbf{x})] \\ &+ \gamma [u^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_k + \mathbf{e}_j) - u^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_k) - u^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_j) + u^n(\mathbf{y}; \mathbf{x})] \\ &\leq 0 \end{aligned}$$

where $\gamma = c\mu_2 - \mu_j - \mu_k + \sum_{i=1}^2 (c_i \mu_i' - x_i \mu_i - y_i \mu_i')$. The first term is less than or equal to 0 since we have shown $\delta_i^n \leq 0$, the second, the third and the last are also non-positive by the induction hypothesis whereas the fourth and fifth terms are 0. Thus, the value functions, u^n , satisfy inequality (5) for all $n \geq 0$.

(ii) We, similarly to the proof of part (i), iterate on $(\mathbf{y} + \mathbf{e}_k; \mathbf{x})$, and then on $(\mathbf{y}; \mathbf{x} + \mathbf{e}_j)$ to obtain the following inequality for $b_j \geq 0$ and $b_k \geq 0$ with $(\mathbf{y} + b_k \mathbf{e}_k; \mathbf{x} + b_j \mathbf{e}_j) \in \mathcal{S}$:

$$u^n(\mathbf{y} + b_k \mathbf{e}_k; \mathbf{x} + b_j \mathbf{e}_j) - u^n(\mathbf{y} + b_k \mathbf{e}_k; \mathbf{x}) \leq u^n(\mathbf{y}; \mathbf{x} + b_j \mathbf{e}_j) - u^n(\mathbf{y}; \mathbf{x}). \quad (9)$$

We want to show that v_i^n 's also satisfy this monotonicity. Hence we define:

$$\delta_i^n = v_i^n(\mathbf{y} + \mathbf{e}_k; \mathbf{x} + \mathbf{e}_j) - v_i^n(\mathbf{y} + \mathbf{e}_k; \mathbf{x}) - v_i^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_j) + v_i^n(\mathbf{y}; \mathbf{x})$$

Let systems A, B, C and D be systems starting from states $(\mathbf{y} + \mathbf{e}_k; \mathbf{x} + \mathbf{e}_j)$, $(\mathbf{y} + \mathbf{e}_k; \mathbf{x})$, $(\mathbf{y}; \mathbf{x} + \mathbf{e}_j)$, and $(\mathbf{y}; \mathbf{x})$, respectively, in period $n + 1$. Unlike part (i), this inequality is not symmetric with respect to j and k , so we need to consider the arrivals of both classes explicitly:

When a class- k customer arrives when $y_k + 1 < c_k$, then all four systems accept the customer to station k , satisfying the inequality trivially by the induction hypothesis. Now assume $y_k + 1 = c_k$: In this case system C and D will follow their optimal policy, thus accepting the incoming customer to station k , whereas system A and B will accept a_A class- k customers to station 0 with a_A being the optimal number of customers that system A accepts to station 0. Then:

$$\begin{aligned} \delta_k^n &\leq u^n(\mathbf{y} + \mathbf{e}_k; \mathbf{x} + a_A \mathbf{e}_k + \mathbf{e}_j) - u^n(\mathbf{y} + \mathbf{e}_k; \mathbf{x} + a_A \mathbf{e}_k) \\ &\quad - u^n(\mathbf{y} + \mathbf{e}_k; \mathbf{x} + \mathbf{e}_j) + u^n(\mathbf{y} + \mathbf{e}_k; \mathbf{x}) \leq 0. \end{aligned}$$

If $a_A = 0$, then the right hand side becomes 0, otherwise, i.e., if $a_A = 1$, then the inequality becomes inequality (5) with $(\mathbf{y}; \mathbf{x}) = (\mathbf{y} + \mathbf{e}_k; \mathbf{x})$ and so it is satisfied by part (i).

If a class- j arrival occurs when $y_j < c_j$, then all four systems accept the customer to station j so that v_j^n 's satisfy the inequality by the induction hypothesis immediately. Hence, consider the states with $y_j = c_j$: We let systems A and D take the optimal actions, and denote the number of customers accepted to station 0 in systems A and D by a_A and a_D , respectively. Then, systems B and C accept a_D and a_A customers to station 0, respectively, which is always feasible since systems B and C have the same number of customers in station 0 with systems D and A, respectively. We have:

$$\begin{aligned} \delta_j^n &\leq u^n(\mathbf{y} + \mathbf{e}_k; \mathbf{x} + (1 + a_A) \mathbf{e}_j) - u^n(\mathbf{y} + \mathbf{e}_k; \mathbf{x} + a_D \mathbf{e}_j) \\ &\quad - u^n(\mathbf{y}; \mathbf{x} + (1 + a_A) \mathbf{e}_j) + u^n(\mathbf{y}; \mathbf{x} + a_D \mathbf{e}_j) \leq 0, \end{aligned}$$

where the first inequality follows from the optimality of v_i^n 's, and the second inequality holds due to inequality (9) by taking $(\mathbf{y}; \mathbf{x} + a_D \mathbf{e}_j)$ equal to $(\mathbf{y}; \mathbf{x})$, $b_k = 1$ and $b_j = 1 + a_A - a_D \geq 0$.

We couple all these systems so that, except for the additional customers, they all behave in the same way. Moreover, we couple the additional class- k customers in systems A and B, and the additional class- j customers in systems A and C, similarly to part (i). Now, consider u^{n+1} 's:

$$\begin{aligned} &u^{n+1}(\mathbf{y} + \mathbf{e}_k; \mathbf{x} + \mathbf{e}_j) - u^{n+1}(\mathbf{y} + \mathbf{e}_k; \mathbf{x}) - u^{n+1}(\mathbf{y}; \mathbf{x} + \mathbf{e}_j) + u^{n+1}(\mathbf{y}; \mathbf{x}) \\ &= \sum_{i=j,k} \left(\lambda_i [v_i^n(\mathbf{y} + \mathbf{e}_k; \mathbf{x} + \mathbf{e}_j) - v_i^n(\mathbf{y} + \mathbf{e}_k; \mathbf{x}) - v_i^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_j) + v_i^n(\mathbf{y}; \mathbf{x})] \right. \\ &+ x_i \mu_i [u^n(\mathbf{y} + \mathbf{e}_k; \mathbf{x} + \mathbf{e}_j - \mathbf{e}_i) - u^n(\mathbf{y} + \mathbf{e}_k; \mathbf{x} - \mathbf{e}_i) - u^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_j - \mathbf{e}_i) + u^n(\mathbf{y}; \mathbf{x} - \mathbf{e}_i)] \\ &+ y_i \mu_i' [u^n(\mathbf{y} + \mathbf{e}_k - \mathbf{e}_i; \mathbf{x} + \mathbf{e}_j) - u^n(\mathbf{y} + \mathbf{e}_k - \mathbf{e}_i; \mathbf{x}) - u^n(\mathbf{y} - \mathbf{e}_i; \mathbf{x} + \mathbf{e}_j) + u^n(\mathbf{y} - \mathbf{e}_i; \mathbf{x})] \\ &+ \mu_j [u^n(\mathbf{y} + \mathbf{e}_k; \mathbf{x}) - u^n(\mathbf{y} + \mathbf{e}_k; \mathbf{x}) - u^n(\mathbf{y}; \mathbf{x}) + u^n(\mathbf{y}; \mathbf{x})] \\ &+ \mu_k' [u^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_j) - u^n(\mathbf{y}; \mathbf{x}) - u^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_j) + u^n(\mathbf{y}; \mathbf{x})] \\ &+ \gamma [u^n(\mathbf{y} + \mathbf{e}_k; \mathbf{x} + \mathbf{e}_j) - u^n(\mathbf{y} + \mathbf{e}_k; \mathbf{x}) - u^n(\mathbf{y}; \mathbf{x} + \mathbf{e}_j) + u^n(\mathbf{y}; \mathbf{x})] \\ &\leq 0 \end{aligned}$$

where $\gamma = c\mu_2 - \mu_j - \mu'_k + \sum_{i=1}^2 (c_i\mu'_i - x_i\mu_i - y_i\mu'_i)$. The first term is less than or equal to 0 since we have shown $\delta_j^n \leq 0$, the second, the third, and the last are also non-positive by the induction hypothesis whereas the fourth and the fifth terms are 0. Thus, the value functions, u^n , satisfy inequality (6) for all n . \square

Part (i) of Lemma 4 guarantees the existence of optimal thresholds regarding the admission control of class- j customers, while part (ii) establishes the monotonicity of these thresholds in the number of customers in station k :

Theorem 4 For $j = 1, 2$, there exist numbers $\{l_j^n(\mathbf{y}; 0), \dots, l_j^n(\mathbf{y}; c-1)\}$ such that it is optimal to accept a class- j customer to station 0 if $y_j = c_j$ and $x_k < l_j^n(\mathbf{y}; x_j)$, and to reject her otherwise, so:

$$a_j^n(\mathbf{y}; \mathbf{x}) = \begin{cases} (\mathbf{y}; \mathbf{x} + \mathbf{e}_j) & : x_k < l_j^n(\mathbf{y}; x_j) \\ (\mathbf{y}; \mathbf{x}) & : \text{otherwise} \end{cases}$$

with $k \neq j$. Moreover, $l_j^n(\mathbf{y}; m) \geq l_j^n(\mathbf{y} + \mathbf{e}_k; m)$.

The first part of this result, i.e., the existence of thresholds, extends the results of Altman et al. [3] on the existence of optimal thresholds in the admission control problem for stochastic knapsacks receiving Poisson arrivals to the same problem for stochastic knapsacks with an overflow arrival process. Indeed, we expect that optimal thresholds for class j are also monotone in the number of class- j customers in station 0. However, this would typically require concavity of value functions in x_j when x_k and \mathbf{y} are kept fixed, which we could not show due to the boundary effects and multiple servers. Note that even with regular stochastic knapsacks, it is difficult to show concavity, as Örmeci et al. [14] are able to show it only under very restrictive conditions.

7 Conclusion

In this paper, we have analyzed the dynamic admission control of a loss system with one shared and two dedicated stations. The main purpose of this analysis is to provide insight on the design and control of large call centers receiving two classes of customers. However, this kind of models is also encountered in certain telecommunications and manufacturing systems, where the assumption of no waiting room becomes critical. Recent trends in these systems, such as just-in-time manufacturing and wider use of synchronous services on the internet, e.g., real-time video or audio, demand for loss models with admission control (see for example [10]). Our model is suitable for systems, which utilize admission restrictions rather than scheduling and routing in order to control the amount of work in the system. Still, our future research plan includes considering finite buffers as well as more than two customer classes that will widen applicability of our model.

References

- [1] O. Z. Akşın and P. T. Harker. Modeling a phone center: Analysis of a multi-channel multi-resource processor shared loss system. *Management Science*, 47:324–336, 2001.
- [2] O. Z. Akşın and P. T. Harker. Capacity sizing in the presence of a common shared resource: Dimensioning an inbound call center. *European Jour. Oper. Res.*, to appear.

- [3] E. Altman, T. Jiménez, and G. M. Koole. On optimal call admission control in a resource-sharing system. *IEEE Trans. on Communications*, 49:1659–1668, 2001.
- [4] S. Bhulai and G. M. Koole. A queueing model for call blending in call centers. *IEEE Trans. on Automatic Cont.*, to appear.
- [5] N. Gans, G. M. Koole, and A. Mandelbaum. Telephone call centers: a tutorial and literature review. *European Jour. Oper. Res.*, to appear.
- [6] N. Gans and Y. P. Zhou. A call routing problem with service-level constraints. *Operations Research*, to appear.
- [7] O. Garnett and A. Mandelbaum. *An introduction to skills-based routing and its operational complexities*. Teaching note, Technion, 2001.
- [8] J. M. Harrison. Dynamic scheduling of a multiclass queue: Discount optimality. *Operations Research*, 23:270–282, 1975.
- [9] G. M. Koole. Structural results for the control of queueing systems using event-based dynamic programming. *Queueing Systems*, 30:323–339, 1998.
- [10] C. Ku and S. Jordan. Access control to two multiserver loss queues in series. *IEEE Transactions on Automatic Control*, 42:1017–1023, 1997.
- [11] S. A. Lippman. Applying a new device in the optimization of exponential queueing systems. *Operations Research*, 23:687–710, 1975.
- [12] S. A. Lippman and S. M. Ross. The streetwalker’s dilemma: A job shop model. *SIAM J. Appl. Math.*, 20:336–342, 1971.
- [13] B. Miller. A queueing reward system with several customer classes. *Management Science*, 16:234–245, 1971.
- [14] E. L. Örmeci, A. Burnetas, and J. van der Wal. Admission policies for a two class loss system. *Stochastic Models*, 17:513–540, 2001.
- [15] K. W. Ross. *Multiservice Loss Models for Broadband Telecommunication Networks*. Springer-Verlag, Great Britain, 1995.
- [16] R. A. Shumsky. *Approximation and analysis of a queueing system with flexible and specialized servers*. Working Paper, University of Rochester, 2000.
- [17] D. A. Stanford and W. K. Grassman. Bilingual server call centres. In D. R. McDonald and S. R. E. Turner, editors, *Analysis of communication networks: Call centres, traffic and performance*. Fields Institute Communications 28: 31-48, 2000.
- [18] J. Walrand. *Introduction to Queueing Networks*. Prentice Hall, Englewood Cliffs, N.J., 1988.
- [19] S. H. Xu, R. Righter, and J. G. Shanthikumar. Optimal dynamic assignment of customers to heterogeneous servers in parallel. *Operations Research*, 40:1126–1138, 1992.