Probability in the Engineering and Informational Sciences (2024), **38:**1 130–149
doi:10.1017/S0269964823000013

CAMBRIDGE
UNIVERSITY PRESS

RESEARCH ARTICLE

Optimal control of supervisors balancing individual and joint responsibilities

Zhuoting Yu, Sigrún Andradóttir and Hayriye Ayhan

H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, 755 Ferst Drive NW, Atlanta, GA 30332-0205, USA. E-mail: hayhan@isye.gatech.edu.

Keywords: Abandonment, Markov decision processes, Optimal server allocation, Pooling systems, Profit maximization, Two-stage service

Abstract

We consider a two-stage service system with two types of servers, namely subordinates who perform the first-stage service and supervisors who have their own responsibilities in addition to collaborating with the subordinates on the second-stage service. Rewards are earned when first- or second-stage service is completed and when supervisors finish one of their own responsibilities. Costs are incurred when impatient customers abandon without completing the second-stage service. Our problem is to determine how the supervisors should distribute their time between their joint work with the subordinates and their own responsibilities. Under the assumptions that service times at both stages are exponentially distributed and that the customers waiting for second-stage service abandon after an exponential amount of time, we prove that one of two policies will maximize the long-run average profit. Namely, it is optimal for supervisors to start collaborating with subordinates either when subordinates can no longer serve new customers or as soon as there is a customer ready for second-stage service. Furthermore, we show that the optimality condition is a simple threshold on the system parameters. We conclude by proving that pooling supervisors (and their associated subordinates) improves system performance, but with limited returns as more supervisors are pooled.

1. Introduction

Consider a system with $N \geq 1$ subordinates and $1 \leq M \leq N$ supervisors. Assume that there is an unlimited supply of work, that each customer requires two stages of service, and that customers are impatient and can leave without receiving the service at the second stage. The first service stage is completed by an assigned subordinate, whereas the second service stage (also referred to as the advanced service) is completed jointly by the assigned subordinate and a supervisor. The subordinate can only start work on a new customer when her previously assigned customer departs. Thus, the subordinate will serve the customer on her own (the first-stage service), wait for a supervisor together with the customer, and then serve the customer together with a supervisor (the second-stage service) if the customer does not leave before the second-stage service starts. In addition to their work with the subordinates, the supervisors have an unlimited supply of their own responsibilities to attend to. Therefore, supervisors may not immediately attend to waiting customers. Our problem is to determine how the supervisors should dynamically divide their time between their joint work with the subordinates and their own responsibilities. Rewards are incurred both when first- and second-stage service is completed for a customer and also when supervisors finish one of their own responsibilities. However, a cost is incurred when customers depart without completing the second-stage service. Our objective is to maximize the long-run average profit per unit time.

Our research is motivated by situations where supervisors must sign off on the work of their subordinates. However, this type of queueing systems may arise in other real-life situations. For example, consider some government service where people need to fill out forms or take other actions by themselves as the first stage. The second stage involves being served by the officials. The limited place for people to finish the self-service can be regarded as the limited number of (subordinate) servers in the first stage. Since there are typically fewer officials than people needing service, and since the officials may have other responsibilities, customers may get impatient and leave while waiting for officials. The assumption on unlimited supply of work for both the first-stage service and the supervisors is consistent with real-life observations and common assumptions in the literature on modern service and manufacturing systems. Specifically, in healthcare facilities, for example, emergency rooms, it is widely acknowledged that overcrowding is prevalent at all times (see, e.g., [12,13,28]), indicating the rationality of the assumption on the unlimited supply of patients in such healthcare facilities. Similarly, in make-to-stock manufacturing systems, it is common to assume ample availability of raw materials (see, e.g., [21,27,31]).

We assume that the amount of time that it takes a supervisor to switch from one activity to the other is negligible. Furthermore, we assume that the service times of each customer in the first stage are exponential random variables with rate $\mu_1 > 0$. The patience time of a customer for the second-stage service is exponentially distributed with rate $\theta > 0$. The corresponding abandonment cost is c. We first assume that abandonments can only occur when the customers are waiting for the second-stage service. Later, we extend the problem to the case where abandonments may also occur during the first- and/or second-stage service. The second-stage service time is exponentially distributed with rate $\mu_2 > 0$. The amounts of time that the supervisors spend on their own responsibilities have an exponential distribution with rate $\mu_s > 0$. We assume that supervisors can switch between tasks in a preemptive manner (rather than only upon completing a task). Finally, all random variables are independent. There is a reward of $r_1 \geq 0$ when a subordinate completes the first-stage service and a reward of $r_2 \geq 0$ when a supervisor and a subordinate complete the second-stage service together. There is also a reward of $r_s \geq 0$ when a supervisor finishes one of her own responsibilities. The abandonment cost c is not restricted to be positive; when c is negative, it can be regarded as the reward for a customer who left the system with the first-stage service only. Note that without loss of generality, we can always set $r_1 = 0$ since the case where $r_1 > 0$ is equivalent to the case where $r_1' = 0$, $c' = c - r_1$, and $r_2' = r_1 + r_2$. The remainder of this paper considers the case where $r_1 = 0$.

For this service system, we are interested in determining the dynamic assignment of the supervisors to their two tasks with the objective of maximizing the long-run average profit. Controlling flexible servers in tandem queueing systems has been studied in many papers. For example, Duenyas *et al.* [16] considered the optimal control of a tandem queueing system with setups where there is only one flexible server and Ahn *et al.* [1] studied the optimal control of two flexible servers in a two-stage tandem queueing system to minimize holding costs. Andradóttir and Ayhan [2] characterized the optimal assignment of M flexible servers to two stations in a tandem queueing system with the objective to maximize the long-run average throughput, and Andradóttir *et al.* [4] considered the assignment of flexible servers in a tandem queueing network with N stations and several dedicated servers. Berman and Sapna-Isotupa [8] studied the optimal server allocation between the front and back rooms of a service facility when the work in the back room is generated by the service provided in the front room and the servers are cross-skilled. The above server assignment problems mainly focus on how to assign servers between different stations within the queueing system, while our work considers the server assignment problem between the queue and other responsibilities.

Moreover, server assignment problems are also seen in call centers with call blending, as well as in other practical applications. Motivated by a Bell Canada call center, Deslauriers *et al.* [14] proposed five Markovian models with inbound and outbound calls where there are two types of servers, that is, inbound-only and blend servers, and compared these models with a benchmark model using simulation. When there are two types of jobs served by a common pool of servers and there is a waiting time constraint on one type of jobs, Bhulai and Koole [9] showed that a trunk reservation policy is optimal for the case where the service rates are the same for the two types of jobs. That is, the optimal server assignment policy is a threshold policy on the number of available servers. Furthermore, Bhulai *et al.* [10] extended the assignment problem in call centers to the case where there is no specific condition

https://doi.org/10.1017/S0269964823000013 Published online by Cambridge University Press

on the service rates, and proposed a stochastic approximation algorithm to find the optimal balanced policy. Pang and Perry [22] proposed a logarithmic safety-staffing rule, combined with a threshold policy, under which the server utilization can be close to 1. Wang *et al.* [24] analyzed an $M/M/c$ queue with two priority classes by reducing the two-dimensional Markov chain to a one-dimensional Markov chain. Meanwhile, in the setting of assigning homecare employees to patients, Koeleman *et al.* [19] showed that a trunk reservation heuristic is close to optimal. Compared with the above work on server assignment, our work involves a tandem queueing system, where a customer and a subordinate will wait together for the second-stage service with a supervisor, and the customer may abandon while waiting.

Abandonment is a natural and ubiquitous phenomenon in queueing systems. We include customer abandonments in our model to reflect this common phenomenon. For example, Garnett *et al.* [17] pointed out that customer abandonment is a key factor for call center operations. Weerasinghe and Mandelbaum [26] studied the trade-off between abandonment and blocking in a one-stage, many-server queue where customers may abandon while waiting for service and will balk once the queue is full. Batt and Terwiesch [7] conducted an empirical study on queue abandonments in a hospital emergency department and identified that the abandonment is correlated with the queue length and queue flows during the waiting exposure.

Abandonment is also considered in two-station tandem queues. For example, Khudyakov *et al.* [18] considered a two-stage queueing system in a call center with Interactive Voice Response (IVR). The customer is served by an IVR processor in the first stage and may leave the system with probability $1 - p$ before proceeding to the second stage. Operational performance measures are approximated in an asymptotic Quality and Efficiency Driven regime. Wang *et al.* [25] evaluated the performance of a tandem queueing network with abandonment using an exact numerical method. Zayas-Cabán *et al.* [29] investigated the server assignment problem between the two stations of a tandem service system with abandonment in both stations. Zayas-Caban *et al.* [30] modeled the triage and treatment processes in an emergency department as a two-phase service system where patients may leave the system without treatment. They provided numerical examples to analyze the rewards and patient waiting times under the policy that treatment is prioritized unless there are K or more patients in triage. However, none of the above works on tandem queueing systems with abandonment characterized the optimal policy explicitly, while we provide an optimal threshold policy with respect to the abandonment cost for a tandem queueing system. Additionally, Atar *et al.* [6] considered a multi-class queueing system with homogeneous servers and abandonment, and provided a server-scheduling policy that is asympotically optimal for minimizing the long-run average holding cost. Ansari *et al.* [5] studied a multi-class queueing system with a single server and abandonment, and characterized the conditions under which the asymptotically optimal policy of Atar *et al.* [6] is indeed optimal. Down *et al.* [15] identified the optimal server control in a two-class service system with abandonments, where they considered two models with different reward/cost structures. However, Atar *et al.* [6], Ansari *et al.* [5], and Down *et al.* [15] all considered single-stage queueing systems, whereas our model is a two-stage service system.

Moreover, most of the related work focuses on allocating flexible servers over time to different stations while we focus on the assignment of the supervisors who have other responsibilities in addition to serving the queueing system. In our model, the supervisors work together with the subordinates in the second stage. Motivated by a healthcare application, Andradóttir and Ayhan [3] considered a two-stage service system where the first stage is the examination of patients done by residents and the second stage is the consultation between residents and their (one) attending physician. By comparison, we consider multiple supervisors, customer abandonments, and a different cost structure (abandonment costs rather than holding costs). The comparison between dedicated versus pooled systems has also been investigated and quantified in many research papers. Cattani and Schmidt [11] reviewed and summarized the related work regarding the effects of pooling. We study the performance of dedicated versus pooled systems in this setting (with collaboration between subordinates and supervisors and abandonments) and show that pooling supervisors (and their subordinates) improves performance.

The remainder of this paper is organized as follows. In Section 2, we provide a Markov decision process formulation of the problem and translate the continuous-time optimization problem into a

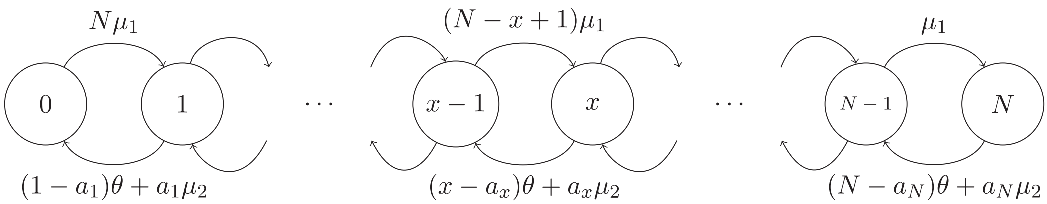


Figure 1. *State-transition diagram for the two-stage service system.*

discrete-time Markov decision process problem. In Section 3, we show that one of two policies is optimal and the optimal policy is defined by a threshold on the abandonment cost c. We also determine the limit of this threshold as the abandonment rate becomes small or large. In Section 4, we prove that pooling supervisors (and their associated subordinates) improves the system performance, but the improvement per pooled supervisor is bounded. Section 5 concludes the paper.

2. Problem formulation

In this section, we consider the stochastic process $\{\mathcal{X}_\pi(t) : t \geq 0\}$ where Π is the set of possible supervisor assignment policies, $\pi \in \Pi$, and $\mathcal{X}_\pi(t) = x \in X = \{0, 1, \ldots, N\}$ is the number of customers who have been served by a subordinate and are waiting for a supervisor at time t under policy π. We assume that Π consists of all Markovian stationary deterministic policies corresponding to the state space X of the stochastic process $\{\mathcal{X}_\pi(t)\}$. The policy $\pi \in \Pi$ specifies if each supervisor is serving the customers or working on her own responsibilities as a function of the current state $x \in X$ (i.e., the number of customers who are waiting for the supervisor). We note that $\{\mathcal{X}_\pi(t)\}$ is a birth-and-death process with finite state space X and there exists a finite scalar q such that the transition rates $\{q_\pi(x, x')\}$ of $\{\mathcal{X}_\pi(t)\}$ satisfy $\sum_{x' \in X, x' \neq x} q_\pi(x, x') \leq q$ for all $x \in X$ and $\pi \in \Pi$. This indicates that $\{\mathcal{X}_\pi(t)\}$ is uniformizable for all $\pi \in \Pi$. Let $\{Y_\pi(k)\}$ denote the corresponding discrete-time Markov chain, so that $\{Y_\pi(k)\}$ has the same state space X as $\{\mathcal{X}_\pi(t)\}$ and transition probabilities $p_\pi(x, x') = q_\pi(x, x')/q$ if $x' \neq x$ and $p_\pi(x, x) = 1 - \sum_{x' \in X, x' \neq x} q_\pi(x, x')/q$ for all $x \in X$. We then translate the continuous-time optimization problem to a discrete-time Markov decision problem (see, e.g., [20]). That is to say, we can generate sample paths of $\{\mathcal{X}_\pi(t)\}$, where $\pi \in \Pi$, by generating a Poisson process $\{K(t)\}$ with rate $q = N\mu_1 + N\theta + M\mu_2 < \infty$ and at the times of events of $\{K(t)\}$, the next state of $\{\mathcal{X}_\pi(t)\}$ is generated using the transition probabilities of $\{Y_\pi(k)\}$.

Let $a \in A = \{0, 1, \ldots, M\}$ denote the assignment of supervisors, where a represents the number of supervisors who are working with the subordinates and A is the action space. Let A_x and a_x denote the set of allowable actions in state $x \in X$. Note that $A_0 = \{0\}$, representing that supervisors can only work on their own responsibilities when there are no customers waiting for the second-stage service, and $A_N = \{1, \ldots, M\}$, representing that we have at least one supervisor serving the customers if the number of customers who are waiting for supervisors attains the maximum (since it would be unethical for supervisors not to serve customers when all the first-stage servers cannot serve any more customers). For $1 \leq x \leq N - 1$, we have $A_x = \{0, 1, \ldots, \min\{x, M\}\}$. Figure 1 illustrates the corresponding rate diagram when action $a_x \in A_x$ is selected in state $x \in X$.

For the discrete-time Markov decision process problem with uniformization constant q, we have, for all $a_x \in A_x$, the following one-step transition probabilities:

$$p(x' \mid x, a_x) = \begin{cases} \dfrac{(N-x)\mu_1}{q} & \text{for } x \in \{0, \ldots, N-1\}, x' = x+1, \\ \dfrac{(x-a_x)\theta + a_x\mu_2}{q} & \text{for } x \in \{1, \ldots, N\}, x' = x-1, \\ 1 - \dfrac{(N-x)\mu_1 + (x-a_x)\theta + a_x\mu_2}{q} & \text{for } x \in \{0, \ldots, N\}, x' = x, \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

https://doi.org/10.1017/S0269964823000013 Published online by Cambridge University Press

Furthermore, for all $x \in X$ and $a_x \in A_x$, we specify the immediate reward $r(x, a_x)$ of choosing action a_x in state x:

$$r(x, a_x) = \frac{(M - a_x)r_s\mu_s + a_x r_2\mu_2 - (x - a_x)c\theta}{q}.$$

Note that due to the abandonments, this Markov decision process problem is unichain. Since X is finite, A_x is finite for each $x \in X$, and $r(x, a_x)$ is bounded, there exists a stationary long-run average optimal policy (see [23], Theorem 8.4.5).

For any policy $\pi \in \Pi$, let $g^{\pi}_{N,M}$ denote the gain (long-run average reward) of the continuous-time problem under policy π for a system with N subordinates and M supervisors. Note that $g^{\pi}_{N,M}/q$ is the gain for the corresponding discrete-time problem. The objective is to identify the optimal policy $\pi^* \in \Pi$ that attains the optimal gain $g^*_{N,M}$, that is, find π^* such that

$$g^{\pi^*}_{N,M} = g^*_{N,M} = \max_{\pi\in\Pi} g^{\pi}_{N,M}.$$

3. Optimal policy

In this section, we show that one of two policies is always optimal and characterize the conditions under which each policy is optimal. Note that a Markovian deterministic decision rule $d : X \to A$ specifies which action $d(x) \in A_x$ to choose in each state $x \in X$. Thus, a stationary policy π can be defined using the corresponding decision rule d which will be denoted as $\pi = d^{\infty}$.

Define $\pi^{\mathcal{S}} = (d_{\mathcal{S}})^{\infty}$, where

$$d_{\mathcal{S}}(x) = \begin{cases} 0 & \text{for } x = 0, \dots, N-1, \\ 1 & \text{for } x = N. \end{cases}$$

Similarly, define $\pi^{\mathcal{C}} = (d_{\mathcal{C}})^{\infty}$ where $d_{\mathcal{C}}(x) = \min\{x, M\}$ for all $x \in \{0, \dots, N\}$. Thus, $\pi^{\mathcal{S}}$ gives priority to the $\mathcal{S}$upervisors' own responsibilities and $\pi^{\mathcal{C}}$ gives priority to the $\mathcal{C}$ustomers. The following theorem completely characterizes the optimal policy.

Theorem 1. *(i) If $c \le c_0 := [(r_s\mu_s - r_2\mu_2)(\theta + \mu_1)]/[(\mu_1 + \mu_2)\theta]$, then $\pi^{\mathcal{S}}$ is optimal;*
(ii) If $c \ge c_0$, then $\pi^{\mathcal{C}}$ is optimal.

Remark 2. It immediately follows from the proof of Theorem 1 that even if the supervisors are not required to serve the customers when there are N customers waiting (i.e., $A_N = A = \{0, 1, \dots, M\}$), a result similar to Theorem 1 remains true. That is to say, if $c \le c_0$, it is optimal for all supervisors to always work on their own responsibilities (even in state N); if $c \ge c_0$, it is optimal for supervisors to start serving customers as soon as there is a customer waiting.

Proof of Theorem 1 and Remark 2. It follows from $1 \le M \le N$ that $N = 1$ implies $M = 1$, in which case $X = \{0, 1\}$, $A_0 = \{0\}$, and $A_1 = \{1\}$. Thus, there is only one feasible policy when $N = 1$ and $\pi^{\mathcal{S}} = \pi^{\mathcal{C}}$ are both optimal. Therefore, we assume $N \ge 2$ in the rest of the proof.

Without loss of generality, we assume that $q = 1$ and use the value iteration algorithm for unichain Markov decision process problems (see p. 364 of [23]).

To prove the optimality of $\pi^{\mathcal{S}}$ and $\pi^{\mathcal{C}}$ under different conditions, for all $x = 0, \dots, N$, we set

$$v_0(x) = (N - x - 1) \times \frac{r_s\mu_s - r_2\mu_2}{\mu_1 + \mu_2}$$

https://doi.org/10.1017/S0269964823000013 Published online by Cambridge University Press

and compute $v_n(x) = \max_{a_x \in A_x} v_n^{a_x}(x)$ for $n \geq 1$, where for $x \in X$ and $a_x \in A_x$,

$$\begin{aligned} v_n^{a_x}(x) = (M - a_x)r_s\mu_s + a_x r_2\mu_2 - (x - a_x)c\theta + (N - x)\mu_1 v_{n-1}(x+1) \\ + [(x - a_x)\theta + a_x\mu_2]v_{n-1}(x-1) \\ + [1 - (N-x)\mu_1 - (x - a_x)\theta - a_x\mu_2]v_{n-1}(x). \end{aligned} \tag{2}$$

Note that since $A_0 = \{0\}$, $v_n(0) = v_n^0(0)$ follows. For $a_x^1, a_x^2 \in A_x$ and $x \in \{1, \ldots, N\}$, define

$$\Delta_n^{a_x^1,a_x^2}(x) = v_n^{a_x^1}(x) - v_n^{a_x^2}(x) = (a_x^2 - a_x^1)(r_s\mu_s - r_2\mu_2 - c\theta + (\theta - \mu_2)[v_{n-1}(x-1) - v_{n-1}(x)]). \tag{3}$$

We first prove part (i). We will show that $\Delta_n^{a_x^1,a_x^2}(x) \geq 0$ for all $n \geq 1$, $x \in \{1, \ldots, N\}$ and $a_x^1 < a_x^2 \in A_x$, which implies that $v_n(x) = v_n^0(x)$ for all $x \in \{1, \ldots, N-1\}$ and $v_n(N) = v_n^1(N)$ ($v_n(N) = v_n^0(N)$ in Remark 2). First assume $\theta = \mu_2$. We then have:

$$\Delta_n^{a_x^1,a_x^2}(x) = (a_x^2 - a_x^1)(r_s\mu_s - r_2\mu_2 - c\theta) \geq 0$$

for all $n \geq 1$ and $x \in \{1, \ldots, N\}$ as long as $c \leq (r_s\mu_s - r_2\mu_2)/\theta = c_0$.

Next assume $\theta \neq \mu_2$. We use induction to prove that $\Delta_n^{a_x^1,a_x^2}(s) \geq 0$ for all $n \geq 1$, $x \in \{1, \ldots, N\}$, and $a_x^1 < a_x^2 \in A_x$. For $n = 1$ and $x \in \{1, \ldots, N\}$, (3) yields

$$\Delta_1^{a_x^1,a_x^2}(x) = (a_x^2 - a_x^1)\left[\frac{(r_s\mu_s - r_2\mu_2)(\theta + \mu_1)}{\mu_1 + \mu_2} - c\theta\right] \geq 0, \tag{4}$$

where the inequality follows since $c \leq c_0$. Now assume that $\Delta_k^{a_x^1,a_x^2}(x) \geq 0$ for $k = 1, \ldots, n-1$, $x \in \{1, \ldots, N\}$ and $a_x^1 < a_x^2$ (i.e., for $k = 1, \ldots, n-1$, $v_k(x) = v_k^0(x)$ for all $x \in \{0, \ldots, N-1\}$ and $v_k(N) = v_k^1(N)$ in Theorem 1; $v_k(N) = v_k^0(N)$ in Remark 2) as long as $c \leq c_0$. We will show that the same assertion holds for $k = n$. From the induction hypothesis, we have

$$v_{n-1}(x) = Mr_s\mu_s - xc\theta + (N-x)\mu_1 v_{n-2}(x+1) + x\theta v_{n-2}(x-1) + [1 - (N-x)\mu_1 - x\theta]v_{n-2}(x)$$

for $x \in \{0, \ldots, N-1\}$ ($x \in \{0, \ldots, N\}$ in Remark 2), and

$$\begin{aligned} v_{n-1}(N) = (M-1)r_s\mu_s + r_2\mu_2 - (N-1)c\theta + [(N-1)\theta + \mu_2]v_{n-2}(N-1) \\ + [1 - (N-1)\theta - \mu_2]v_{n-2}(N). \end{aligned}$$

Furthermore, it follows from $\Delta_{n-1}^{a_x^1,a_x^2}(x) \geq 0$ and (3) that for $x = 1, \ldots, N$,

$$r_s\mu_s - r_2\mu_2 - c\theta + (\theta - \mu_2)[v_{n-2}(x-1) - v_{n-2}(x)] \geq 0,$$

which implies that

$$(\theta - \mu_2)[v_{n-2}(x-1) - v_{n-2}(x)] \geq -r_s\mu_s + r_2\mu_2 + c\theta. \tag{5}$$

Note that for $x \in \{1, \ldots, N-1\}$ ($x \in \{1, \ldots, N\}$ in Remark 2),

$$\begin{aligned} v_{n-1}(x-1) - v_{n-1}(x) = c\theta + (N-x)\mu_1[v_{n-2}(x) - v_{n-2}(x+1)] \\ + (x-1)\theta[v_{n-2}(x-2) - v_{n-2}(x-1)] \\ + [1 - (N-x+1)\mu_1 - x\theta][v_{n-2}(x-1) - v_{n-2}(x)] \end{aligned} \tag{6}$$

https://doi.org/10.1017/S0269964823000013 Published online by Cambridge University Press

and

$$v_{n-1}(N-1) - v_{n-1}(N) = r_s\mu_s - r_2\mu_2 + (N-1)\theta[v_{n-2}(N-2) - v_{n-2}(N-1)] + [1-(N-1)\theta - \mu_1 - \mu_2][v_{n-2}(N-1) - v_{n-2}(N)]. \tag{7}$$

Note that since $q = N\mu_1 + N\theta + M\mu_2$ and we assumed, without loss of generality, that $q = 1$, we have that $1-(N-x+1)\mu_1 - x\theta$ and $1-(N-1)\theta - \mu_1 - \mu_2$ are positive for $x \in \{1,\ldots,N\}$.

Observe that the multipliers $(N-x)\mu_1$ and $(x-1)\theta$ of $v_{n-2}(x)-v_{n-2}(x+1)$ and $v_{n-2}(x-2)-v_{n-2}(x-1)$ equal zero when $x = N$ and $x = 1$, respectively. Equations (5), (6), and (7) yield

$$\begin{aligned}(\theta-\mu_2)[v_{n-1}(x-1) - v_{n-1}(x)] &\geq c\theta(\theta-\mu_2) + (1-\theta-\mu_1)(-r_s\mu_s + r_2\mu_2 + c\theta)\\ &= (1-\mu_1-\mu_2)\theta c + (\theta+\mu_1-1)(r_s\mu_s - r_2\mu_2)\end{aligned}$$

for $x \in \{1,\ldots,N-1\}$ ($x \in \{1,\ldots,N\}$ in Remark 2), and

$$\begin{aligned}(\theta-\mu_2)[v_{n-1}(N-1) - v_{n-1}(N)] &\geq (\theta-\mu_2)(r_s\mu_s - r_2\mu_2) + (1-\mu_1-\mu_2)(-r_s\mu_s + r_2\mu_2 + c\theta)\\ &= (1-\mu_1-\mu_2)\theta c + (\theta+\mu_1-1)(r_s\mu_s - r_2\mu_2).\end{aligned}$$

Now Eq. (3) yields that for all $x = 1,\ldots,N$ and $a_x^1 < a_x^2 \in A_x$,

$$\begin{aligned}\Delta_n^{a_x^1,a_x^2}(x) &\geq (a_x^2 - a_x^1)[r_s\mu_s - r_2\mu_2 - c\theta + (1-\mu_1-\mu_2)\theta c + (\theta+\mu_1-1)(r_s\mu_s - r_2\mu_2)]\\ &= (a_x^2 - a_x^1)[-(\mu_1+\mu_2)c\theta + (\theta+\mu_1)(r_s\mu_s - r_2\mu_2)] \geq 0\end{aligned} \tag{8}$$

as long as $c \leq c_0$.

From (8), we have $\Delta_n^{a_x^1,a_x^2}(x) \geq 0$ for all $n \geq 1, x \in \{1,\ldots,N\}$, and $a_x^1 < a_x^2$ when $c \leq c_0$. Therefore, $v_n(x) = v_n^0(x)$ for $x = 0,\ldots,N-1$ ($x = 0,\ldots,N$ in Remark 2) and $v_n(N) = v_n^1(N)$ for all $n \geq 1$ when $c \leq c_0$. Since we have a finite state space X and A_x is finite for all x, $r(x,a_x)$ is bounded and the model is unichain, there exists a stationary long-run average optimal policy (see [23], Theorem 8.4.5). Note that from (1) and $q = N\mu_1 + N\theta + M\mu_2 = 1$, regardless of the action a_x chosen in each state x, we have $p(x\,|\,x,a_x) = 1-[(N-x)\mu_1 + (x-a_x)\theta + a_x\mu_2]/q = x\mu_1 + (N-x+a_x)\theta + (M-a_x)\mu_2 > 0$ for $\forall x \in X$ and $a_x \in A_x$, which indicates that the transition matrix for any feasible stationary policy is aperiodic. Therefore, since the stationary policies are unichain and every optimal policy has an aperiodic transition matrix, it follows from Theorems 8.5.4 and 8.5.6 of Puterman [23] that for any $\epsilon > 0$, value iteration will stop after a finite number of iterations with an ϵ-optimal policy. Furthermore, since ϵ is arbitrary and the state and action spaces are finite, an ϵ-optimal policy (for ϵ small enough) is indeed an optimal policy.

For part (ii), it follows from the proof of part (i) that $\Delta_n^{a_x^1,a_x^2}(x) \leq 0$ for all $n \geq 1, x \in \{1,\ldots,N\}$, and $a_x^1 < a_x^2 \in A_x$, when $c \geq c_0$ and $\theta = \mu_2$. When $\theta \neq \mu_2$, we again use induction to prove that $\Delta_n^{a_x^1,a_x^2}(x) \leq 0$ for all $n \geq 1, x \in \{1,\ldots,N\}$, and $a_x^1 < a_x^2 \in A_x$. From (4), we know that $\Delta_1^{a_x^1,a_x^2}(x) \leq 0$ for $x \in \{1,\ldots,N\}$ and $a_x^1 < a_x^2 \in A_x$. Assume that $\Delta_k^{a_x^1,a_x^2}(x) \leq 0$ for $k = 1,\ldots,n-1, x \in \{1,\ldots,N\}$, and $a_x^1 < a_x^2 \in A_x$. Therefore, we have $v_{n-1}(x) = v_{n-1}^{\min\{x,M\}}(x)$ for $x \in \{0,\ldots,N\}$.

Note that from $\Delta_{n-1}^{a_x^1,a_x^2}(x) \leq 0$ and (3), for $x = 1,\ldots,N$, we have

$$(\theta-\mu_2)[v_{n-2}(x-1) - v_{n-2}(x)] \leq -r_s\mu_s + r_2\mu_2 + c\theta. \tag{9}$$

https://doi.org/10.1017/S0269964823000013 Published online by Cambridge University Press

Furthermore, for $x \in \{1, \ldots, N\}$, (2) yields

$$\begin{aligned}
&v_{n-1}(x-1) - v_{n-1}(x) \\
&\quad = v_{n-1}^{\min\{x-1,M\}}(x-1) - v_{n-1}^{\min\{x,M\}}(x) \\
&\quad = (M - \min\{x-1, M\}) r_s\mu_s + \min\{x-1, M\} r_2\mu_2 - (x - 1 - \min\{x-1, M\}) c\theta \\
&\qquad + (N - x + 1)\mu_1 v_{n-2}(x) + [(x - 1 - \min\{x-1, M\})\theta + \min\{x-1, M\}\mu_2] v_{n-2}(x-2) \\
&\qquad + [1 - (N - x + 1)\mu_1 - (x - 1 - \min\{x-1, M\})\theta - \min\{x-1, M\}\mu_2] v_{n-2}(x-1) \\
&\qquad - (M - \min\{x, M\}) r_s\mu_s - \min\{x, M\} r_2\mu_2 + (x - \min\{x, M\}) c\theta \\
&\qquad - (N - x)\mu_1 v_{n-2}(x+1) - [(x - \min\{x, M\})\theta + \min\{x, M\}\mu_2] v_{n-2}(x-1) \\
&\qquad - [1 - (N - x)\mu_1 - (x - \min\{x, M\})\theta - \min\{x, M\}\mu_2] v_{n-2}(x) \\
&\quad = c\theta + (r_s\mu_s - r_2\mu_2 - c\theta) \times 1_{\{x \le M\}} + (N - x)\mu_1 [v_{n-2}(x) - v_{n-2}(x+1)] \\
&\qquad + [(x - 1 - \min\{x-1, M\})\theta + \min\{x-1, M\}\mu_2] \times [v_{n-2}(x-2) - v_{n-2}(x-1)] \\
&\qquad + [1 - (N - x + 1)\mu_1 - (x - \min\{x, M\})\theta - \min\{x, M\}\mu_2] \times [v_{n-2}(x-1) - v_{n-2}(x)],
\end{aligned}$$

where $1_{\{x \le M\}}$ is an indicator function defined as

$$1_{\{x \le M\}} = \begin{cases} 1 & \text{when } x \le M, \\ 0 & \text{when } x > M. \end{cases}$$

Since $q = N\mu_1 + N\theta + M\mu_2$ and we assumed, without loss of generality, that $q = 1$, we have that $1 - (N - x + 1)\mu_1 - (x - \min\{x, M\})\theta - \min\{x, M\}\mu_2$ is positive for $x \in \{1, \ldots, N\}$. Therefore, for $x \in \{1, \ldots, M\}$, (9) yields

$$\begin{aligned}
(\theta - \mu_2)[v_{n-1}(x-1) - v_{n-1}(x)] &\le (\theta - \mu_2)(r_s\mu_s - r_2\mu_2) + (N - x)\mu_1 \times (-r_s\mu_s + r_2\mu_2 + c\theta) \\
&\quad + (x-1)\mu_2 \times (-r_s\mu_s + r_2\mu_2 + c\theta) \\
&\quad + [1 - (N - x + 1)\mu_1 - x\mu_2] \times (-r_s\mu_s + r_2\mu_2 + c\theta) \\
&= (\theta - \mu_2)(r_s\mu_s - r_2\mu_2) + (1 - \mu_1 - \mu_2)(-r_s\mu_s + r_2\mu_2 + c\theta) \\
&= (1 - \mu_1 - \mu_2)c\theta - (1 - \mu_1 - \theta)(r_s\mu_s - r_2\mu_2),
\end{aligned}$$

and for $x \in \{M + 1, \ldots, N\}$, we have

$$\begin{aligned}
(\theta - \mu_2)[v_{n-1}(x-1) - v_{n-1}(x)] &\le (\theta - \mu_2)c\theta + (N - x)\mu_1 \times (-r_s\mu_s + r_2\mu_2 + c\theta) \\
&\quad + [(x - 1 - M)\theta + M\mu_2] \times (-r_s\mu_s + r_2\mu_2 + c\theta) \\
&\quad + [1 - (N - x + 1)\mu_1 - (x - M)\theta - M\mu_2] \times (-r_s\mu_s + r_2\mu_2 + c\theta) \\
&= (\theta - \mu_2)c\theta + (1 - \mu_1 - \theta)(-r_s\mu_s + r_2\mu_2 + c\theta) \\
&= (1 - \mu_1 - \mu_2)c\theta - (1 - \mu_1 - \theta)(r_s\mu_s - r_2\mu_2).
\end{aligned}$$

Now for all $x = 1, \ldots, N$ and $a_x^1 < a_x^2 \in A_s$, Eq. (3) yields that

$$\begin{aligned}
\Delta_n^{a_x^1, a_x^2}(x) &\le (a_x^2 - a_x^1)[r_s\mu_s - r_2\mu_2 - c\theta + (1 - \mu_1 - \mu_2)c\theta - (1 - \mu_1 - \theta)(r_s\mu_s - r_2\mu_2)] \\
&= (a_x^2 - a_x^1)[-(\mu_1 + \mu_2)c\theta + (\theta + \mu_1)(r_s\mu_s - r_2\mu_2)] \le 0 \qquad (10)
\end{aligned}$$

as long as $c \ge c_0$.

Equation (10) shows that $\Delta_n^{a_x^1, a_x^2}(x) \le 0$ for all $n \ge 1, x \in \{1, \ldots, N\}$, and $a_x^1 < a_x^2 \in A_x$ when $c \ge c_0$. Therefore, we have $v_n(x) = v_n^{\min\{x,M\}}(x)$ for all $x \in \{0, \ldots, N\}$. The remaining proof of part (ii) regarding the ϵ-optimality of the policy generated from value iteration is identical to the corresponding arguments in part (i). □

https://doi.org/10.1017/S0269964823000013 Published online by Cambridge University Press

The threshold c_0 increases in r_s, μ_s and decreases in r_2, μ_2. That is to say, the threshold on the abandonment cost where the supervisors switch from focusing on their own responsibilities to focusing on the customers increases with the rewards and processing rate of the supervisor on their own responsibilities, and decreases when the rewards or processing rate of the supervisors on the customers increase. This is because larger r_s, μ_s and smaller r_2, μ_2 all imply relatively greater rewards when the supervisors are working on their own responsibilities. The fact that c_0 does not depend on M, N reflects the linearity of the rewards and lack of switching times and costs, as well as the fact that each supervisor's choices on whether to work with a subordinate or not has limited immediate impact on other supervisors and subordinates.

Moreover, when $r_s\mu_s > r_2\mu_2$, i.e., π^C is ineffective from the perspective of immediate revenue, then c_0 decreases when θ increases. This means that when the supervisors earn greater rewards per unit time working on their own responsibilities, then as the abandonment rate θ increases, the supervisors switch from prioritizing their own responsibilities to prioritizing customers earlier (for lower abandonment costs). The condition $\mu_2 > \theta$ determines whether the rate of supervisors finishing the second-stage is larger than the abandonment rate. Therefore, if $\mu_2 > \theta$, then π^C is effective in reducing future abandonments. If $\mu_2 < \theta$, then c_0 decreases in μ_1. In this case, π^C is ineffective in both increasing immediate revenue and reducing future abandonments, and if π^C is optimal for a particular μ_1, then Policy π^C will remain optimal for a larger μ_1. However, if $\mu_2 > \theta$, then c_0 increases in μ_1. In this case, π^C is of mixed effectiveness in improving immediate revenue and reducing future abandonments, and if π^S is optimal for a particular μ_1, then Policy π^S will remain optimal for a larger μ_1.

When $r_s\mu_s < r_2\mu_2$, i.e., c_0 is negative, then c_0 increases in θ. When c_0 is negative, π^C is always optimal if there is a cost when a customer leaves the system without the second-stage service. However, when there is a reward for each customer leaving the system with the completion of the first-stage service only, an increase in the abandonment rate can lead supervisors to switch to serving the customers earlier (for lower abandonment rewards). If $\mu_2 > \theta$, then c_0 decreases in μ_1 and if $\mu_2 < \theta$, then c_0 increases in μ_1. This is because if $r_s\mu_s < r_2\mu_2$ and $\mu_2 > \theta$, then Policy π^C is effective in both increasing immediate revenues and reducing future abandonments, which leads to the conclusion that when Policy π^C is optimal for a specific μ_1, it will remain optimal for larger μ_1. Conversely, if $r_s\mu_s < r_2\mu_2$ and $\mu_2 < \theta$, then Policy π^C is effective in increasing immediate revenues but ineffective in reducing future abandonments. In this case, if π^S is optimal for a particular μ_1, then Policy π^S will remain optimal for a larger μ_1.

Remark 3. If $c = 0$, the two extreme policies are still optimal. In particular, if $r_s\mu_s - r_2\mu_2 \geq 0$, supervisors prioritize their own responsibilities; otherwise, they prioritize the customers. Thus, if there is no abandonment cost, the supervisors will focus on optimizing immediate revenue whenever they can.

Remark 4. If $r_s\mu_s - r_2\mu_2 = 0$, i.e., $c_0 = 0$, the optimality of π^S (π^C) depends on the whether c is negative (positive) only. When $r_s\mu_s = r_2\mu_2$, the rewards per unit time do not depend on the chosen action. Therefore, the optimal assignment of the supervisors only depends on whether there is a cost or a reward when a customer leaves the system without the second-stage service.

The next corollary specifies the optimal policy when the abandonment rate θ is small or large.

Corollary 5. *When $\theta \searrow 0$, π^S is optimal if $r_s\mu_s - r_2\mu_2 > 0$ and π^C is optimal if $r_s\mu_s - r_2\mu_2 < 0$. When $\theta \nearrow \infty$, π^S is optimal if $c \leq (r_s\mu_s - r_2\mu_2)/(\mu_1 + \mu_2)$ and π^C is optimal if $c \geq (r_s\mu_s - r_2\mu_2)/(\mu_1 + \mu_2)$.*

Proof. Theorem 1 introduces optimal policies with a threshold c_0 on c. We have

$$c_0 = \frac{(r_s\mu_s - r_2\mu_2)(\theta + \mu_1)}{(\mu_1 + \mu_2)\theta} = \frac{r_s\mu_s - r_2\mu_2}{\mu_1 + \mu_2} + \frac{(r_s\mu_s - r_2\mu_2)\mu_1}{(\mu_1 + \mu_2)\theta},$$

https://doi.org/10.1017/S0269964823000013 Published online by Cambridge University Press

which leads to:

$$\lim_{\theta\to 0} c_0 = \begin{cases} +\infty & \text{when } r_s\mu_s - r_2\mu_2 > 0, \\ 0 & \text{when } r_s\mu_s - r_2\mu_2 = 0, \\ -\infty & \text{when } r_s\mu_s - r_2\mu_2 < 0, \end{cases} \qquad \lim_{\theta\to\infty} c_0 = \frac{r_s\mu_s - r_2\mu_2}{\mu_1+\mu_2}.$$

□

Corollary 5 indicates that when the abandonment rate approaches 0, the optimal policy (whether a supervisor prioritizes her own responsibilities or serving customers) will maximize the immediate reward associated with the action. On the other hand, when the abandonment rate approaches infinity, the optimal policy still depends on how the abandonment cost c compares with a threshold. To better understand the value of the threshold, consider the case where $N > M = 1$ as an example. When π^S is adopted, in the limit all customers will abandon and the long-run average reward of the system approaches $r_s\mu_s - Nc\mu_1$. When π^C is adopted, in the limit the system behaves as a birth-death process with states 0, 1, birth rate $N\mu_1$, death rate μ_2, and the long-run average reward approaches $(\mu_2 r_s\mu_s + N\mu_1[r_2\mu_2 - (N-1)c\mu_1])/(\mu_2 + N\mu_1)$. The comparison of the long-run average rewards of the two systems leads to a threshold of $(r_s\mu_s - r_2\mu_2)/(\mu_1+\mu_2)$ for the parameter c.

The next proposition shows the closed-form expressions of the gains for policies π^S and π^C.

Proposition 6. *For $1 \le M \le N$, the gains of π^S and π^C are*

$$g^{\pi^S}_{N,M} = \frac{\begin{array}{c}\sum_{j=0}^{N-1}\binom{N}{j}\theta^{N-j}\mu_1^j[\mu_2 + (N-1)\theta](Mr_s\mu_s - jc\theta) \\ +N\theta\mu_1^N[(M-1)r_s\mu_s + r_2\mu_2 - (N-1)c\theta]\end{array}}{(\theta+\mu_1)^N[\mu_2+(N-1)\theta] + \mu_1^N(\theta-\mu_2)} \tag{11}$$

and

$$\begin{aligned} g^{\pi^C}_{N,M} = {} & \frac{1}{\sum_{k=0}^{M}\binom{N}{k}(\frac{\mu_1}{\mu_2})^k + \sum_{k=M+1}^{N}\frac{\prod_{i=1}^{k}[(N+1-i)\mu_1]}{M!\mu_2^M(\prod_{l=M+1}^{k}[M\mu_2+(l-M)\theta])}} \\ & \times\left\{\sum_{k=0}^{M}\binom{N}{k}\left(\frac{\mu_1}{\mu_2}\right)^k[(M-k)r_s\mu_s + kr_2\mu_2]\right. \\ & \left. + \sum_{k=M+1}^{N}\frac{\prod_{i=1}^{k}[(N+1-i)\mu_1]}{M!\mu_2^M\prod_{l=M+1}^{k}[M\mu_2+(l-M)\theta]}[Mr_2\mu_2 - (k-M)c\theta]\right\}, \end{aligned} \tag{12}$$

respectively (with the convention that the summation over an empty set is 0).

Proof. The long-run average rewards $g^{\pi^C}_{N,M}$ and $g^{\pi^S}_{N,M}$ can be computed using the birth-death structure of the underlying Markov chains under π^C and π^S, respectively. Specifically, the closed-form expression of the gain (long-run average reward) for any specific policy can be uniquely determined by

$$g^{\pi}_{N,M} = \sum_{x=0}^{N}\eta^{\pi}_x r(x, d_\pi(x))q, \tag{13}$$

where η^π_x is the limiting probability of $\{X_\pi(t)\}$ being in state x under policy π.

Let $\boldsymbol{\eta}^\pi$ denote the limiting probability vector under policy π. The limiting probabilities can be obtained by solving the set of equations $\eta^\pi_x[\sum_{k\neq x} q_\pi(x,k)] = \sum_{k\neq x}\eta^\pi_k q_\pi(k,x)$ for all $x \in X$, along with the equation $\sum_{k=0}^{N}\eta^\pi_k = 1$.

https://doi.org/10.1017/S0269964823000013 Published online by Cambridge University Press

Therefore, for policy π^S, we have

$$\eta_x^{\pi^S} = \begin{cases} \dfrac{\binom{N}{x}\theta^{N-x}\mu_1^x[\mu_2+(N-1)\theta]}{\sum_{j=0}^{N-1}\binom{N}{j}\theta^{N-j}\mu_1^j[\mu_2+(N-1)\theta]+N\theta\mu_1^N} & \text{for } x=0,\ldots,N-1, \\ \dfrac{N\theta\mu_1^N}{\sum_{j=0}^{N-1}\binom{N}{j}\theta^{N-j}\mu_1^j[\mu_2+(N-1)\theta]+N\theta\mu_1^N} & \text{for } x=N, \end{cases} \tag{14}$$

and for policy π^C, we have

$$\eta_x^{\pi^C} = \begin{cases} \dfrac{\binom{N}{x}(\frac{\mu_1}{\mu_2})^x}{\sum_{k=0}^{M}\binom{N}{k}(\frac{\mu_1}{\mu_2})^k+\sum_{k=M+1}^{N}\frac{\prod_{i=1}^{k}[(N+1-i)\mu_1]}{M!\mu_2^M(\prod_{l=M+1}^{k}[M\mu_2+(l-M)\theta])}} & \text{for } x=0,\ldots,M, \\ \dfrac{\frac{\prod_{i=1}^{x}[(N+1-i)\mu_1]}{M!\mu_2^M(\prod_{l=M+1}^{x}[M\mu_2+(l-M)\theta])}}{\sum_{k=0}^{M}\binom{N}{k}(\frac{\mu_1}{\mu_2})^k+\sum_{k=M+1}^{N}\frac{\prod_{i=1}^{k}[(N+1-i)\mu_1]}{M!\mu_2^M(\prod_{l=M+1}^{k}[M\mu_2+(l-M)\theta])}} & \text{for } x=M+1,\ldots,N. \end{cases} \tag{15}$$

By plugging Eqs. (14) ((15)) and the corresponding rewards $r(x,d_{\pi^S}(x))$ $(r(x,d_{\pi^C}(x)))$ into (13), we can obtain the gains of π^S (π^C). □

Until now, we have assumed that customers will not abandon while they are receiving (first- or second-stage) service. This is motivated by service applications where it is unlikely that customers will abandon when in service. However, there are situations where abandonments may occur during the service (e.g., in healthcare applications).

The next corollary extends Theorem 1 to the case where abandonments can also occur during the first- and second-stage service. In particular, the corollary show that the structure of the optimal threshold policy remains the same when abandonments can also occur during service. Let θ_1 (θ_2) denote the abandonment rate during the first-stage (second-stage) service and c_1 (c_2) denote the corresponding abandonment cost. Note that θ_1, θ_2 or c_1, c_2 do not necessarily equal θ or c. Then, one of π^C and π^S is always optimal, but the threshold on the value of c is different.

Corollary 7. *When abandonments can also occur during the first- and second-stage service,*

(i) if $c \le c_0' := \frac{c_1\theta_1(\mu_2+\theta_2-\theta)+(r_s\mu_s-r_2\mu_2+c_2\theta_2)(\theta+\mu_1)}{(\mu_1+\mu_2+\theta_2)\theta}$*, then* π^S *is optimal;*
(ii) if $c \ge c_0'$*, then* π^C *is optimal.*

The proof of Corollary 7 follows similar techniques as the proof of Theorem 1 by setting

$$v_0(x) = (N-x-1)\times\frac{r_s\mu_s - r_2\mu_2 + c_2\theta_2 - c_1\theta_1}{\mu_1+\mu_2+\theta_2}.$$

Alternatively, the new threshold can be obtained in an intuitive way as follows. Note that when abandonments can also occur during the first- and second-stage service, the birth rates in Figure 1 remain the same, while the death rate in state $x \in \{1,\ldots,N\}$ is now $(x-a_x)\theta + a_x(\mu_2+\theta_2)$ due to the abandonments that may take place during the second-stage service. Similarly, the immediate reward of the second-stage service is now $r_2\times\mu_2/(\mu_2+\theta_2) - c_2\times\theta_2/(\mu_2+\theta_2)$. Moreover, for state $x\in\{1,\ldots,N\}$, the costs from the first-stage abandonments are $(N-x)c_1\theta_1$ per unit time. That is to say, the immediate reward $r(x, a_x)$ of choosing action a_x in state x now is

$$r(x,a_x) = \frac{(M-a_x)r_s\mu_s + a_x(r_2 \times \frac{\mu_2}{\mu_2+\theta_2} - c_2 \times \frac{\theta_2}{\mu_2+\theta_2})(\mu_2+\theta_2) - (x-a_x)c\theta - (N-x)c_1\theta_1}{q}$$
$$= \frac{(M-a_x)r_s\mu_s + a_x(\frac{r_2\mu_2}{\mu_2+\theta_2} + \frac{c_1\theta_1}{\mu_2+\theta_2} - \frac{c_2\theta_2}{\mu_2+\theta_2})(\mu_2+\theta_2) - (x-a_x)(c - \frac{c_1\theta_1}{\theta})\theta - Nc_1\theta_1}{q}.$$

By ignoring the $Nc_1\theta_1/q$ term as it is constant in x, replacing the r_2 term in c_0 by $r_2\mu_2/(\mu_2+\theta_2) + c_1\theta_1/(\mu_2+\theta_2) - c_2\theta_2/(\mu_2+\theta_2)$, replacing the c term by $c - c_1\theta_1/\theta$, replacing the μ_2 term by $\mu_2+\theta_2$, and replacing the $(\mu_1+\mu_2)$ term by $(\mu_1+\mu_2+\theta_2)$ in (8) and (10), the structure of the optimal policy remains unchanged, and the new threshold c_0' should satisfy

$$c_0' = \frac{c_1\theta_1(\mu_2+\theta_2-\theta) + (r_s\mu_s - r_2\mu_2 + c_2\theta_2)(\theta+\mu_1)}{(\mu_1+\mu_2+\theta_2)\theta},$$

as in Corollary 7. We note that c_0' increases in c_2. This is because when the second-stage abandonment cost increases, the actual reward of a supervisor serving a customer $(r_2 \times \mu_2/(\mu_2+\theta_2) - c_2 \times \theta_2/(\mu_2+\theta_2))$ decreases. Therefore, the supervisors will only switch to serve the customers for larger abandonment costs while they are waiting for the second-stage service.

Moreover, c_0' is constant in c_1 or θ_1 when $\mu_2+\theta_2 = \theta$; c_0' increases in c_1 or θ_1 when $\mu_2+\theta_2 > \theta$; and c_0' decreases in c_1 or θ_1 when $\mu_2+\theta_2 < \theta$. This is because when $\mu_2+\theta_2 = \theta$, the death rate in state $x \in \{1,\ldots,N\}$ is $a_x(\mu_2+\theta_2) + (x-a_x)\theta = x\theta$. Thus, the death rate in state x is the same regardless of the chosen action a_x, which leads to the threshold c_0' remaining the same. However, when $\mu_2+\theta_2 > \theta$, larger a_x results in higher death rates. Since the c_1, θ_1 terms in the immediate reward $r(x,a_x)$ equal $-(N-x)c_1\theta_1$, the effects of c_1, θ_1 are less for smaller x, leading to an increment in the threshold c_0'. On the contrary, when $\mu_2+\theta_2 < \theta$, larger a_x results in lower death rates, and hence, the supervisors will switch to serve the customers for smaller abandonment costs c while they are waiting for the second-stage service.

Meanwhile, we note that c_0' increases in θ_2 when $c_1\theta_1 + c_2(\mu_1+\mu_2) > r_s\mu_s - r_2\mu_2$; decreases in θ_2 when $c_1\theta_1 + c_2(\mu_1+\mu_2) < r_s\mu_s - r_2\mu_2$; and is constant in θ_2 when $c_1\theta_1 + c_2(\mu_1+\mu_2) = r_s\mu_s - r_2\mu_2$. Thus, when the abandonment costs c_1, c_2 and rate θ_1 during service are large (small) relative to the benefit $r_s\mu_s - r_2\mu_2$ of supervisors focusing on their own responsibilities, the supervisors will switch later (earlier) from their own responsibilities to serving the customers as the abandonment rate θ_2 increases.

Remark 8. When abandonments can also occur during the first- and second-stage service, if $c_1 = c_2 = c$ and $\theta_1 = \theta_2 = \theta$, we have:

(i) if $r_2\mu_2 \le r_s\mu_s$, then π^S is optimal;
(ii) if $r_2\mu_2 \ge r_s\mu_s$, then π^C is optimal.

Note that when $c_1 = c_2 = c$ and $\theta_1 = \theta_2 = \theta$, the immediate reward $r(x,a_x)$ of choosing action a_x in state x is

$$r(x,a_x) = \frac{(M-a_x)r_s\mu_s + a_x(r_2 \times \frac{\mu_2}{\mu_2+\theta} - c \times \frac{\theta}{\mu_2+\theta})(\mu_2+\theta) - (x-a_x)c\theta - (N-x)c\theta}{q}$$
$$= \frac{Mr_s\mu_s + a_x(r_2\mu_2 - r_s\mu_s) - Nc\theta}{q}. \tag{16}$$

Since $Mr_s\mu_s/q$ and $Nc\theta/q$ in (16) are constant in x, the optimal policy depends solely on the comparison of $r_2\mu_2$ and $r_s\mu_s$ in this case.

https://doi.org/10.1017/S0269964823000013 Published online by Cambridge University Press

Remark 9. When abandonments can also occur during the first-stage (but not during the second-stage service), if $c_1 = c$ and $\theta_1 = \theta$, we have

(i) if $c \leq (r_s\mu_s - r_2\mu_2)/\theta$, then π^S is optimal;
(ii) if $c \geq (r_s\mu_s - r_2\mu_2)/\theta$, then π^C is optimal.

Note that in this case, π^S is optimal when $r_s\mu_s \geq r_2\mu_2 + c\theta$ and π^C is optimal otherwise. This is because the immediate reward $r(x, a_x)$ of choosing action a_x in state x now is

$$\begin{aligned} r(x, a_x) &= \frac{(M - a_x)r_s\mu_s + a_x(r_2\mu_2 + c\theta) - Nc\theta}{q} \\ &= \frac{Mr_s\mu_s + a_x(r_2\mu_2 + c\theta - r_s\mu_s) - Nc\theta}{q}, \end{aligned}$$

which does not depend on x. Therefore, the optimal policy depends on the comparison of $r_2\mu_2 + c\theta$ and $r_s\mu_s$ in this case.

Remark 10. When abandonments can also occur during the second-stage (but not during the first-stage service), if $c_2 = c$ and $\theta_2 = \theta$, we have

(i) if $c \leq [(r_s\mu_s - r_2\mu_2)(\theta + \mu_1)]/\theta\mu_2$, then π^S is optimal;
(ii) if $c \geq [(r_s\mu_s - r_2\mu_2)(\theta + \mu_1)]/\theta\mu_2$, then π^C is optimal.

Note that in this case, the immediate reward $r(x, a_x)$ of choosing action a_x in state x is

$$\begin{aligned} r(x, a_x) &= \frac{(M - a_x)r_s\mu_s + a_x(\frac{r_2\mu_2}{\mu_2+\theta} - \frac{c\theta}{\mu_2+\theta})(\mu_2 + \theta) - (x - a_x)c\theta}{q} \\ &= \frac{Mr_s\mu_s + a_x(r_2\mu_2 - r_s\mu_s) - xc\theta}{q}, \end{aligned}$$

which depends on the state x. However, when $r_s\mu_s - r_2\mu_2 > 0$, the threshold in Remark 10 satisfies $[(r_s\mu_s - r_2\mu_2)(\theta + \mu_1)]/\theta\mu_2 > c_0 = [(r_s\mu_s - r_2\mu_2)(\theta + \mu_1)]/(\mu_1 + \mu_2)\theta > 0$. This implies that when the supervisors earn greater rewards per unit time working on their own responsibilities, if abandonments can also occur during the second-stage service and $c_2 = c > 0$, the expected rewards for supervisors serving the customers decrease. Therefore, the supervisors will switch to serve the customers for larger abandonment costs. However, when $r_s\mu_s - r_2\mu_2 < 0$, we have $[(r_s\mu_s - r_2\mu_2)(\theta + \mu_1)]/\theta\mu_2 < c_0 < 0$. That is to say, when the supervisors earn greater rewards per unit time serving the customers, if abandonments can also occur during the second-stage service and there is a reward for customers leaving the system without second-stage service (i.e., $-c$ as $c < 0$), there are added benefits for supervisors to serve the customers and hence they will switch to serving the customers earlier.

4. Benefits of pooling the subordinates of several supervisors

In this section, we investigate the effects of pooling supervisors (and their subordinates) on the system performance. When there are multiple supervisors, each of whom has her own subordinates, a natural question arises: should each supervisor work with her subordinates only, or should all supervisors work with all the subordinates? Consider our example of government services in Section 1. One possibility is that the waiting people form several queues and each official is responsible for one queue in addition to her other responsibilities. Alternatively, all officials can be jointly responsible for serving all the waiting people (in addition to their own responsibilities).

In particular, we consider two cases. In case 1, there are KM subordinates and M supervisors (a pooled system with K subordinates per supervisor); in case 2, there are M systems each with $K \geq 1$

https://doi.org/10.1017/S0269964823000013 Published online by Cambridge University Press

subordinates and one supervisor (M dedicated systems). Note that since c_0 does not depend on the number of subordinates or supervisors, the optimality condition is the same for both cases. We then have Proposition 12. Before we elaborate on Proposition 12, we first prove Lemma 11.

Lemma 11. *For* $1 \le k \le KM$, $\binom{KM}{k}k \ge \sum_{j=1}^{\min\{K,k\}} \binom{KM}{k-j}\binom{K}{j} j!(M-k+j)$.

Proof. Observe that

$$\binom{KM}{k}k = \binom{KM}{k-1}[KM-(k-1)],$$

and

$$\begin{aligned}\binom{K}{j} j!(M-k+j) &\le K(K-1)^{j-1}(M-k+j)\\ &= [KM-(k-j)-(k-j)(K-1)](K-1)^{j-1}\end{aligned}$$

for $1 \le k \le KM$ and $1 \le j \le K$. When $1 \le k \le KM$, we have

$$\begin{aligned}&\binom{KM}{k}k - \sum_{j=1}^{\min\{K,k\}} \binom{KM}{k-j}\binom{K}{j} j!(M-k+j)\\
&\ge \binom{KM}{k-1}[KM-(k-1)] - \sum_{j=1}^{\min\{K,k\}} \binom{KM}{k-j}[KM-(k-j)-(k-j)(K-1)](K-1)^{j-1}\\
&= \binom{KM}{k-1}(k-1)(K-1) - \sum_{j=2}^{\min\{K,k\}} \binom{KM}{k-j}[KM-(k-j)-(k-j)(K-1)](K-1)^{j-1}\\
&= \binom{KM}{k-2}[KM-(k-2)](K-1) - \sum_{j=2}^{\min\{K,k\}} \binom{KM}{k-j}[KM-(k-j)-(k-j)(K-1)](K-1)^{j-1}\\
&\ge \binom{KM}{k-2}(k-2)(K-1)^2 - \sum_{j=3}^{\min\{K,k\}} \binom{KM}{k-j}[KM-(k-j)-(k-j)(K-1)](K-1)^{j-1}\\
&\ \ \vdots\\
&\ge \binom{KM}{k-\min\{K,k\}}(k-\min\{K,k\})(K-1)^{\min\{K,k\}} \ge 0.\end{aligned}$$

□

Proposition 12. *(i) If* $c \le c_0$ *(i.e.,* π^S *is optimal), then* $g_{KM,M}^{\pi^S} \ge Mg_{K,1}^{\pi^S}$ *for* $M, K \ge 1$ *and*

$$\lim_{M\to\infty} \frac{g_{KM,M}^{\pi^S} - Mg_{K,1}^{\pi^S}}{M} = \frac{K\theta\mu_1^K[(\mu_1+\theta)(r_s\mu_s - r_2\mu_2) - c\theta(\mu_1+\mu_2)]}{(\mu_1+\theta)\{[\mu_2+(K-1)\theta](\theta+\mu_1)^K + \mu_1^K(\theta-\mu_2)\}}$$

for $K \ge 1$;

(ii) If $c \ge c_0$ *(i.e.,* π^C *is optimal), then* $g_{KM,M}^{\pi^C} \ge Mg_{K,1}^{\pi^C}$ *for* $M, K \ge 1$ *and*

$$0 \le \frac{g_{KM,M}^{\pi^C} - Mg_{K,1}^{\pi^C}}{M} \le \frac{c\theta(\mu_1+\mu_2) - (\mu_1+\theta)(r_s\mu_s - r_2\mu_2)}{(\mu_1+\theta)[1+T(K)]}$$

for $K \ge 1$*, where* $T(K) = \sum_{k=1}^{K} \prod_{i=1}^{k}[(K+1-i)\mu_1]/\Pi_{l=1}^{k}[\mu_2+(l-1)\theta]$.

https://doi.org/10.1017/S0269964823000013 Published online by Cambridge University Press

Proof. (i) Using the closed-form expression of the gain for policy π^S in (11), with some algebra, we have

$$
\begin{aligned}
g^{\pi^S}_{KM,M} - Mg^{\pi^S}_{K,1} = &\frac{1}{(\mu_1+\theta)\{[(K-1)\theta+\mu_2](\theta+\mu_1)^K+\mu_1^K(\theta-\mu_2)\}} \\
&\quad \{[(KM-1)\theta+\mu_2](\theta+\mu_1)^{KM}+\mu_1^{KM}(\theta-\mu_2)\} \\
&\quad \times KM\theta[(\mu_1+\theta)(r_s\mu_s-r_2\mu_2)-c\theta(\mu_1+\mu_2)] \\
&\quad \times \{\mu_1^K(\mu_1+\theta)^{KM}[(KM-1)\theta+\mu_2]-\mu_1^{KM}(\mu_1+\theta)^K[(K-1)\theta+\mu_2]\}. \quad (17)
\end{aligned}
$$

We now proceed to show that each term here is non-negative. Since $c \le c_0$, $(\mu_1+\theta)(r_s\mu_s-r_2\mu_2)-c\theta(\mu_1+\mu_2) \ge 0$. Furthermore,

$$
\mu_1^K(\mu_1+\theta)^{KM}[(KM-1)\theta+\mu_2]-\mu_1^{KM}(\mu_1+\theta)^K[(K-1)\theta+\mu_2] \ge \mu_1^{KM}(\mu_1+\theta)^K K(M-1)\theta \ge 0.
$$

Similarly,

$$
[(K-1)\theta+\mu_2](\theta+\mu_1)^K+\mu_1^K(\theta-\mu_2) \ge K\theta\mu_1^K > 0
$$

and

$$
[(KM-1)\theta+\mu_2](\theta+\mu_1)^{KM}+\mu_1^{KM}(\theta-\mu_2) \ge KM\theta\mu_1^{KM} > 0.
$$

Thus, $g^{\pi^S}_{KM,M} - Mg^{\pi^S}_{K,1} \ge 0$ with the equality holding when $c = c_0$ or $M = 1$. We now proceed to obtain the limit of $(g^{\pi^S}_{KM,M} - Mg^{\pi^S}_{K,1})/M$ when M goes to infinity. Based on (17), we have:

$$
\begin{aligned}
&\lim_{M\to\infty} \frac{g^{\pi^S}_{KM,M} - Mg^{\pi^S}_{K,1}}{M} \\
&= \frac{K\theta\mu_1^K[(\mu_1+\theta)(r_s\mu_s-r_2\mu_2)-c\theta(\mu_1+\mu_2)]}{(\mu_1+\theta)\{[(K-1)\theta+\mu_2](\theta+\mu_1)^K+\mu_1^K(\theta-\mu_2)\}} \\
&\quad \times \lim_{M\to\infty} \frac{(\mu_1+\theta)^{KM}[(KM-1)\theta+\mu_2]-\mu_1^{K(M-1)}(\mu_1+\theta)^K[(K-1)\theta+\mu_2]}{[(KM-1)\theta+\mu_2](\theta+\mu_1)^{KM}+\mu_1^{KM}(\theta-\mu_2)}.
\end{aligned}
$$

The result now follows from the fact that

$$
\lim_{M\to\infty} \frac{(\mu_1+\theta)^{KM}[(KM-1)\theta+\mu_2]-\mu_1^{K(M-1)}(\mu_1+\theta)^K[(K-1)\theta+\mu_2]}{[(KM-1)\theta+\mu_2](\theta+\mu_1)^{KM}+\mu_1^{KM}(\theta-\mu_2)} = 1.
$$

(ii) Using the closed-form expression of the gain for policy π^C in (12), with some algebra, we have

$$
\begin{aligned}
g^{\pi^C}_{KM,M} - Mg^{\pi^C}_{K,1} = &\frac{[c\theta(\mu_1+\mu_2)-(\mu_1+\theta)(r_s\mu_s-r_2\mu_2)]}{(\mu_1+\theta)[1+T(K)]} \\
&\times \frac{\Gamma}{\sum_{k=0}^{M}\binom{KM}{k}(\frac{\mu_1}{\mu_2})^k+\sum_{k=M+1}^{KM}\frac{\prod_{i=1}^{k}[(KM+1-i)\mu_1]}{M!\mu_2^M\prod_{l=M+1}^{k}[M\mu_2+(l-M)\theta]}}, \quad (18)
\end{aligned}
$$

https://doi.org/10.1017/S0269964823000013 Published online by Cambridge University Press

where

$$\Gamma = M \sum_{k=M+1}^{KM} \frac{\prod_{i=1}^{k}[(KM+1-i)\mu_1]}{M!\mu_2^M \Pi_{l=M+1}^{k}[M\mu_2+(l-M)\theta]} + \sum_{k=1}^{M} \binom{KM}{k} \left(\frac{\mu_1}{\mu_2}\right)^k k$$
$$- \left(\sum_{k=1}^{M} \binom{KM}{M-k} \left(\frac{\mu_1}{\mu_2}\right)^{M-k} k\right) \times T(K).$$

We now show that the expression (18) is non-negative. Note that the term $c\theta(\mu_1+\mu_2)-(r_s\mu_s-r_2\mu_2)(\mu_1+\theta) \geq 0$ since $c \geq c_0$. We will show that $\Gamma \geq 0$. Define

$$\alpha_k = \sum_{j=\max\{0,k-K\}}^{k-1} \frac{\binom{KM}{j}\binom{K}{k-j}(k-j)!(M-j)}{\mu_2^j \Pi_{l=1}^{k-j}[\mu_2+(l-1)\theta]}$$

for $1 \leq k \leq M$, and

$$\beta_k = \sum_{j=\max\{0,k-K\}}^{M-1} \frac{\binom{KM}{j}\binom{K}{k-j}(k-j)!(M-j)}{\mu_2^j \Pi_{l=1}^{k-j}[\mu_2+(l-1)\theta]}$$

for $M+1 \leq k \leq M+K-1$. Note that by expanding $(\sum_{k=1}^{M}\binom{KM}{M-k}(\frac{\mu_1}{\mu_2})^{M-k}k) \times T(K)$ and ordering the terms by ascending exponent of μ_1, we have

$$\left(\sum_{k=1}^{M} \binom{KM}{M-k} \left(\frac{\mu_1}{\mu_2}\right)^{M-k} k\right) \times T(K) = \left(\sum_{j=0}^{M-1} \binom{KM}{j} \left(\frac{\mu_1}{\mu_2}\right)^{j} (M-j)\right)$$
$$\times \left(\sum_{k=1}^{K} \frac{\prod_{i=1}^{k}[(K+1-i)\mu_1]}{\Pi_{l=1}^{k}[\mu_2+(l-1)\theta]}\right)$$
$$= \sum_{k=1}^{M+K-1} \mu_1^k \sum_{j=\max\{0,k-K\}}^{\min\{k-1,M-1\}} \frac{\binom{KM}{j}\binom{K}{k-j}(k-j)!(M-j)}{\mu_2^j \Pi_{l=1}^{k-j}[\mu_2+(l-1)\theta]}$$
$$= \sum_{k=1}^{M} \mu_1^k \alpha_k + \sum_{k=M+1}^{M+K-1} \mu_1^k \beta_k.$$

By grouping the terms in Γ based on the exponent of μ_1, we obtain $\Gamma = \Gamma_1 + \Gamma_2 + \Gamma_3$, where

$$\Gamma_1 = \sum_{k=1}^{M} \mu_1^k \left[\binom{KM}{k}\frac{k}{\mu_2^k} - \alpha_k\right],$$
$$\Gamma_2 = \sum_{k=M+1}^{M+K-1} \mu_1^k \left[\frac{M\prod_{i=1}^{k}(KM+1-i)}{M!\mu_2^M \Pi_{l=M+1}^{k}[M\mu_2+(l-M)\theta]} - \beta_k\right],$$
$$\Gamma_3 = \sum_{k=M+K}^{KM} \mu_1^k \frac{M\prod_{i=1}^{k}(KM+1-i)}{M!\mu_2^M \Pi_{l=M+1}^{k}[M\mu_2+(l-M)\theta]}$$

(recall that the summation over an empty set equals zero).

Note that Γ_3 is positive if $M, K \geq 2$ and Γ_3 is zero if $M = 1$ or $K = 1$. Moreover, when $M = 1$, $\Gamma_1 = 0$ and $\Gamma_2 = \sum_{k=M+1}^{M+N-1} \mu_1^k \times 0 = 0$. Thus, $\Gamma = 0$ when $M = 1$ and it suffices to show that Γ_1 and Γ_2 are non-negative, which we will prove by showing that each term of the respective summation is

https://doi.org/10.1017/S0269964823000013 Published online by Cambridge University Press

non-negative. For Γ_1, when $1 \le k \le M$, we have

$$\binom{KM}{k}\frac{k}{\mu_2^k} - \alpha_k \ge \frac{1}{\mu_2^k}\left[\binom{KM}{k}k - \sum_{j=\max\{0,k-K\}}^{k-1}\binom{KM}{j}\binom{K}{k-j}(k-j)!(M-j)\right]$$
$$= \frac{1}{\mu_2^k}\left[\binom{KM}{k}k - \sum_{j=1}^{\min\{K,k\}}\binom{KM}{k-j}\binom{K}{j}j!(M-k+j)\right] \ge 0, \tag{19}$$

where the last inequality follows from Lemma 11. Similarly, note that

$$\beta_k \le \sum_{j=\max\{0,k-K\}}^{M-1}\frac{\binom{KM}{j}\binom{K}{k-j}(k-j)!(M-j)}{\mu_2^M\prod_{l=1}^{k-M}(\mu_2+l\theta)}, \quad \text{for } M+1 \le k \le M+K-1.$$

Therefore, for $M+1 \le k \le M+K-1$,

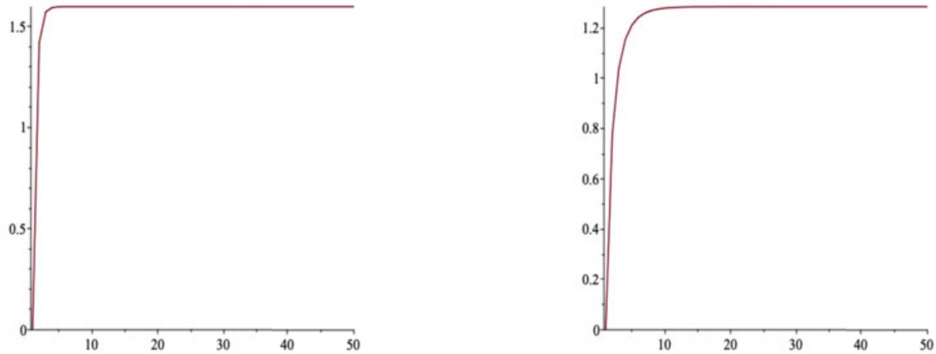
$$\frac{M\prod_{i=1}^{k}(KM+1-i)}{M!\mu_2^M\prod_{l=M+1}^{k}[M\mu_2+(l-M)\theta]} - \beta_k$$
$$\ge \frac{M\binom{KM}{k}k!}{M!\mu_2^M M^{k-M}\prod_{l=1}^{k-M}(\mu_2+l\theta)} - \frac{\sum_{j=\max\{0,k-K\}}^{M-1}\binom{KM}{j}\binom{K}{k-j}(k-j)!(M-j)}{\mu_2^M\prod_{l=1}^{k-M}(\mu_2+l\theta)}$$
$$= \frac{1}{\mu_2^M\prod_{l=1}^{k-M}(\mu_2+l\theta)}\left[\frac{\binom{KM}{k}k!}{(M-1)!M^{k-M}} - \sum_{j=\max\{0,k-K\}}^{M-1}\binom{KM}{j}\binom{K}{k-j}(k-j)!(M-j)\right]$$
$$\ge \frac{1}{\mu_2^M\prod_{l=1}^{k-M}(\mu_2+l\theta)}\left[\binom{KM}{k}k - \sum_{j=\max\{0,k-K\}}^{M-1}\binom{KM}{j}\binom{K}{k-j}(k-j)!(M-j)\right]$$
$$= \frac{1}{\mu_2^M\prod_{l=1}^{k-M}(\mu_2+l\theta)}\left[\binom{KM}{k}k - \sum_{j=k-M+1}^{\min\{K,k\}}\binom{KM}{k-j}\binom{K}{j}j!(M-k+j)\right]$$
$$\ge \frac{1}{\mu_2^M\prod_{l=1}^{k-M}(\mu_2+l\theta)}\left[\binom{KM}{k}k - \sum_{j=1}^{\min\{K,k\}}\binom{KM}{k-j}\binom{K}{j}j!(M-k+j)\right] \ge 0, \tag{20}$$

where the last inequality follows from Lemma 11. It follows from (19) and (20) that Γ_1 and Γ_2 are non-negative, which implies that Γ is positive when $M \ge 2$ and $K \ge 2$. When $M \ge 2$ and $K = 1$, since $\Gamma_2 = \Gamma_3 = 0$, we have

$$\Gamma = \Gamma_1 = \sum_{k=1}^{M}\mu_1^k\left[\binom{M}{k}\frac{k}{\mu_2^k} - \alpha_k\right] = \sum_{k=1}^{M}\left(\frac{\mu_1}{\mu_2}\right)^k\left[\binom{M}{k}k - \binom{M}{k-1}(M-k+1)\right] = 0.$$

Therefore, we have $g_{KM,M}^{\pi^C} - Mg_{K,1}^{\pi^C} \ge 0$ with equality holding only when $c = c_0$ or $M = 1$ or $K = 1$ ($\Gamma = 0$ in the last two cases).

https://doi.org/10.1017/S0269964823000013 Published online by Cambridge University Press



(a) The value of $\frac{g^{\pi^S}_{KM,M} - M g^{\pi^S}_{K,1}}{M}$ when $c = 2$. (b) The value of $\frac{g^{\pi^C}_{KM,M} - M g^{\pi^C}_{K,1}}{M}$ when $c = 10$.

Figure 2. *The incremental value of pooling M supervisors as a function of M.*

We now proceed to obtain the lower and upper bounds of $(g^{\pi^C}_{KM,M} - M g^{\pi^C}_{K,1})/M$ when M goes to infinity. Note that

$$\frac{\Gamma}{\sum_{k=0}^{M} \binom{KM}{k} \left(\frac{\mu_1}{\mu_2}\right)^k + \sum_{k=M+1}^{KM} \frac{\prod_{i=1}^{k}[(KM+1-i)\mu_1]}{M!\mu_2^M \prod_{l=M+1}^{k}[M\mu_2 + (l-M)\theta]}} \times \frac{1}{M}$$

$$= \frac{M \sum_{k=M+1}^{KM} \frac{\prod_{i=1}^{k}[(KM+1-i)\mu_1]}{M!\mu_2^M \prod_{l=M+1}^{k}[M\mu_2 + (l-M)\theta]} + \sum_{k=1}^{M} \binom{KM}{k} \left(\frac{\mu_1}{\mu_2}\right)^k k - T(K) \times \left[\sum_{k=0}^{M-1} \binom{KM}{k} \left(\frac{\mu_1}{\mu_2}\right)^k (M-k)\right]}{M \sum_{k=0}^{M} \binom{KM}{k} \left(\frac{\mu_1}{\mu_2}\right)^k + M \sum_{k=M+1}^{KM} \frac{\prod_{i=1}^{k}[(KM+1-i)\mu_1]}{M!\mu_2^M \prod_{l=M+1}^{k}[M\mu_2 + (l-M)\theta]}}$$

$$= 1 - \frac{[1 + T(K)] \sum_{k=0}^{M} \binom{KM}{k} \left(\frac{\mu_1}{\mu_2}\right)^k \times (M-k)}{M \sum_{k=0}^{M} \binom{KM}{k} \left(\frac{\mu_1}{\mu_2}\right)^k + M \sum_{k=M+1}^{KM} \frac{\prod_{i=1}^{k}[(KM+1-i)\mu_1]}{M!\mu_2^M \prod_{l=M+1}^{k}[M\mu_2 + (l-M)\theta]}} < 1.$$

Since $g^{\pi^C}_{KM,M} - M g^{\pi^C}_{K,1} \geq 0$, it now follows from (18) that for all $M \geq 1$,

$$0 \leq \frac{g^{\pi^C}_{KM,M} - M g^{\pi^C}_{K,1}}{M} \leq \frac{c\theta(\mu_1 + \mu_2) - (\mu_1 + \theta)(r_s\mu_s - r_2\mu_2)}{(\mu_1 + \theta)[1 + T(K)]}.$$

□

Remark 13. It follows directly from the proof of Proposition 12 that pooling supervisors and their associated subordinates is a strict improvement over the unpooled system, as long as $M > 1$ and either $c < c_0$ or $c > c_0$ and $K > 1$; otherwise, the pooled and unpooled systems have identical performance. While it may at first seem surprising that pooling is not beneficial when $c > c_0$ and $K = 1$, observe that the Markov chain models of the pooled and unpooled systems are identical under π^C when $K = 1$.

Proposition 12 shows that pooling supervisors and their subordinates improves the performance of the system in terms of the long-run average reward. However, the improvement per pooled supervisor is bounded. We utilize numerical examples to illustrate the comparison of dedicated and pooled systems and to quantify the incremental benefit per pooled supervisor of pooling systems as more supervisors

https://doi.org/10.1017/S0269964823000013 Published online by Cambridge University Press

are pooled. In our numerical examples, we have: $\mu_1 = 4$, $r_1 = 5$, $\mu_2 = 6$, $r_2 = 8$, $\mu_s = 11$, $r_s = 6$, $\theta = 2$. Consider the dedicated and pooled systems where there are M supervisors, each of whom has $K = 4$ subordinates. Note that the threshold c_0 is $\frac{27}{5} = 5.4$. Figure 2 shows the value of pooling M supervisors as a function of M when $c = 2$ (where π^S is optimal) and when $c = 10$ (where π^C is optimal). The incremental value of pooling more than 5 (10) supervisors is small when $c = 2$ ($c = 10$).

5. Conclusion

In this paper, we characterize the optimal policy for a two-stage service system with customer abandonments. There are subordinates who perform the first-stage service on their own and supervisors who work together with the subordinates to complete the second-stage service and also have other responsibilities beyond serving the customers in the system. We show that there are only two optimal policies, namely the supervisors start working on the second-stage service either when the subordinates can no longer serve new customers in the first stage or as soon as there is a customer waiting for the second-stage service. The optimality of the two policies depends on how the abandonment cost compares with a threshold that is a function of the other model parameters. We also investigate the effects of pooling supervisors (and their associated subordinates) and show that pooling improves the system performance. In a future research, we are interested in characterizing the optimal policies when there is not unlimited work, and instead the customers and/or other responsibilities of the supervisors arrive according to Poisson processes.

Acknowledgments. This work was supported in part by the National Science Foundation under the grant CMMI-1536990. The authors thank Dr. E. Lerzan Örmeci and two anonymous referees for their valuable suggestions for improving this paper.

Conflict of interest. The authors have not disclosed any competing interests.

References

[1] Ahn, H.-S., Duenyas, I., & Lewis, M.E. (2002). Optimal control of a two-stage tandem queuing system with flexible servers. *Probability in the Engineering and Informational Sciences* 16(4): 453–469.
[2] Andradóttir, S. & Ayhan, H. (2005). Throughput maximization for tandem lines with two stations and flexible servers. *Operations Research* 53(3): 516–531.
[3] Andradóttir, S. & Ayhan, H. (2021). Optimizing the interaction between residents and attending physicians. *European Journal of Operational Research* 290(1): 210–218.
[4] Andradóttir, S., Ayhan, H., & Down, D.G. (2007). Dynamic assignment of dedicated and flexible servers in tandem lines. *Probability in the Engineering and Informational Sciences* 21(4): 497–538.
[5] Ansari, S., Debo, L., & Iravani, S. (2019). Optimal policy in single-server multi-class queuing systems with abandonment. Available at SSRN: https://ssrn.com/abstract{\mathsurround=\opskip$=$}3453227 or http://dx.doi.org/10.2139/ssrn.3453227.
[6] Atar, R., Giat, C., & Shimkin, N. (2010). The $c\mu/\theta$ rule for many-server queues with abandonment. *Operations Research* 58(5): 1427–1439.
[7] Batt, R.J. & Terwiesch, C. (2015). Waiting patiently: An empirical study of queue abandonment in an emergency department. *Management Science* 61(1): 39–59.
[8] Berman, O. & Sapna-Isotupa, K.P. (2005). Optimal control of servers in front and back rooms with correlated work. *IIE Transactions* 37(2): 167–173.
[9] Bhulai, S. & Koole, G. (2003). A queueing model for call blending in call centers. *IEEE Transactions on Automatic Control* 48: 1434–1438.
[10] Bhulai, S., Farenhorst-Yuan, T., Heidergott, B., & van der Laan, D. (2012). Optimal balanced control for call centers. *Annals of Operations Research* 201: 39–62.
[11] Cattani, K. & Schmidt, G.M. (2005). The pooling principle. *INFORMS Transactions on Education* 5(2): 17–24.
[12] Derlet, R.W. & Richards, J.R. (2000). Overcrowding in the nation's emergency departments: Complex causes and disturbing effects. *Annals of Emergency Medicine* 35(1): 63–68.
[13] Derlet, R.W., Richards, J.R., & Kravitz, R.L. (2001). Frequent overcrowding in U.S. emergency departments. *Academic Emergency Medicine* 8(2): 151–155.
[14] Deslauriers, A., L'Ecuyer, P., Pichitlamken, J., Ingolfsson, A., & Avramidis, A.N. (2007). Markov chain models of a telephone call center with call blending. *Computers & Operations Research* 34: 1616–1645.

https://doi.org/10.1017/S0269964823000013 Published online by Cambridge University Press

[15] Down, D.G., Koole, G., & Lewis, M.E. (2011). Dynamic control of a single-server system with abandonments. *Queueing Systems* 67: 63–90.
[16] Duenyas, I., Gupta, D., & Olsen, T.L. (1998). Control of a single-server tandem queueing system with setups. *Operations Research* 46(2): 218–230.
[17] Garnett, O., Mandelbaum, A., & Reiman, M. (2002). Designing a call center with impatient customers. *Manufacturing & Service Operations Management* 4(3): 208–227.
[18] Khudyakov, P., Feigin, P.D., & Mandelbaum, A. (2010). Designing a call center with an IVR (interactive voice response). *Queueing Systems* 66(3): 215–237.
[19] Koeleman, P., Bhulai, S., & van Meersbergen, M. (2012). Optimal patient and personnel scheduling policies for care-at-home service facilities. *European Journal of Operational Research* 219: 557–563.
[20] Lippman, S.A. (1975). Applying a new device in the optimization of exponential queuing systems. *Operations Research* 23(4): 687–710.
[21] Muth, E.J. (1979). The reversibility property of production lines. *Management Science* 25: 152–158.
[22] Pang, G. & Perry, O. (2015). A logarithmic safety staffing rule for contact centers with call blending. *Management Science* 61: 73–91.
[23] Puterman, M.L. (2014). *Markov decision processes: Discrete stochastic dynamic programming*. Hoboken, New Jersey: John Wiley & Sons.
[24] Wang, J., Abouee-Mehrizi, H., Baron, O., & Berman, O. (2019). Tandem queues with impatient customers. *Performance Evaluation* 135: Article 102011.
[25] Wang, J., Baron, O., & Scheller-Wolf, A. (2015). M/M/c queue with two priority classes. *Operations Research* 63: 733–749.
[26] Weerasinghe, A.P. & Mandelbaum, A. (2013). Abandonment versus blocking in many-server queues: Asymptotic optimality in the QED regime. *Queueing Systems* 75: 279–337.
[27] Wein, L.M. (1992). Dynamic scheduling of a multiclass make-to-stock queue. *Operations Research* 40: 724–735.
[28] Yarmohammadian, M.H., Rezaei, F., Haghshenas, A., & Tavakoli, N. (2017). Overcrowding in emergency departments: A review of strategies to decrease future challenges. *Journal of Research in Medical Sciences* 22.
[29] Zayas-Cabán, G., Xie, J., Green, L.V., & Lewis, M.E. (2016). Dynamic control of a tandem system with abandonments. *Queueing Systems* 84(3): 279–293.
[30] Zayas-Caban, G., Xie, J., Green, L.V., & Lewis, M.E. (2019). Policies for physician allocation to triage and treatment in emergency departments. *IISE Transactions on Healthcare Systems Engineering* 9(4): 342–356.
[31] Zhao, Y. & Melamed, B. (2007). IPA derivatives for make-to-stock production-inventory systems with lost sales. *IEEE Transactions on Automatic Control* 52: 1491–1495.

Cite this article: Yu Z, Andradóttir S and Ayhan H (2024). Optimal control of supervisors balancing individual and joint responsibilities. *Probability in the Engineering and Informational Sciences* **38**, 130–149. https://doi.org/10.1017/S0269964823000013