Service Quality versus Efficiency in Tandem Systems

Eser Kırkızlar
School of Management
State University of New York – Binghamton
Binghamton, NY 13902-6000, U.S.A.

Sigrún Andradóttir • Hayriye Ayhan
H. Milton Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0205, U.S.A.

Abstract

We study the tradeoff between efficiency and service quality in tandem systems. We reward efficiency by assuming that a revenue is earned each time a job is completed, and we penalize poor service quality by incorporating positive holding costs. We study the dynamic assignment of servers to tasks with the objective of maximizing the long-run average profit. For systems of arbitrary size, generalist servers, and linear or nonlinear holding costs, we determine the server assignment policy that maximizes the profit. For systems with two stations, two specialist servers, and linear holding costs, we show that the optimal server assignment policy is of threshold type and determine the value of this threshold as a function of the revenue and holding cost. We also provide numerical results that suggest that the optimal policy has a threshold structure for nonlinear holding costs. Finally, for larger systems with specialist servers, we propose effective server assignment heuristics.

Keywords: Flexible servers, finite buffers, profit maximization, holding cost, queueing control.

1 Introduction

The tradeoff between service quality and efficiency is a common issue in manufacturing and service systems. Attempts to improve one of these performance measures may worsen the other. For example, increasing server utilization causes longer waiting times, and enforcing shorter waiting times may decrease the throughput. In this paper, we study the tradeoff between service quality and efficiency in tandem queueing systems with flexible servers. More specifically, we assume that the service quality deteriorates as a customer spends more time in the system, and that the efficiency is directly proportional to the throughput. Our main objective is to determine a dynamic server assignment policy that maximizes the profit by simultaneously achieving high throughput and short waiting times.

The queueing literature concentrates on the goal of either increasing service quality (by reducing waiting times, queue sizes, abandonments, holding costs, etc.) or increasing efficiency (by increasing throughput, revenue, utilization, etc.). Very few papers study these two conflicting objectives at the same time. We now provide a summary of the work done for the systems with flexible workforce, by classifying them according to their objective.

Most of the previous work concentrates on the holding cost minimization problem. More specifically, Harrison and López [20], Bell and Williams [15, 16], Ahn, Duenyas, and Zhang [4], and Mandelbaum and Stolyar [29] study holding cost minimization in parallel queueing systems; and Rosberg, Varaiya, and Walrand [34], Farrar [18], Iravani, Posner, and Buzacott [22], Ahn, Duenyas, and Zhang [2], Ahn, Duenyas, and Lewis [3], Kaufman, Ahn, and Lewis [23], Wu, Lewis, and Veatch [39], Pandelis [32], and Wu, Down, and Lewis [40] study tandem systems. These papers provide guidelines on how to effectively assign the flexible servers dynamically to stations in order to minimize the holding cost in the system. These papers study either systems with infinite buffers and outside arrivals or clearing systems with no outside arrivals.

The papers that study the throughput maximization problem mostly concentrate on tandem systems, although a few study general queueing networks. Andradóttir, Ayhan, and Down [6, 9, 10], Andradóttir and Ayhan [5], and Kırkızlar, Andradóttir, and Ayhan [24, 25] study the dynamic server assignment problem in various settings with finite buffers, including tandem systems with failures or non-exponential service time distributions. Tassiulas and Ephrimedes [36], Tassiulas and Bhattacharya [35], and Andradóttir, Ayhan, and Down [7, 8] consider the throughput maximization problem in a queueing network with infinite buffers and outside arrivals. Finally, McClain, Thomas, and Sox [31], Zavadlav, McClain, and Thomas [41], Bartholdi and Eisenstein [12], Bartholdi, Bunimovich, and Eisenstein [13], Bartholdi, Eisenstein, and Foley [14], Gel, Hopp, and Van Oyen [19], Hopp, Tekin, and Van Oyen [21], Ahn and Righter [1], and Lim and Yang [27] are some of the main papers that study line balancing via server flexibility.

We are aware of fewer queueing papers that study optimization problems that include both throughput and costs. In particular, Mayorga, Taaffe, and Arumugam [30] study a finite-horizon discounted profit maximization problem with holding and setup costs. They provide heuristic server assignment policies for a tandem line with two homogeneous stations, two flexible servers, and an infinite buffer between the stations. By contrast, in this paper we consider the long-run average profit maximization problem with holding costs and provide the exact optimal server assignment policy for a system with general service rates and finite buffers. Note that general service rates (as opposed to systems with homogeneous servers or homogeneous tasks) and finite buffers are more realistic representations of the actual systems, however they also make our problem substantially more difficult to analyze. Andradóttir, Ayhan, and Kırkızlar [11] study the profit maximization problem in a tandem line with two stations, two flexible servers, a finite buffer between the stations, and a positive setup cost. However, due to the complexity of the problem, Andradóttir et al. [11] only provide the optimal server assignment policy for systems with small buffer sizes, but in this work we study a system with a buffer of any finite size.

The main contributions of this paper can be summarized as follows:

- We identify the optimal server assignment policy for systems with arbitrary number of stations, arbitrary number of generalist servers, and linear or nonlinear holding costs. As systems with homogeneous tasks or servers are special cases of systems with generalist servers, this result is applicable to a wide range of systems.
- For systems with two stations, two specialist servers, and linear holding costs, we show that the optimal policy is of threshold type and determine the value of the threshold. This is an important achievement because determining the optimal policy in a revenue maximization problem is significantly harder than throughput maximization and cost minimization problems (such problems can be considered as special cases of the profit maximization problem).
- For systems with two stations, two specialist servers, and nonlinear holding costs, we provide results of numerical experiments that suggest that the optimal server assignment policy also has a threshold structure in this setting.
- For systems with more than two stations and more than two specialist servers, we propose effective server assignment heuristics.

Furthermore, we believe that this work has the potential of starting a new line of research in the area of dynamic server assignment. In particular, the problems studied in the papers cited above (including the ones studying parallel systems and line balancing problems) can be potentially adapted to the profit maximization problem. This would also determine which server assignment policies and system configurations can deal with conflicting objectives better than the others.

The remainder of this paper is organized as follows. In Section 2, we provide a detailed description of our system and formulate our problem as a discrete-time Markov decision problem (MDP). In Section 3, we determine the optimal server assignment policy for systems with two stations and two flexible servers. In Section 4, we present our results about the optimal server assignment policy for different holding cost structures and for larger systems. Finally, in Section 5, we make some concluding remarks. Proofs of our main results are given in the Appendix.

2 Problem Formulation

We consider a tandem line with $N \geq 2$ stations and $M \geq 1$ flexible servers. We assume that there is an infinite supply of jobs in front of the first station, infinite room for completed jobs after the last station, and a buffer of size B_j between stations $j-1$ and j, where $j \in \{2, \dots, N\}$. A station is blocked after a service completion at that station if the consecutive buffer is full (i.e., the line operates under the blocking-after-service mechanism), and travel times of the servers and setup times at the stations are assumed to be negligible. Let μ_{ij} denote the deterministic rate with which server $i \in \{1, \dots, M\}$ works at station $j \in \{1, \dots, N\}$. Multiple servers can collaborate on the same job, in which case their service rates are additive. Service times at each station $j \in \{1, \dots, N\}$ are independent and exponentially distributed with mean $m(j)$. Without loss of generality, we assume that $\sum_{j=1}^{N} \mu_{ij} > 0$ for $i \in \{1, \dots, M\}$ (otherwise some servers have zero service rate at all tasks, and this is equivalent to the case with fewer servers) and that $\sum_{i=1}^{M} \mu_{ij} > 0$ for $j \in \{1, \dots, N\}$ (because otherwise there is a station where nobody is trained at, and all policies will result in zero throughput). Each time a job is completed at station N, we assume that a revenue $r > 0$ is earned. Furthermore, for each job that has completed service at station $j-1$ but has not completed service at station j, a holding cost $h_j \geq 0$ per time unit is incurred for $j \in \{2, \dots, N\}$. We further assume that $\sum_{j=2}^{N} \frac{h_j}{\Sigma_j} < r$, where $\Sigma_j = \sum_{i=1}^{M} \mu_{ij}$, so that the expected profit is positive when the jobs do not wait in the buffer at any station and are served with the maximum service rate at each station, because otherwise the policy that idles all servers is optimal. Without loss of generality, we assume that $m(j) = 1$ for all $j \in \{1, \dots, N\}$ and that $r = 1$. Our objective is to determine the dynamic server assignment policy that maximizes the long-run average profit in the system described above.

For all server assignment policies π and $t \geq 0$, let $X_{\pi,j}(t) \in \{0, 1, \dots, B_j + 2\}$ denote the number of jobs that have completed service at station j and are either waiting for service or in service at station $j+1$ at time t under policy π for $j \in \{1, \dots, N-1\}$. Let $\mathcal{S} \subset \mathbb{N}^{N-1}$ be the state space corresponding to the stochastic process $X_\pi(t) = (X_{\pi,1}(t), \dots, X_{\pi,N-1}(t))$.

Decision epochs are the service completion times at any station, so that decisions are made when changes to the state of the system are observable. Theorem 9.1.8 of Puterman [33] shows the existence of an optimal stationary deterministic policy when the state and action spaces are finite. Hence, without loss of generality, we restrict ourselves to the set Π of all Markovian stationary deterministic policies corresponding to the state space $\mathcal{S}$.

For all $x \in \mathcal{S}$, let A_x denote the set of allowable actions at state x. We let $a_{\sigma_1\sigma_2\ldots\sigma_M} \in A_x$ denote an action at state x, where σ_i is the station to which server $i \in \{1, \ldots, M\}$ is assigned (with the convention that $\sigma_i = 0$ when server i is idle). The decision rule d is chosen such that $d(x) \in A_x$ for all $x \in \mathcal{S}$, and hence the policy $\pi \in \Pi$ corresponding to the decision rule d can be represented as $\pi = (d)^\infty$.

For all $\pi \in \Pi$ and $t \geq 0$, let $D_\pi(t)$ be the number of departures under policy π by time t. Moreover, for all $\pi \in \Pi$ and $t \geq 0$, let $H_\pi(t)$ be the (cumulative) holding cost incurred under the server assignment policy π in the period $[0, t]$. Let

$$P_\pi = \lim_{t\to\infty} \mathbb{E}\left\{\frac{D_\pi(t)}{t} - \frac{H_\pi(t)}{t}\right\} \tag{1}$$

be the long-run average profit under policy $\pi \in \Pi$ (since $r = 1$). Note that the existence of the limit in (1) follows from the strong law of large numbers for Markov chains (see, e.g., Wolff [38], page 164) because the state space of $\{X_\pi(t)\}$ and the immediate rewards are finite (although this limit may depend on the initial distribution of the Markov chain). We are interested in solving the optimization problem:

$$\max_{\pi\in\Pi} P_\pi. \tag{2}$$

Next, we translate our original optimization problem (2) into an equivalent (discrete time) Markov decision problem.

Our assumptions ensure that the stochastic process $\{X_\pi(t)\}$ is a continuous time Markov chain. Under the server assignment policy $\pi = (d)^\infty$ and for all $x, x' \in \mathcal{S}$, let $q_d(x, x')$ denote the rate at which the continuous time Markov chain $\{X_\pi(t)\}$ goes from state x to state x'. Then, for all $\pi = (d)^\infty \in \Pi$, there exists a scalar $q_\pi \leq \sum_{i=1}^M \max_{1\leq j\leq N} \mu_{ij} < \infty$ such that the transition rates $\{q_d(x, x')\}$ of $\{X_\pi(t)\}$ satisfy $\sum_{x'\in\mathcal{S}, x'\neq x} q_d(x, x') \leq q_\pi$ for all $x \in \mathcal{S}$. Hence, $\{X_\pi(t)\}$ is uniformizable for all $\pi \in \Pi$ (as suggested in Lippman [28]). The fact that $\{X_\pi(t)\}$ is uniformizable will be used to translate the original optimization problem (2) into an equivalent (discrete time) Markov decision problem. We denote the corresponding discrete time Markov chain by $\{X'_\pi(k)\}$. Hence, $\{X'_\pi(k)\}$ has state space $\mathcal{S}$ and transition probabilities $p_d(x, x') = q_d(x, x')/q_\pi$ if $x' \neq x$ and $p_d(x, x) = 1 - \sum_{x'\in\mathcal{S}, x'\neq x} q_d(x, x')/q_\pi$ for all $x \in \mathcal{S}$. We will generate sample paths of the continuous time Markov chain $\{X_\pi(t)\}$, where $\pi \in \Pi$, by generating a Poisson process $\{K_\pi(t)\}$ with rate q_π and the next state of the continuous time Markov chain $\{X_\pi(t)\}$ is generated (independently from $\{K_\pi(t)\}$) using the transition probabilities of the discrete time Markov chain $\{X'_\pi(k)\}$ at the times of the events of $\{K_\pi(t)\}$ (so that there may be transitions from a state back to itself).

For all $x, x' \in \mathcal{S}$, let

$$R_d(x, x') = \begin{cases} 1 - \frac{\sum_{j=2}^N h_j x_j}{q_\pi} & \text{if } x' \in \mathcal{D}_x, \\ -\frac{\sum_{j=2}^N h_j x_j}{q_\pi} & \text{otherwise,} \end{cases}$$

where $\mathcal{D} = \{x \in \mathcal{S} : x_{N-1} > 0\}$, and $\mathcal{D}_x = \{(x_1, \ldots, x_{N-2}, x_{N-1} - 1)\}$ for all $x \in \mathcal{D}$, and $\mathcal{D}_x = \emptyset$ for all $x \notin \mathcal{D}$ (note that the set $\mathcal{D}_x$ has at most one state x'). Note that $\sum_{j=2}^N h_j x_j$ is the holding cost per unit time in state x and that the expected time between transitions is $\frac{1}{q_\pi}$. Hence an expected holding cost of $\frac{\sum_{j=2}^N h_j x_j}{q_\pi}$ is incurred for each transition in state x. Moreover, a revenue of $r = 1$ is gained if there is a departure from the system. Hence, $R_d(x, x')$ equals the expected immediate reward associated with each transition from state x to state x'. Consequently, for all $\pi = (d)^\infty \in \Pi$,

$$P_\pi = \lim_{t \to \infty} \mathbb{E}\left\{ \frac{K_\pi(t)}{t} \times \frac{1}{K_\pi(t)} \sum_{k=1}^{K_\pi(t)} R_d(X'_\pi(k-1), X'_\pi(k)) \right\}. \tag{3}$$

It is clear that $K_\pi(t)/t \to q_\pi$ almost surely as $t \to \infty$ for all $\pi \in \Pi$ by the elementary renewal theorem. Moreover, it is clear that for all $\pi \in \Pi$, the limit $\lim_{K \to \infty} \sum_{k=1}^K R_d(X'_\pi(k-1), X'_\pi(k))/K$ exists almost surely by the strong law of large numbers for Markov chains (see, e.g., Wolff [38], page 164) , although the limit may depend on $\{X'_\pi(0)\}$ and it may be random (see also Section 3.8 of Kulkarni [26]). Uniform integrability shows that for all $\pi \in \Pi$, we have

$$\begin{aligned} P_\pi &= q_\pi \mathbb{E}\left\{ \lim_{K \to \infty} \frac{1}{K} \sum_{k=1}^K R_d(X'_\pi(k-1), X'_\pi(k)) \right\} \\ &= q_\pi \lim_{K \to \infty} \mathbb{E}\left\{ \frac{1}{K} \sum_{k=1}^K R_d(X'_\pi(k-1), X'_\pi(k)) \right\} \end{aligned}$$

(see for example the corollary to Theorem 25.12 in Billingsley [17]) since

$$\left| \frac{1}{K} \sum_{k=1}^K R_d(X'_\pi(k-1), X'_\pi(k)) \right| \le 1 + \frac{N}{q_\pi} \Big(\max_{2 \le k \le N} |B_k| + 2 \Big) \Big(\max_{2 \le k \le N} |h_k| \Big) < \infty$$

for all $K \geq 1$ and $\sup_{t \geq 0} \mathbb{E}\{[K_\pi(t)/t]^2\} < \infty$ (because $K_\pi(t)$ is a Poisson random variable with mean $q_\pi t$). Hence, the optimization problem (2) has the same solution as the optimization problem

$$\max_{\pi \in \Pi} q_\pi \lim_{K \to \infty} \mathbb{E}\left\{ \frac{1}{K} \sum_{k=1}^K R_d(X'_\pi(k-1), X'_\pi(k)) \right\}.$$

Then, using the strong law of large numbers for Markov chains, we obtain for all $\pi = (d)^\infty \in \Pi$ that

$$\lim_{K\to\infty} \frac{1}{K}\sum_{k=1}^{K} R_d(X'_\pi(k-1), X'_\pi(k)) = \lim_{K\to\infty} \frac{1}{K}\sum_{k=1}^{K} R'_d(X'_\pi(k-1)) \quad \text{a.s.},$$

where

$$\begin{aligned} R'_d(x) &= \sum_{x'\in\mathcal{S}} p_d(x,x')R_d(x,x') \\ &= \sum_{x'\in\mathcal{D}_x} p_d(x,x') - \sum_{x'\in\mathcal{S}} p_d(x,x')\frac{\sum_{j=2}^{N} h_j x_j}{q_\pi} \\ &= \sum_{x'\in\mathcal{D}_x} \frac{q_d(x,x')}{q_\pi} - \frac{\sum_{j=2}^{N} h_j x_j}{q_\pi} \end{aligned}$$

for all $x \in \mathcal{S}$ (note that both limits may be random and may depend on $\{X'_\pi(0)\}$). Uniform integrability now gives that

$$\lim_{K\to\infty} \mathbb{E}\left\{\frac{1}{K}\sum_{k=1}^{K} R_d(X'_\pi(k-1), X'_\pi(k))\right\} = \lim_{K\to\infty} \mathbb{E}\left\{\frac{1}{K}\sum_{k=1}^{K} R'_d(X'_\pi(k-1))\right\}.$$

Hence, the optimization problem (2) has the same solution as the optimization problem

$$\max_{\pi\in\Pi} q_\pi \lim_{K\to\infty} \mathbb{E}\left\{\frac{1}{K}\sum_{k=1}^{K} R'_d(X'_\pi(k-1))\right\}.$$

Finally, it is clear that if

$$R''_d(x) = \sum_{x'\in\mathcal{D}_x} q_d(x,x') - \sum_{j=2}^{N} h_j x_j$$

for all $x \in \mathcal{S}$ and $\pi = (d)^\infty \in \Pi$, then the optimization problem (2) has the same solution as the Markov decision problem

$$\max_{\pi\in\Pi} \lim_{K\to\infty} \mathbb{E}\left\{\frac{1}{K}\sum_{k=1}^{K} R''_d(X'_\pi(k-1))\right\}. \tag{4}$$

In the remainder of this paper, we analyze the alternative formulation (4) of the original optimization problem (2).

3 Main Results

In this section, we provide our main results about the profit maximization problem in tandem systems with finite buffers. More specifically, we characterize the optimal server assignment policy for systems with arbitrary numbers of generalist servers and stations in Section 3.1, and for systems with two specialist servers and two stations in Section 3.2.

3.1 Systems with Generalist Servers

In this section, we consider a tandem line with generalist servers whose service rates at a station can be written as the product of the server's speed and a constant related to the complexity level of the task at the station. In other words, we assume that $\mu_{ij} = \mu_i \gamma_j$ for $i, j \in \{1, 2\}$. The following technical assumption is used in the description of the optimal server assignment policy:

Assumption S. *The service requirements $S_{j,k}$ of job $k \geq 1$ at station $j \in \{1, \ldots, N\}$ are independent and identically distributed with mean 1. Moreover, if there is a job in service at station j at time $t \geq 0$, then the expected remaining service requirement at station j of that job is bounded above by a scalar $1 \leq S < \infty$. Finally, the service discipline is either nonpreemptive or preemptive-resume.*

The following theorem characterizes the optimal server assignment policy for systems with arbitrary size:

Theorem 3.1. *Suppose that Assumption* **S** *holds. For a system with N stations in tandem, M generalist servers, and finite buffers between the stations, the expedite policy π^e of Van Oyen, Gel, and Hopp [37] (in which all servers work as a team that moves with each job through the line) is optimal, with the long-run average profit*

$$P_{\pi^e} = \frac{\sum_{i=1}^{M} \mu_i}{\sum_{j=1}^{N} 1/\gamma_j} \Big(1 - \sum_{j=2}^{N} \frac{h_j}{\Sigma_j}\Big).$$

Proof: Define the holding cost per item produced up to time t under policy $\pi \in \Pi$ as $C_\pi(t) = H_\pi(t)/D_\pi(t)$. Moreover, note that for all $\pi \in \Pi$, we have

$$\begin{aligned} P_\pi &= \lim_{t\to\infty} \mathbb{E}\left\{\frac{D_\pi(t)}{t}\Big(1 - \frac{H_\pi(t)}{D_\pi(t)}\Big)\right\} \\ &= \lim_{t\to\infty} \mathbb{E}\left\{\frac{D_\pi(t)}{t}(1 - C_\pi(t))\right\} \\ &= \lim_{t\to\infty} \mathbb{E}\left\{\frac{D_\pi(t)}{t}\right\} \lim_{t\to\infty} \mathbb{E}\left\{(1 - C_\pi(t))\right\}. \end{aligned} \tag{5}$$

Note that the existence of the limits in (5) follows because the state space and the immediate rewards are finite. First, we determine an upper bound on the long-run average profit. Andradóttir, Ayhan, and Down [9] show that any nonidling server assignment policy maximizes the long-run average throughput, with the throughput being equal to $\sum_{i=1}^{M} \mu_i \Big/ \sum_{j=1}^{N} 1/\gamma_j$. Note that even if no waiting occurs before starting a job in any station, an expected holding cost of $\sum_{j=2}^{N} \frac{h_j}{\Sigma_j}$ is incurred. Hence, we can conclude that $\lim_{t\to\infty} \mathbb{E}(1 - C_\pi(t)) \leq (1 - \sum_{j=2}^{N} \frac{h_j}{\Sigma_j})$. Combining these with (5), for all $\pi \in \Pi$ we obtain

$$P_\pi \leq \frac{\sum_{i=1}^{M} \mu_i}{\sum_{j=1}^{N} 1/\gamma_j} \Big(1 - \sum_{j=2}^{N} \frac{h_j}{\Sigma_j}\Big). \tag{6}$$

Next, we show that the long-run average profit of the expedite policy is equal to the upper bound provided in (6). Note that under the expedite policy there is only one job in the system that incurs a holding cost. Hence, no waiting occurs before service at any station. Hence, we conclude that the long-run average profit under the expedite policy is equal to $\left(\sum_{i=1}^{M}\mu_i \Big/ \sum_{j=1}^{N} 1/\gamma_j\right)\,(1-\sum_{j=2}^{N}\frac{h_j}{\Sigma_j})$ because expedite policy is a nonidling policy. This, together with inequality (6), proves the result. □

Note that the expedite policy is also optimal when the holding cost is not a linear function of the number of jobs. More specifically, assume that a holding cost of $h_j^c(s)$ is incurred when there are s jobs waiting to be served or being served at station j, where $h_j^c(s)$ is a nondecreasing function of s for $j \in \{1,\dots,N\}$ and $s \in \{0,\dots,B_j+2\}$. The proof of Theorem 3.1 shows that the expedite policy is optimal with the long-run average profit

$$P_{\pi^e} = \frac{\sum_{i=1}^{M}\mu_i}{\sum_{j=1}^{N}1/\gamma_j}\Big(1-\sum_{j=2}^{N}\frac{h_j^c(1)}{\Sigma_j}\Big).$$

3.2 Systems with Specialist Servers

In this section, we consider systems with specialist servers whose service rates cannot be written as products of two terms. Specialist servers have a general service rate structure, and hence their analysis is more complicated than that of generalist servers. Consider a tandem line with two stations and two flexible servers. Since we can relabel the servers if necessary, without loss of generality assume that $\mu_{11}\mu_{22} \ge \mu_{12}\mu_{21}$. For $i,j \in \mathbb{N}^+$ and $i<j$, let $f_1(i,j)$ and $f_2(i,j)$ be defined as follows:

$$f_1(i,j) = \begin{cases} \mu_{22}^{i-1}(\mu_{12}+\mu_{22})(\mu_{22}-\mu_{11})(\mu_{11}\mu_{22}-\mu_{12}\mu_{21})(\mu_{22}^{j-i}-\mu_{11}^{j-i}) & \text{if } \mu_{11}\neq\mu_{22}; \\ (j-i)\mu_{22}(\mu_{12}+\mu_{22})(\mu_{22}^2-\mu_{12}\mu_{21}) & \text{if } \mu_{11}=\mu_{22}. \end{cases}$$

and

$$f_2(i,j) = \begin{cases} \frac{\mu_{22}^{i-2}}{\mu_{11}^{i-1}}(\mu_{22}-\mu_{11})\Big[\mu_{22}^{j-i}\Big(\mu_{12}\sum_{k=1}^{j-1}k(\frac{\mu_{11}}{\mu_{22}})^{k-1}+\mu_{22}\sum_{k=1}^{j}k(\frac{\mu_{11}}{\mu_{22}})^{k-1}\Big) \\ \times\Big(\mu_{12}\mu_{21}(\mu_{22}^{i-1}-\mu_{11}^{i-1})+(\mu_{12}+\mu_{21})(\mu_{22}^{i}-\mu_{11}^{i})+\mu_{22}^{i+1}-\mu_{11}^{i+1}\Big) \\ -\Big(\mu_{12}\sum_{k=1}^{i-1}k(\frac{\mu_{11}}{\mu_{22}})^{k-1}+\mu_{22}\sum_{k=1}^{i}k(\frac{\mu_{11}}{\mu_{22}})^{k-1}\Big) \\ \times\Big(\mu_{12}\mu_{21}(\mu_{22}^{j-1}-\mu_{11}^{j-1})+(\mu_{12}+\mu_{21})(\mu_{22}^{j}-\mu_{11}^{j})+\mu_{22}^{j+1}-\mu_{11}^{j+1}\Big)\Big] & \text{if } \mu_{11}\neq\mu_{22}; \\ \Big(\mu_{12}\frac{(j-1)j-(i-1)i}{2}+\mu_{22}\frac{j(j+1)-i(i+1)}{2}\Big)\Big(\mu_{12}\mu_{21}+(\mu_{12}+\mu_{21})\mu_{22}+\mu_{22}^2\Big) & \text{if } \mu_{11}=\mu_{22}. \end{cases}$$

We need the following four lemmas in the proof of Theorem 3.2 that identifies a server assignment policy that maximizes the long-run average profit. The proofs of the following lemmas are provided in the Appendix.

Lemma 3.1. *For a tandem line with $M=2$, there exists an optimal policy that is non-idling.*

Lemma 3.2. *We have* $\frac{f_1(i,i+1)}{f_2(i,i+1)} > \frac{f_1(i+1,i+2)}{f_2(i+1,i+2)}$ *for* $i \in I\!N^+$.

Lemma 3.3. *We have* $\frac{f_1(i,j)}{f_2(i,j)} \geq 0$ *for* $i, j \in I\!N^+$ *and* $i < j$.

Lemma 3.4. $\frac{f_1(i,j)}{f_2(i,j)}$ *is nonincreasing in* i *for all* $i, j \in I\!N^+$ *and* $i < j$.

Note that for $i < j$, $f_1(i,j) \geq 0$ trivially and $f_2(i,j) \geq 0$ as shown by Lemma 3.3. Moreover, we show in the Appendix that the functions $f_1(i,j)$ and $f_2(i,j)$ are obtained in the policy improvement step of policy iteration algorithm. More specifically, these functions are found by comparing the actions a_{12} and a_{22} in state i, when state j is the smallest state where both servers are assigned to station 2.

Now we characterize the optimal server assignment policy for a tandem line with two stations and two specialist servers. The proof of the following theorem is provided in the Appendix.

Theorem 3.2. *Consider a system with two stations, two flexible servers, and a buffer of size* B *between the stations. If* $h > \frac{f_1(1,2)}{f_2(1,2)}$, *then let* $s = 1$; *or if* $h \leq \frac{f_1(B+1,B+2)}{f_2(B+1,B+2)}$, *then let* $s = B + 2$; *or let* $s \in \{2, \ldots, B+1\}$ *be such that* $\frac{f_1(s,s+1)}{f_2(s,s+1)} < h \leq \frac{f_1(s-1,s)}{f_2(s-1,s)}$. *Let* $\pi = (d)^\infty$, *where*

$$d(i) = \begin{cases} a_{11} & \text{for } i = 0, \\ a_{12} & \text{for } 1 \leq i \leq s-1, \\ a_{22} & \text{for } s \leq i \leq B+2. \end{cases}$$

Then π *maximizes the long-run average profit and the recurrent states are* $\{0, \ldots, s\}$. *Moreover, the optimal actions* $d(i)$ *are unique for* $i \in \{0, \ldots, s\}$ *if* $\mu_{11}\mu_{22} > \mu_{12}\mu_{21}$ *and* $h \neq \frac{f_1(s-1,s)}{f_2(s-1,s)}$. *However, in the transient states* $\{s+1, \ldots, B+2\}$, *any action that eventually takes the process to a state in* $\{0, \ldots, s\}$ *can be selected.*

Note that Lemmas 3.2 and 3.3 guarantee that h can belong to only one of the intervals described in Theorem 3.2.

The policy of Theorem 3.2 is a “threshold” policy, where each server is primarily assigned to a station, the server assigned to station 2 only switches to station 1 to avoid idleness, and the server assigned to station 1 only switches to station 2 when the number of jobs in the buffer reaches a certain threshold (in other words, this policy has the effect of reducing the size of the buffer that is used in the system). More specifically, when the service rates satisfy the condition $\mu_{11}\mu_{22} \geq \mu_{12}\mu_{21}$, server 1 (2) is primarily assigned to station 1 (2), server 2 moves to station 1 when station 2 is idle, and server 1 moves to station 2 when the jobs in the buffer reach a certain threshold. The value of this threshold decreases as the holding cost increases (this follows from Lemma 3.2).

Remark 3.1. *The optimal server assignment policy of Theorem 3.2 is the same as*

- *the policy π_T^* that maximizes the throughput (see, Andradóttir et al. [6]) when $h \leq \frac{f_1(B+1,B+2)}{f_2(B+1,B+2)}$;*
- *the policy π_H^* that minimizes the holding cost (i.e., the expedite policy of Van Oyen et al. [37]) when $h > \frac{f_1(1,2)}{f_2(1,2)}$.*

Moreover, for intermediate values of h, the optimal policy is always a compromise between the two policies mentioned above (i.e., it is still a threshold policy but neither π_T^* nor π_H^*).

Note that the optimal policy may change or remain the same when the buffer size changes, depending on the value of the threshold. More specifically, suppose that s_B is the optimal threshold for B and assume that the new buffer size is B'. If $s_B \leq B+1$, then the optimal policy does not change if $B' > B$ or if $s_B \leq B'+2$. However, if $s_B = B+2$, the optimal policy may change if either $B' < B$ or both $B' > B$ and $h \leq \frac{f_1(B'+1,B'+2)}{f_2(B'+1,B'+2)}$. Moreover, for a system with holding cost h that satisfies $\frac{f_1(s,s+1)}{f_2(s,s+1)} < h \leq \frac{f_1(s-1,s)}{f_2(s-1,s)}$, having a buffer size larger than $s-2$ does not improve the performance of the system.

Finally, note that the performance loss associated with using the wrong threshold policy can be large. For example, consider the system with $\mu_{11} = 10$, $\mu_{12} = 0.5$, $\mu_{21} = 0.5$, $\mu_{22} = 15$, and $B = 20$. When $h = 0.35$, the profit of π_H^* is 6.12. However, the optimal policy has $s = 15$ and the optimal profit is 9.45 (hence, using the optimal threshold improves the profit of π_H^* by more than 50%). Similarly, when $h = 4$, the profit of π_T^* is 2.05. However, the optimal policy has $s = 2$ and the optimal profit is 5.15 (hence, using the optimal threshold improves the profit of π_T^* by more than 100%).

4 Numerical Results

In this section, we present numerical results for the profit maximization problem in systems with either nonlinear holding costs or more than two stations and servers. More specifically, we provide simulation results for systems with two stations, two flexible servers, and nonlinear holding costs in Section 4.1 that show that the optimal server assignment policy is of threshold type even when the holding cost is nonlinear. In Section 4.2, we provide several server assignment heuristics and demonstrate how their performance depend on the length of the line, the buffer size, and the relations between the holding costs at each stage.

4.1 Systems with Nonlinear Holding Costs

Theorem 3.2 shows that the optimal server assignment policy for systems with two stations and two flexible servers is a threshold policy when the holding cost is a linear function of the number of jobs waiting for service or in service at the second station. In this section, we provide numerical results for systems with nonlinear holding costs. More specifically, we assume that either a superlinear holding cost of $h_2^c(s) = hs^2$ or a sublinear holding cost of $h_2^c(s) = h\sqrt{s}$ per unit time is incurred when the system is in state $s \in \{0, \ldots, B+2\}$.

We perform simulations where the service rates are randomly generated from a uniform distribution in the interval $(0.5, 2.5)$, the size of the intermediate buffer is randomly generated from a discrete uniform distribution with range $\{0, \ldots, 10\}$, and the holding cost is randomly generated from a uniform distribution with range $(0, 1)$. Note that the service rates and the holding satisfy the assumption that $\frac{h}{\mu_{12}+\mu_{22}} < 1$ (and hence it is not optimal to idle all servers). In particular, in each experiment we create 100,000 random systems for each nonlinear holding cost structure (i.e., when $h_2^c(s)$ is equal to hs^2 or $h\sqrt{s}$) and determine the optimal server assignment policy using the policy iteration algorithm for weakly communicating Markov chains (note that we can do this because we start the policy iteration algorithm with a policy that results in a communicating Markov chain). In each system, we observe that the structure of the optimal server assignment policy is the same as the one in Theorem 3.2, i.e., regardless of the holding cost structure, the optimal server assignment policy is a threshold-type policy. These results suggest that the structure of the optimal server assignment policy is robust with respect to the holding cost, as long as it is a nondecreasing function of the number of jobs in the system.

Note that this result can be used to develop effective heuristics for systems with $M = N = 2$ and nonlinear holding costs. More specifically, the number of nonidling policies in the policy space Π is 4^{B+1} (because there is one possible action in states $\{0, B+2\}$ and four possible actions in states $\{1, \ldots, B+1\}$). However, the total number of threshold policies is $2 \times (B+2)$ (because there can be two different primary assignments of servers and there can be at most $B+2$ different threshold policies for each primary assignment). Hence, an enumeration of all possible threshold policies is a better alternative to other search algorithms, including policy iteration, in terms of the required computational time.

4.2 Larger Systems

The optimal assignment policy for systems with $M = N > 2$ can be difficult to identify and implement. In particular, numerical results suggest that the optimal policy, obtained by using the policy iteration algorithm, for systems with $M = N = 3$ is still a threshold policy. However, the servers do not necessarily have primary assignments, and each server has multiple thresholds where the server moves between the stations according to the number of jobs at each station and the server's current location. In this section, we propose four easily implementable server assignment heuristics and compare their performance to the optimal server assignment policy.

We study server assignment heuristics with two main parts, namely a Primary Assignment (PA) and a Contingency Plan (CP). The primary assignment determines the station that the server is initially assigned to. The contingency plan determines where the server works if it is not working at the station it is primarily assigned to. Note that among the earlier papers that study the throughput maximization problem, the ones that are most closely related to our work (see, e.g., Andradóttir et al. [6] and Kırkızlar et al. [24]) propose heuristics that employ a contingency plan when a station is starved or blocked. Our heuristics for the profit maximization problem employ a contingency plan when a station is starved or blocked, as well as when a threshold on the number of jobs is reached or the thresholds satisfy a certain condition. The optimal server assignment policies for systems with two stations and two servers also have this form, see Theorem 3.2.

As a primary assignment, we employ the one that maximizes $\Pi_{j=1}^{M}\mu_{ij_i}$, where $j_i \in \{1, \ldots, N\}$ denotes the station server $i \in \{1, \ldots, M\}$ is assigned to. Note that this PA was proposed by Andradóttir et al. [6] and has been shown to outperform other primary assignments considered in Kırkızlar et al. [24] when the objective is to maximize system throughput. Furthermore, this primary assignment has the advantage of assigning only one server to each station, which consequently simplifies the corresponding contingency plans. Finally, this primary assignment is also in accordance with the optimal server assignment policy for small systems, as shown in Theorem 3.2.

In the heuristics, we consider the following contingency plans:

Heuristic 1: Determine a threshold t_j for the number of jobs being served or waiting to be served at each station $j \in \{1, \ldots, N\}$, considering the servers that are primarily assigned to station $j-1$ and j. More specifically, label station $j-1$ as the first station and station j as the second station, and use the logic in Theorem 3.2. When a station is starved but not blocked (in that the threshold at the next station is not exceeded), the server with a PA at that station moves to the closest upstream station that is operating (neither blocked or starved); when it is blocked (i.e., the threshold at the next station is reached), the server with a PA at that station moves to the closest downstream station that is not blocked.

Heuristic 2: Employ Heuristic 1 with the modification that all servers not working at their assigned station are working at the station closest to the end of the line that is not blocked (i.e., the following threshold is not exceeded).

Heuristic 3: Whenever a station is blocked, starved, or its threshold is reached, the server with a PA at this station works at the station that is operating and whose threshold is not exceeded where it has the highest relative rate with respect to the cumulative rate of all servers with primary assignment at that station (compared to the other stations that are operating); i.e., server i works at station j^* where $j^* = \arg\max_{k \in I} \mu_{ik}/SR_k$ and I is again the set of stations that are operating.

Heuristic 4: Employ Heuristic 3 with the modification that if a condition on the thresholds is satisfied, the teamwork policy is used. More specifically, if some or all of the holding costs are high enough so that $\sum_{j=2}^{N} t_j > \frac{1}{2}\sum_{j=2}^{N}(B_j + 2)$ (i.e., the average of all thresholds is larger than half of the average of all maximum possible values for thresholds), then all the servers work as a team at the closest station to the end of the line that is operating.

Note that Heuristic 1 is inspired by the local heuristic defined in Andradóttir et al. [6]. It does not take into account the rest of the line, and its main objective is to make a server get back to working at the station it is primarily assigned to. Heuristic 2 considers the performance of the whole line by employing the servers to pull the jobs from the system when their primarily assigned station is blocked or its threshold is exceeded. Heuristic 3 is a modified version of the best heuristic provided in Kırkızlar et al. [24] for the throughput maximization problem that considers exceeding the thresholds as blocking. Finally, using teamwork policy as a special case, Heuristic 4 attempts to balance the throughput maximization and cost minimization objectives at the same time.

We study systems with $M = N = 3$, $M = N = 4$, and $M = N = 5$. The service rates are drawn independently from a uniform distribution with range (0.5,2.5). In other words, we randomly generate sets of service rates $\{\mu_{ij}\}$, for $1 \leq i \leq M$ and $1 \leq j \leq N$, and each experiment consists of estimating the long-run average profit of such a random system. For each experiment, we also determine the long-run average profit of the optimal policy by using the policy iteration algorithm. In the tables below, we provide the 95% confidence interval for long-run average profit of each heuristic as well as the percentage of the average optimal profit each heuristic achieves. As a benchmark, we also compare our results to the profit of the teamwork policy of Van Oyen et al. [37]. More specifically, in Section 4.2.1 we compare the performance of our heuristics under holding costs of different magnitude, in Section 4.2.2 we study how the buffer sizes affect our heuristics, and in Section 4.2.3 we investigate the effects of increasing and random holding costs at different buffers. Finally, in Section 4.2.4 we summarize the insights gained from the numerical experiments for larger systems, and shortly describe how our heuristics can be improved.

4.2.1 Effects of the Magnitude of the Holding Cost

In this section, we study systems where a constant holding cost at each station was randomly generated from a uniform distribution with range $(0, 0.01)$, $(0, 0.05)$, or $(0, 0.1)$. Note that these ranges were selected because for higher values of the holding cost, the teamwork policy was optimal for most systems. Furthermore, these ranges satisfy the assumption $\sum_{j=2}^{N} \frac{h_j}{\Sigma_j} < 1$, so that the policy that idles all servers is not the optimal policy. Moreover, for all systems, the buffer sizes at each station were randomly chosen independently from a discrete uniform distribution. More specifically, the range for the discrete uniform distributions were $\{0, \ldots, 10\}$, $\{0, \ldots, 5\}$, and $\{0, 1, 2\}$ for systems with $M = N = 3$, $M = N = 4$, and $M = N = 5$, respectively. The results for systems with $M = N = 3$ and $M = N = 4$ were obtained from 10,000 experiments, and the results for systems with $M = N = 5$

was obtained from 1,000 experiments. Note that the magnitude of the buffers as well as the number of experiments were decreased due to the computational effort required to determine the optimal profit in each system.

When the holding cost is selected from the ranges $(0, 0.01)$ or $(0, 0.05)$, we observe in Tables 1 and 2 that Heuristics 1 and 2 perform better when $M = N = 3$, but Heuristic 4 performs the best when $M = N \in \{4, 5\}$. Moreover, we observe that the best heuristic attains at least 96.9% (95.5%) of the optimal profit in each system configuration when $h = 0.01$ ($h = 0.05$), and that the optimality gap between the heuristics and the teamwork policy increases as h decreases.

When the holding cost is selected from the range $(0, 0.1)$, we observe in Table 3 that the best heuristic attains at least 92.9% of the optimal profit in each system configuration and each holding cost structure. Moreover, we observe that Heuristic 2 performs the best when $M = N \in \{3, 4\}$, but Heuristic 4 performs the best when $M = N = 5$.

Table 1: Average Performance of Heuristics for Constant Holding Cost with Range $(0, 0.01)$ and Buffer Sizes with Varying Ranges

	$M = N = 3$, $0 \leq B_2, B_3 \leq 10$		$M = N = 4$, $0 \leq B_2, B_3, B_4 \leq 5$		$M = N = 5$, $0 \leq B_2, B_3, B_4, B_5 \leq 2$	
Policy	performance	% of the optimal	performance	% of the optimal	performance	% of the optimal
Teamwork	1.448 ± 0.004	% 83.092	1.453 ± 0.003	% 79.452	1.463 ± 0.002	% 78.118
Heuristic 1	1.705 ± 0.005	% 97.844	1.755 ± 0.004	% 95.928	1.755 ± 0.003	% 93.738
Heuristic 2	1.705 ± 0.005	% 97.843	1.749 ± 0.004	% 95.618	1.741 ± 0.003	% 93.009
Heuristic 3	1.676 ± 0.005	% 96.182	1.727 ± 0.004	% 94.414	1.730 ± 0.003	% 92.416
Heuristic 4	1.684 ± 0.005	% 96.625	1.778 ± 0.004	% 97.208	1.815 ± 0.003	% 96.921
optimal	1.743 ± 0.005	—	1.829 ± 0.004	—	1.872 ± 0.003	—

Table 2: Average Performance of Heuristics for Constant Holding Cost with Range $(0, 0.05)$ and Buffer Sizes with Varying Ranges

	$M = N = 3$, $0 \leq B_2, B_3 \leq 10$		$M = N = 4$, $0 \leq B_2, B_3, B_4 \leq 5$		$M = N = 5$, $0 \leq B_2, B_3, B_4, B_5 \leq 2$	
Policy	performance	% of the optimal	performance	% of the optimal	performance	% of the optimal
Teamwork	1.442 ± 0.004	% 84.232	1.443 ± 0.003	% 81.169	1.451 ± 0.002	% 79.817
Heuristic 1	1.644 ± 0.005	% 96.081	1.674 ± 0.004	% 94.155	1.691 ± 0.003	% 93.040
Heuristic 2	1.655 ± 0.005	% 96.727	1.686 ± 0.004	% 94.808	1.689 ± 0.003	% 92.934
Heuristic 3	1.615 ± 0.005	% 94.343	1.637 ± 0.004	% 92.074	1.654 ± 0.003	% 90.998
Heuristic 4	1.623 ± 0.005	% 94.844	1.699 ± 0.004	% 95.512	1.750 ± 0.003	% 96.280
optimal	1.711 ± 0.005	—	1.778 ± 0.004	—	1.817 ± 0.003	—

Table 3: Average Performance of Heuristics for Constant Holding Cost with Range $(0, 0.1)$ and Buffer Sizes with Varying Ranges

	$M = N = 3$, $0 \leq B_2, B_3 \leq 10$		$M = N = 4$, $0 \leq B_2, B_3, B_4 \leq 5$		$M = N = 5$, $0 \leq B_2, B_3, B_4, B_5 \leq 2$	
Policy	performance	% of the optimal	performance	% of the optimal	performance	% of the optimal
Teamwork	1.433 ± 0.004	% 85.117	1.431 ± 0.003	% 82.568	1.436 ± 0.002	% 81.429
Heuristic 1	1.571 ± 0.006	% 93.324	1.577 ± 0.005	% 90.966	1.612 ± 0.004	% 91.444
Heuristic 2	1.597 ± 0.005	% 94.838	1.610 ± 0.004	% 92.891	1.613 ± 0.004	% 91.517
Heuristic 3	1.541 ± 0.006	% 91.495	1.527 ± 0.005	% 88.109	1.559 ± 0.004	% 88.401
Heuristic 4	1.551 ± 0.005	% 92.136	1.601 ± 0.005	% 92.388	1.670 ± 0.003	% 94.710
optimal	1.684 ± 0.005	—	1.733 ± 0.004	—	1.763 ± 0.003	—

Note that when $M = N = 2$, the teamwork policy is optimal for some values of the holding cost. Hence, we believe that superior performance of Heuristic 4 for longer lines results from the fact that it employs the teamwork policy for some values of the holding cost. Note that Heuristics 1, 2, and 3 are myopic in the sense that they only consider the holding cost and number of jobs at each station. Hence, they give priority to increase the revenue rather than to limit the holding costs. However, Heuristic 4 considers the holding costs over the all line and hence its performance improves as the line gets longer. The performance of the teamwork policy seems to be stable in all three systems. However, note that the good performance of teamwork policy in longer lines also results from the decrease in the buffer sizes. Since the heuristics performed similarly for different ranges of the holding cost, we will concentrate on the intermediate case where the holding cost is selected from the range $(0, 0.05)$ in the rest of this paper.

4.2.2 Effects of the Buffer Size

In this section, in order to eliminate the effect of buffer sizes, we compare the performance of heuristics for $M = N \in \{3, 4, 5\}$ when constant holding cost was randomly drawn from a continuous uniform distribution with range $(0, 0.05)$, and buffer sizes were randomly drawn from a discrete uniform distribution with ranges $\{0, 1, 2\}$ and $\{0, \ldots, 5\}$. The number of experiments performed were as in Section 4.2.1. Tables 4 and 5 show that the performance of the heuristics become even better for systems with smaller buffer sizes. However, the best heuristic for each system does not change (compared to the systems with the same holding cost range and varying buffer sizes). More specifically, for systems with buffer sizes from the range $\{0, 1, 2\}$, we observe that performance of Heuristics 1 and 2 reach more than 97% of the optimal profit when $M = N = 3$, and performance of Heuristic 4 reaches more than 96% of the optimal profit when $M = N \in \{4, 5\}$. Similarly, for systems with buffer sizes from the range $\{0, \ldots, 5\}$, Heuristics 1 and 2 perform the best when $M = N = 3$, and Heuristic 4 performs the best when $M = N \in \{4, 5\}$. Furthermore, for some heuristics, we observe that even the actual profit (in addition to the percentage of the optimal profit achieved) increases as the buffer sizes decrease. We believe that this results from the fact that the effects of holding costs become more prominent in systems with larger buffer sizes. Note that, due to the long computational time, we have not been able to determine the optimal profit for systems with $M = N = 5$ and the buffer sizes with range $\{0, \ldots, 5\}$. To summarize, our heuristics' performance compared to each other does not change as the buffer sizes change, however they perform slightly better for systems with small buffer sizes.

Table 4: Average Performance of Heuristics for Constant Holding Cost with Range $(0, 0.05)$ and Buffer Sizes with Range $\{0, 1, 2\}$

	$M = N = 3$		$M = N = 4$		$M = N = 5$	
Policy	performance	% of the optimal	performance	% of the optimal	performance	% of the optimal
Teamwork	1.442 ± 0.004	% 86.014	1.443 ± 0.003	% 82.036	1.451 ± 0.002	% 79.817
Heuristic 1	1.636 ± 0.005	% 97.634	1.673 ± 0.004	% 95.090	1.655 ± 0.003	% 93.040
Heuristic 2	1.637 ± 0.005	% 97.689	1.675 ± 0.004	% 95.186	1.615 ± 0.003	% 92.934
Heuristic 3	1.580 ± 0.005	% 94.286	1.626 ± 0.004	% 92.401	1.623 ± 0.003	% 90.998
Heuristic 4	1.593 ± 0.005	% 95.057	1.697 ± 0.004	% 96.430	1.750 ± 0.003	% 96.280
optimal	1.676 ± 0.005	—	1.760 ± 0.004	—	1.817 ± 0.003	—

Table 5: Average Performance of Heuristics for Constant Holding Cost with Range $(0, 0.05)$ and Buffer Sizes with Range $\{0, \ldots, 5\}$

	$M = N = 3$		$M = N = 4$		$M = N = 5$	
Policy	performance	% of the optimal	performance	% of the optimal	performance	% of the optimal
Teamwork	1.442 ± 0.004	% 84.931	1.443 ± 0.003	% 81.169	1.436 ± 0.002	—
Heuristic 1	1.649 ± 0.005	% 97.177	1.674 ± 0.004	% 94.155	1.601 ± 0.004	—
Heuristic 2	1.653 ± 0.005	% 97.411	1.686 ± 0.004	% 94.808	1.611 ± 0.004	—
Heuristic 3	1.607 ± 0.005	% 94.707	1.637 ± 0.004	% 92.074	1.571 ± 0.004	—
Heuristic 4	1.619 ± 0.005	% 95.381	1.699 ± 0.004	% 95.512	1.697 ± 0.004	—
optimal	1.697 ± 0.005	—	1.778 ± 0.004	—	—	—

4.2.3 Effects of the Increasing and Random Holding Costs at Different Buffers

In this section, we try to understand the effects of increasing and random holding costs at different buffers (as opposed to a constant holding cost at all buffers). More specifically, In each experiment three different holding cost structures were considered: (i) a constant holding cost at each station was randomly generated from a uniform distribution with range $(0, 0.05)$, (ii) h_2 was chosen randomly from a range $(0, 0.05)$, and h_j was generated from a uniform distribution with range $(h_{j-1}, 0.05)$ for $j > 2$, and (iii) each holding cost was generated randomly and independently from each other from a uniform distribution with range $(0, 0.05)$. Moreover, the buffer sizes and the number of experiments were chosen as in Section 4.2.1.

Comparison of Tables 2, 6, and 7 show that the effects of dependency between the holding costs at different buffers is negligible. More specifically, we observe that performance of the heuristics and the optimal server assignment policy is almost identical when constant holding costs or independent holding costs are used at each buffer. When the increasing holding cost structure is used, the performance of our heuristics are slightly worse when $M = N \in \{4, 5\}$ (however the difference as a percentage of the optimal profit is less than 1%). Hence, we can conclude that our heuristics are robust to increasing and random holding costs at different buffers.

Table 6: Average Performance of Heuristics for Increasing Holding Costs with Range $(0, 0.05)$ and Buffer Sizes with Varying Ranges

	$M = N = 3$, $0 \leq B_2, B_3 \leq 10$		$M = N = 4$, $0 \leq B_2, B_3, B_4 \leq 5$		$M = N = 5$, $0 \leq B_2, B_3, B_4, B_5 \leq 2$	
Policy	performance	% of the optimal	performance	% of the optimal	performance	% of the optimal
Teamwork	1.441 ± 0.004	% 84.247	1.440 ± 0.003	% 81.651	1.444 ± 0.002	% 80.642
Heuristic 1	1.644 ± 0.005	% 96.059	1.650 ± 0.005	% 93.553	1.659 ± 0.003	% 92.606
Heuristic 2	1.655 ± 0.005	% 96.699	1.668 ± 0.004	% 94.559	1.663 ± 0.003	% 92.880
Heuristic 3	1.612 ± 0.005	% 94.230	1.610 ± 0.004	% 91.274	1.615 ± 0.003	% 90.191
Heuristic 4	1.621 ± 0.005	% 94.717	1.676 ± 0.004	% 95.003	1.718 ± 0.003	% 95.939
optimal	1.711 ± 0.005	—	1.764 ± 0.004	—	1.791 ± 0.003	—

Table 7: Average Performance of Heuristics for Independent Holding Costs with Range $(0, 0.05)$ and Buffer Sizes with Varying Ranges

	$M = N = 3$, $0 \leq B_2, B_3 \leq 10$		$M = N = 4$, $0 \leq B_2, B_3, B_4 \leq 5$		$M = N = 5$, $0 \leq B_2, B_3, B_4, B_5 \leq 2$	
Policy	performance	% of the optimal	performance	% of the optimal	performance	% of the optimal
Teamwork	1.442 ± 0.004	% 84.247	1.443 ± 0.003	% 81.166	1.451 ± 0.002	% 79.817
Heuristic 1	1.643 ± 0.005	% 96.022	1.674 ± 0.004	% 94.148	1.691 ± 0.003	% 93.060
Heuristic 2	1.654 ± 0.005	% 96.675	1.686 ± 0.004	% 94.789	1.689 ± 0.003	% 92.933
Heuristic 3	1.613 ± 0.005	% 94.300	1.637 ± 0.004	% 92.034	1.654 ± 0.003	% 91.004
Heuristic 4	1.621 ± 0.005	% 94.747	1.698 ± 0.004	% 95.491	1.750 ± 0.003	% 96.269
optimal	1.711 ± 0.005	—	1.778 ± 0.004	—	1.817 ± 0.003	—

4.2.4 Insights and Limitations

Our numerical experiments imply that the threshold policies perform well even for larger systems. However, it is better to consider the holding costs, servers' rates, and number of jobs at each station (as opposed to two consecutive stations) when determining the thresholds. Hence, it is not practically possible to calculate each threshold exactly, however simple server assignment heuristics that use an effective primary assignment together with an effective contingency plan still have near-optimal performance. Moreover, we observe that our heuristics provide significant improvements over the teamwork policy. For example, when $M = N = 3$ and a constant holding cost of 0.05 is used, Heuristic 2 closes the optimality gap by more than 85% compared to the teamwork policy. Moreover, we observe that our heuristics perform better for systems with smaller buffer sizes, and the effects of increasing and random holding costs at different buffers are negligible. Although our heuristics reached near-optimal profit, we believe that it may be possible to improve some of our heuristics even further. For example, Heuristic 2 can be changed to involve the value of the holding costs in the contingency plan, and the condition in Heuristic 4 that determines when the servers work as a team perhaps can be improved. Finally, note that our heuristics have been devised for systems with $M = N$, but they can be easily adapted to systems with $M > N$. However, we believe that understaffed systems with $M < N$ necessitate different types of server assignment heuristics that include zones instead of primary assignments.

5 Conclusion

In this work we have studied the dynamic assignment of servers in tandem lines with the objective of maximizing the profit. More specifically, we assumed that a positive revenue is earned every time a job leaves the system and a holding cost is incurred for the jobs that have completed the service at the first station. We have characterized the optimal server assignment policy for systems with either arbitrary number of generalist servers and arbitrary number of stations or two specialist servers and two stations.

We have observed that the optimal policy is of "threshold" type where each server moves between the stations according to the number of jobs waiting for service or being served at each station. In particular, we have shown that the policy where all the servers work together as a team is optimal (note that this is a threshold policy as well) for systems with generalist servers. For systems with two specialist servers and two stations, we have shown that as the value of the holding cost increases, the value of the threshold decreases (we have also determined this threshold). Furthermore, we have observed the teamwork policy is optimal for high values of the holding cost, and the optimal policy of the throughput maximization problem is optimal for small values of the holding cost.

Moreover, we have studied systems with nonlinear holding cost structures. We have observed that the teamwork policy is still optimal for systems with generalist servers, and we have provided numerical results that support the conjecture that the optimal server assignment policy is of threshold type for systems with two specialist servers and two stations.

Finally, we have provided numerical results that compared the performances of various server assignment heuristics with the optimal server assignment policy in larger systems. More specifically, for systems with specialist servers we have observed that the optimal server assignment policy of the systems with two specialist servers and two stations can not simply be generalized to larger systems. Moreover, the teamwork policy (that is optimal for generalist systems) have performed well even for systems with specialist servers, however its performance have deteriorated as the system got larger.

This paper provides a starting point for future research in dynamic server assignment problem with conflicting objectives. We plan to study the quality and efficiency tradeoff for systems with different configurations, collaboration policies, and service time distributions. We believe that this will provide significant improvements to earlier results for problems with single objectives, because multiple conflicting objectives are more common in real production and service systems and their analysis will provide interesting insights to the practitioners as well as academicians.

Appendix

Proof of Lemma 3.1: If all the servers are idle at a state s, then s is an absorbing state and the long-run average profit is zero. Hence, at least one of the servers should not be idle at a state s. Let π be any policy that idles one server when the other server is assigned to a station j in state s. Now compare π with the policy π' that assigns both servers to station j in state s and agrees with π otherwise. The transition time from state s to another state s' is never longer for π' compared to π. Hence, the number of departures is never smaller under π'. Moreover, the holding cost per item produced is never bigger under π' because the time required to reach the same number of departures (going through the same states) is never longer under π'. Hence, $\frac{D_\pi(t)}{t} \geq \frac{D'_\pi(t)}{t}$ and $C_\pi(t) \leq C_{\pi'}(t)$ for all $t \geq 0$. Equation (5) shows that $P_\pi \geq P_{\pi'}$, and consequently there exists an optimal policy that does not idle any servers in state s. □

Proof of Lemma 3.2: When $\mu_{11} \neq \mu_{22}$, some algebra shows that

$$\begin{aligned}
&\frac{f_1(i,i+1)}{f_2(i,i+1)} > \frac{f_1(i+1,i+2)}{f_2(i+1,i+2)} \\
&\Leftrightarrow \mu_{11}(\frac{\mu_{11}}{\mu_{22}})^{i-1}(\mu_{11}+\mu_{12})(\mu_{22}-\mu_{11}) \\
&\times\Big(\mu_{12}\mu_{21}(\mu_{22}^i-\mu_{11}^i)+(\mu_{12}+\mu_{21})(\mu_{22}^{i+1}-\mu_{11}^{i+1})+(\mu_{22}^{i+2}-\mu_{11}^{i+2})\Big) > 0.
\end{aligned}$$

The last inequality is correct because $\mu_{11} \neq \mu_{22}$, hence the lemma holds. When $\mu_{11} = \mu_{22}$, we can show that the lemma holds because

$$\frac{f_1(i,i+1)}{f_2(i,i+1)} > \frac{f_1(i+1,i+2)}{f_2(i+1,i+2)} \Leftrightarrow 2(\mu_{12}+\mu_{22})\Big(\mu_{12}\mu_{21}+(\mu_{12}+\mu_{21})\mu_{22}+\mu_{22}^2\Big) > 0.$$

Hence the proof is complete. □

Proof of Lemma 3.3: Let $i, j \in \mathbb{N}^+$ and $i < j$. Since $f_1(i,j) \geq 0$, we only need to show that $f_2(i,j) \geq 0$. Some algebra shows that

$$\begin{aligned}
f_2(i,j) \quad = \quad & \frac{\mu_{22}^{j+i-3}}{\mu_{11}^{i-1}}(\mu_{22}-\mu_{11})^2\Big[\mu_{12}\mu_{22}^2\Big(\sum_{k=1}^{j-1}k(\frac{\mu_{11}}{\mu_{22}})^{k-1}+\sum_{k=1}^{j-1}k(\frac{\mu_{11}}{\mu_{22}})^{k}\Big) \\
&+\mu_{12}\mu_{22}(\mu_{12}+\mu_{21})\sum_{k=1}^{j-1}k(\frac{\mu_{11}}{\mu_{22}})^{k-1}+(\mu_{11}\mu_{22}-\mu_{12}\mu_{21})\mu_{22}\sum_{k=i}^{j-1}(\frac{\mu_{11}}{\mu_{22}})^{k-1} \\
&+(\mu_{11}+\mu_{12})(\mu_{11}+\mu_{21})\Big(\mu_{12}\sum_{k=i}^{j-1}\sum_{l=1}^{i-1}(k-l)(\frac{\mu_{11}}{\mu_{22}})^{k+l-2}+\mu_{22}\sum_{k=i+1}^{j}\sum_{l=1}^{i}(k-l)(\frac{\mu_{11}}{\mu_{22}})^{k+l-3}\Big)\Big].
\end{aligned}$$

Hence, it is clear that $f_2(i,j) > 0$ because $\mu_{11} \neq \mu_{22}$ and $\mu_{11}\mu_{22} \geq \mu_{12}\mu_{21}$. Finally, when $\mu_{11} = \mu_{22}$, it is clear that $f_2(i,j) > 0$ because $i < j$. Hence the proof is complete. □

Proof of Lemma 3.4: It is sufficient to show that

$$\frac{f_1(i,j)}{f_2(i,j)} \geq \frac{f_1(i+1,j)}{f_2(i+1,j)}. \tag{7}$$

First assume that $\mu_{22} \neq \mu_{11}$. Then, some algebra shows that the inequality (7) holds if and only if

$$\mu_{22}^{j-i-1}(\mu_{22}-\mu_{11})\Big(\mu_{22}^{j+1}-\mu_{11}^{j+1}+(\mu_{22}^{j}-\mu_{11}^{j})(\mu_{12}+\mu_{21})+(\mu_{22}^{j-1}-\mu_{11}^{j-1})\mu_{12}\mu_{21}\Big)$$
$$\times\Big(\mu_{12}\sum_{k=i+1}^{j-1}(k-i)(\frac{\mu_{11}}{\mu_{22}})^{k-1}+\mu_{22}\sum_{k=i+2}^{j}(k-i-1)(\frac{\mu_{11}}{\mu_{22}})^{k-1}\Big)\geq 0. \quad (8)$$

Note that the inequality (8) holds because $j > i$. When $\mu_{11} = \mu_{22}$ we observe that the inequality (7) holds if and only if

$$\frac{\mu_{12}+\mu_{22}}{2}(j-i)(j-i-1)\geq 0.$$

This final inequality holds because $j > i$ and $i \geq 1$. Hence the proof is complete. □

Proof of Theorem 3.2: Lemma 3.1 shows that it suffices to consider the policies that are non-idling. Then, in each state $i \in \mathcal{S}$, it is sufficient to consider the following set of allowable actions:

$$A_i = \begin{cases} a_{11} & \text{for } i = 0, \\ \{a_{11}, a_{12}, a_{21}, a_{22}\} & \text{for } i \in \{1,\ldots,B+1\}, \\ a_{22} & \text{for } i = B+2. \end{cases}$$

Furthermore, for all $\pi \in \Pi$, let q be the uniformization constant(see, Lippman [28]). The policy described in the theorem corresponds to a Markov chain with a single recurrent class and possibly some transient states. Thus, we can use the policy iteration algorithm for weakly communicating models as described in Section 9.5.1 of Puterman [33].

Let P_d be the probability transition matrix corresponding to the policy π, and $r_d(i)$ denote the reward in state i when policy π is employed.

Let us define $\frac{f_1(0,1)}{f_2(0,1)} = \infty$ and $\frac{f_1(B+2,B+3)}{f_2(B+2,B+3)} = 0$, and assume that $\frac{f_1(s,s+1)}{f_2(s,s+1)} < h \leq \frac{f_1(s-1,s)}{f_2(s-1,s)}$ (Lemma 3.2 shows that this interval is nonempty). We start the policy iteration algorithm by considering the policy $\pi_0 = (d_0)^\infty$, where

$$d_0(i) = \begin{cases} a_{11} & \text{for } i = 0, \\ a_{12} & \text{for } 1 \leq i \leq s-1, \\ a_{22} & \text{for } s \leq i \leq B+2. \end{cases}$$

Then we obtain

$$r_{d_0}(i) = \begin{cases} 0 & \text{for } i = 0, \\ \mu_{22} - ih & \text{for } 1 \leq i \leq s-1, \\ \mu_{12}+\mu_{22} - ih & \text{for } s \leq i \leq B+2, \end{cases}$$

and

$$P_{d_0}(i,i') = \begin{cases} \frac{\mu_{11}+\mu_{21}}{q} & \text{for } i=0 \text{ and } i'=1, \\ \frac{\mu_{12}+\mu_{22}}{q} & \text{for } i=i'=0, \\ \frac{\mu_{11}}{q} & \text{for } i \in \{1,\ldots,s-1\} \text{ and } i'=i+1, \\ \frac{\mu_{22}}{q} & \text{for } i \in \{1,\ldots,s-1\} \text{ and } i'=i-1, \\ \frac{\mu_{12}+\mu_{21}}{q} & \text{for } i=i' \in \{1,\ldots,s-1\}, \\ \frac{\mu_{12}+\mu_{22}}{q} & \text{for } i \in \{s,\ldots,B+2\} \text{ and } i'=i-1, \\ \frac{\mu_{11}+\mu_{21}}{q} & \text{for } i=i' \in \{s,\ldots,B+2\}. \end{cases}$$

For all $i, i' \in \mathcal{S}$ and $a \in A_i$, we use $r(i,a)$ to denote the immediate reward in state i when action a is taken and $p(i'|i,a)$ to denote the one-step probability of going from state i to state i' when action a is chosen in state i. Under our assumptions on the service rates ($\mu_{11}\mu_{22} \geq \mu_{12}\mu_{21}$, $\sum_{k=1}^{M} \mu_{kj} > 0$ for $j \in \{1,\ldots,N\}$, and $\sum_{j=1}^{N} \mu_{kj} > 0$ for $k \in \{1,\ldots,M\}$), it is clear that $\mu_{11} > 0$ and $\mu_{22} > 0$. Hence, the discrete time Markov chain $\{X'_{\pi_0}(t)\}$ found by uniformization of $\{X_{\pi_0}(t)\}$ is unichain, and we can solve the following set of equations to find a scalar g_0 and a vector h_0, letting $h_0(0,0) = 0$,

$$r_{d_0} - g_0 e + (P_{d_0} - I)h_0 = 0, \tag{9}$$

where e is the column vector of ones and I is the identity matrix. We can show that

$$g_0 = \begin{cases} (\mu_{11}+\mu_{21})(\mu_{12}+\mu_{22})\Big((\mu_{22}^s - \mu_{11}^s) - \frac{h}{(\mu_{22}-\mu_{11})(\mu_{12}+\mu_{22})} \\ \times\Big[(\mu_{12}+\mu_{22})\Big(\mu_{22}^s - s\mu_{11}^{s-1}\mu_{22} + (s-1)\mu_{11}^s\Big) + s\mu_{11}^{s-1}(\mu_{22}-\mu_{11})^2\Big]\Big) \\ \Big/\Big(\mu_{22}^{s+1} - \mu_{11}^{s+1} + (\mu_{22}^s - \mu_{11}^s)(\mu_{12}+\mu_{21}) + (\mu_{22}^{s-1} - \mu_{11}^{s-1})\mu_{12}\mu_{21}\Big) & \text{if } \mu_{11} \neq \mu_{22}; \\ (\mu_{11}+\mu_{21})(\mu_{11}+\mu_{22})\Big(s\mu_{22} - \frac{h}{2(\mu_{12}+\mu_{22})}\Big[s(s-1)(\mu_{12}+\mu_{22}) + 2s\mu_{22}\Big]\Big) \\ \Big/\Big(\mu_{22}(\mu_{11}+\mu_{21}) + (s-1)(\mu_{11}+\mu_{21})(\mu_{12}+\mu_{22}) + \mu_{22}(\mu_{12}+\mu_{22})\Big) & \text{if } \mu_{11} = \mu_{22}; \end{cases}$$

and $h_0(0) = 0$,

$$h_0(i) = \begin{cases} g_0 q \frac{(\mu_{11}+\mu_{21})\sum_{j=0}^{i-2}(j+1)\mu_{11}^j\mu_{22}^{i-2-j} + \sum_{j=0}^{i-1}\mu_{11}^j\mu_{22}^{i-1-j}}{\mu_{11}^{i-1}(\mu_{11}+\mu_{21})} - q\mu_{22}\frac{\sum_{j=0}^{i-2}(j+1)\mu_{11}^j\mu_{22}^{i-2-j}}{\mu_{11}^{i-1}} \\ +hq(\mu_{11}+\mu_{22})\frac{\sum_{j=0}^{s-2}\mu_{11}^j\mu_{22}^{i-2-j}\sum_{k=0}^{j+1}k}{\mu_{11}^{i-1}} \text{ for } 1 \leq i \leq s; \\ h_0(s) + q(i-s) - qg_0\frac{(i-s)}{\mu_{12}+\mu_{22}} - hq\sum_{k=s+1}^{i}k \text{ for } s+1 \leq i \leq B+1; \end{cases}$$

is a solution to (9), with the convention that summation over an empty set is equal to zero.

Next, we compute $d_1(i)$, where

$$d_1(i) \in \arg\max_{a \in A_i}\Big\{r(i,a) + \sum_{i' \in \mathcal{S}} p(i'|i,a)h_0(i')\Big\}, \quad \forall i \in \mathcal{S},$$

and set $d_1(i) = d_0(i)$ whenever possible. If one can show $d_1(i) = d_0(i)$ for all $i \in S$, then the policy $\pi_0 = (d_0)^\infty$ is optimal according to Theorem 9.5.1 of Puterman [33]. Consequently, for all $i \in \mathcal{S}$ and $a \in A_i$, we want to show that the following inequality holds:

$$\Delta(i,a) = \Big(r(i,d_0(i)) + \sum_{i'\in\mathcal{S}} p(i'|i,d_0(i))h_0(i')\Big) - r(i,a) - \sum_{i'\in\mathcal{S}} p(i'|i,a)h_0(i') \geq 0. \tag{10}$$

For the remainder of the proof, we assume that $\mu_{11} \neq \mu_{22}$ since the proof for the case with $\mu_{11} = \mu_{22}$ is similar. It is clear that the action $d_0(i)$ is optimal for $i \in \{0, B+2\}$ because the action space consists of only one action in these states. Hence, it is sufficient to check if the inequality (10) is satisfied in states $1 \leq i \leq B+1$. For $i, l \in \mathbb{N}^+$ and $i < l$, let us define $f_j(i,l)$ for $j \in \{3,4,5,6\}$ as follows:

$$\begin{aligned}
f_3(i,l) &= \mu_{11}^{l-i-1}(\mu_{11}+\mu_{21})(\mu_{22}^i - \mu_{11}^i)(\mu_{22}-\mu_{11})(\mu_{11}\mu_{22}-\mu_{12}\mu_{21}),\\
f_4(i,l) &= \frac{\mu_{22}^{i-1}}{\mu_{11}^i}(\mu_{22}-\mu_{11})\Big[\Big(\mu_{11}\sum_{k=1}^{i-1} k(\frac{\mu_{11}}{\mu_{22}})^{k-1} + \mu_{21}\sum_{k=1}^{i} k(\frac{\mu_{11}}{\mu_{22}})^{k-1}\Big)\\
&\quad \times\Big(\mu_{12}\mu_{21}(\mu_{22}^{l-1}-\mu_{11}^{l-1}) + (\mu_{12}+\mu_{21})(\mu_{22}^l - \mu_{11}^l) + \mu_{22}^{l+1} - \mu_{11}^{l+1}\Big)\\
&\quad -\mu_{22}^{l-i-1}\Big(\mu_{12}\sum_{k=1}^{l-1} k(\frac{\mu_{11}}{\mu_{22}})^{k-1} + \mu_{22}\sum_{k=1}^{l} k(\frac{\mu_{11}}{\mu_{22}})^{k-1}\Big)(\mu_{11}+\mu_{21})(\mu_{21}+\mu_{22})(\mu_{22}^i-\mu_{11}^i)\Big],\\
f_5(i,l) &= (\mu_{22}-\mu_{11})(\mu_{11}\mu_{22}-\mu_{12}\mu_{21})\Big(\mu_{22}^l - \mu_{11}^l + \mu_{12}\mu_{22}^{i-1}(\mu_{22}^{l-i}-\mu_{11}^{l-i}) + \mu_{11}^{l-i-1}\mu_{12}(\mu_{22}^i - \mu_{11}^i)\Big),\\
f_6(i,l) &= \frac{\mu_{22}^{i-2}}{\mu_{11}^i}(\mu_{22}-\mu_{11})\Big[\mu_{22}^{l-i}\Big(\mu_{12}\sum_{k=1}^{j-1} k(\frac{\mu_{11}}{\mu_{22}})^{k-1} + \mu_{22}\sum_{k=1}^{j} k(\frac{\mu_{11}}{\mu_{22}})^{k-1}\Big)\\
&\quad \times\Big(\mu_{11}^i(\mu_{11}+\mu_{21})(\mu_{21}+\mu_{22}-\mu_{11}-\mu_{12}) + \mu_{22}^{i-1}(\mu_{21}+\mu_{22})(\mu_{11}\mu_{12}-\mu_{21}\mu_{22})\Big)\\
&\quad -\Big(\mu_{11}\mu_{12}\sum_{k=1}^{i-1} k(\frac{\mu_{11}}{\mu_{22}})^{k-1} - \mu_{21}\mu_{22}\sum_{k=1}^{i} k(\frac{\mu_{11}}{\mu_{22}})^{k-1} + i\frac{\mu_{11}^i}{\mu_{22}^{i-2}}\Big)\\
&\quad \times\Big(\mu_{12}\mu_{21}(\mu_{22}^{l-1}-\mu_{11}^{l-1}) + (\mu_{12}+\mu_{21})(\mu_{22}^l-\mu_{11}^l) + \mu_{22}^{j+1} - \mu_{11}^{l+1}\Big)\Big].
\end{aligned}$$

First, let us consider the states $i \in \{1,\ldots,s-1\}$. Recall that $d_0(i) = a_{12}$ for $i \in \{1,\ldots,s-1\}$. With some algebra we have

$$\Delta(i,a_{11}) = \frac{f_3(i,s) - hf_4(i,s)}{(\mu_{22}-\mu_{11})\Big(\mu_{22}^{s+1} - \mu_{11}^{s+1} + (\mu_{12}+\mu_{21})(\mu_{22}^s - \mu_{11}^s) + \mu_{12}\mu_{21}(\mu_{22}^{s-1}-\mu_{11}^{s-1})\Big)}.$$

The denominator of this expression is always positive. We observe that $f_3(i,s) \geq 0$ for $i \in \{1,\ldots,s-1\}$. If $f_4(i,s) \leq 0$, then $\Delta(i,a_{11}) \geq 0$. If $f_4(i,s) > 0$, we will prove that $\Delta(i,a_{11}) \geq 0$ by showing that $h \leq \frac{f_3(i,s)}{f_4(i,s)}$ (note that we do not have to show this when $f_4(i,s) \leq 0$). Extensive algebra together with Lemma 3.3 shows that for $i \in \{1,\ldots,s-1\}$, we have

$$\frac{f_1(i,s)}{f_2(i,s)} \leq \frac{f_3(i,s)}{f_4(i,s)} \quad\Leftrightarrow\quad (\mu_{22}-\mu_{11})\Big(\mu_{22}^{s+1} - \mu_{11}^{s+1} + (\mu_{12}+\mu_{21})(\mu_{22}^s-\mu_{11}^s) + \mu_{12}\mu_{21}(\mu_{22}^{s-1}-\mu_{11}^{s-1})\Big)$$

$$\times\Big[(\mu_{11}+\mu_{21})(\mu_{12}+\mu_{22})\sum_{k=i}^{s-1}\sum_{j=1}^{i-1}(k-j)(\frac{\mu_{11}}{\mu_{22}})^{k+j-1}$$

$$+(s-i)(\mu_{11}+\mu_{21})\mu_{11}^{s-1}\mu_{22}\sum_{k=i}^{s-1}(\frac{\mu_{11}}{\mu_{22}})^{k}$$

$$+(\mu_{12}+\mu_{22})\mu_{22}^{s-1}\Big(\mu_{11}\sum_{k=i}^{s-1}k(\frac{\mu_{11}}{\mu_{22}})^{k-1}+\mu_{21}\sum_{k=i}^{s-1}(k-i)(\frac{\mu_{11}}{\mu_{22}})^{k-1}\Big)\Big]\geq 0. \quad (11)$$

We observe that the inequality (11) holds trivially. Our assumption that $h \leq \frac{f_1(s-1,s)}{f_2(s-1,s)}$ together with Lemma 3.4 implies that $h \leq \frac{f_1(i,s)}{f_2(i,s)}$ for $i \in \{1,\ldots,s-1\}$. Consequently, the inequality (11) shows that $h \leq \frac{f_3(i,s)}{f_4(i,s)}$ and hence $\Delta(i,a_{11}) \geq 0$ for $i \in \{1,\ldots,s-1\}$.

Similarly, we can show that

$$\Delta(i,a_{21}) = \frac{f_5(i,s)-hf_6(i,s)}{(\mu_{22}-\mu_{11})\Big(\mu_{22}^{s+1}-\mu_{11}^{s+1}+(\mu_{12}+\mu_{21})(\mu_{22}^{s}-\mu_{11}^{s})+\mu_{12}\mu_{21}(\mu_{22}^{s-1}-\mu_{11}^{s-1})\Big)}.$$

The denominator of this expression is always positive. We see that $f_5(i,s) \geq 0$ for $i \in \{1,\ldots,s-1\}$. If $f_6(i,s) \leq 0$, then $\Delta(i,a_{21}) \geq 0$. If $f_6(i,s) > 0$, we will prove that $\Delta(i,a_{21}) \geq 0$ by showing that $h \leq \frac{f_5(i,s)}{f_6(i,s)}$ (note that we do not have to show this when $f_6(i,s) \leq 0$). Extensive algebra together with Lemma 3.3 shows that for $i \in \{1,\ldots,s-1\}$, we have

$$\frac{f_1(i,s)}{f_2(i,s)} \leq \frac{f_5(i,s)}{f_6(i,s)} \quad\Leftrightarrow\quad \frac{(\mu_{22}-\mu_{11})}{\mu_{11}}\Big(\mu_{22}^{s+1}-\mu_{11}^{s+1}+(\mu_{12}+\mu_{21})(\mu_{22}^{s}-\mu_{11}^{s})+\mu_{12}\mu_{21}(\mu_{22}^{s-1}-\mu_{11}^{s-1})\Big)$$

$$\times\Big[(\mu_{11}+\mu_{21})(\mu_{12}+\mu_{22})\sum_{k=i}^{s-1}\sum_{j=1}^{i-1}(k-j)(\frac{\mu_{11}}{\mu_{22}})^{k+j-1}$$

$$+(s-i)(\mu_{11}+\mu_{21})\mu_{11}^{s-1}\mu_{22}(\mu_{22}^{i}-\mu_{11}^{i})\sum_{k=0}^{i-1}(\frac{\mu_{11}}{\mu_{22}})^{k}$$

$$+(\mu_{12}+\mu_{22})\mu_{22}^{s-1}\Big(\mu_{11}\sum_{k=i}^{s-1}k(\frac{\mu_{11}}{\mu_{22}})^{k-1}+\mu_{21}\sum_{k=i}^{s-1}(k-i)(\frac{\mu_{11}}{\mu_{22}})^{k-1}\Big)\Big]\geq 0. \quad (12)$$

We observe that the inequality (12) holds trivially. Our assumption that $h \leq \frac{f_1(s-1,s)}{f_2(s-1,s)}$ together with Lemma 3.4 implies that $h \leq \frac{f_1(i,s)}{f_2(i,s)}$ for $i \in \{1,\ldots,s-1\}$. Consequently, the inequality (12) shows that $h \leq \frac{f_5(i,s)}{f_6(i,s)}$ and hence $\Delta(i,a_{21}) \geq 0$ for $i \in \{1,\ldots,s-1\}$.

Finally, some algebra shows that

$$\Delta(i,a_{22}) = \frac{f_1(i,s)-hf_2(i,s)}{(\mu_{22}-\mu_{11})\Big(\mu_{22}^{s+1}-\mu_{11}^{s+1}+(\mu_{12}+\mu_{21})(\mu_{22}^{s}-\mu_{11}^{s})+\mu_{12}\mu_{21}(\mu_{22}^{s-1}-\mu_{11}^{s-1})\Big)}.$$

The denominator of this expression is always positive, and our assumption that $h \leq \frac{f_1(s-1,s)}{f_2(s-1,s)}$ together with Lemma 3.4 shows that $\Delta(i,a_{22}) \geq 0$ for $i \in \{1,\ldots,s-1\}$.

For $l \in \mathbb{N}^+$ and $l \geq s$, let us define $f'_j(l)$ for $j \in \{1, 2, 3, 4, 5\}$ as follows:

$$\begin{aligned}
f'_1(l) &= \frac{\mu_{22}^{s-2}}{\mu_{11}^{s-1}}(\mu_{22}-\mu_{11})\Big[\Big(l\mu_{12}+(l+1)\mu_{11}\Big) \\
&\quad\times\Big(\mu_{12}\mu_{21}(\mu_{22}^{s-1}-\mu_{11}^{s-1})+(\mu_{12}+\mu_{21})(\mu_{22}^{s}-\mu_{11}^{s})+\mu_{22}^{s+1}-\mu_{11}^{s+1}\Big) \\
&\quad-\mu_{22}^{s-2}(\mu_{22}-\mu_{11})(\mu_{11}+\mu_{12})(\mu_{11}+\mu_{21})\Big(\mu_{12}\sum_{k=1}^{s-1}k(\frac{\mu_{11}}{\mu_{22}})^{k-1}+\mu_{22}\sum_{k=1}^{s}k(\frac{\mu_{11}}{\mu_{22}})^{k-1}\Big)\Big], \\
f'_2(l) &= (\mu_{11}+\mu_{21})(\mu_{11}\mu_{22}-\mu_{12}\mu_{21})(\mu_{22}-\mu_{11})(\mu_{22}^{l-1}-\mu_{11}^{l-1}), \\
f'_3(l) &= (\mu_{22}-\mu_{11})\Big[\Big(l(\mu_{12}+\mu_{22})+(l+1)(\mu_{11}+\mu_{21})\Big)\times\Big(\mu_{12}\mu_{21}(\mu_{22}^{s-1}-\mu_{11}^{s-1}) \\
&\quad+(\mu_{12}+\mu_{21})(\mu_{22}^{s}-\mu_{11}^{s})+\mu_{22}^{s+1}-\mu_{11}^{s+1}\Big)-\mu_{22}^{s-2}(\mu_{22}-\mu_{11})(\mu_{11}+\mu_{21})(\mu_{11}+\mu_{12}+\mu_{21}+\mu_{22}) \\
&\quad\times\Big(\mu_{12}\sum_{k=1}^{s-1}k(\frac{\mu_{11}}{\mu_{22}})^{k-1}+\mu_{22}\sum_{k=1}^{s}k(\frac{\mu_{11}}{\mu_{22}})^{k-1}\Big)\Big], \\
f'_4(l) &= (\mu_{11}\mu_{22}-\mu_{12}\mu_{21})(\mu_{22}-\mu_{11})\Big((\mu_{22}^{l}-\mu_{11}^{l})+\mu_{21}(\mu_{22}^{l-1}-\mu_{11}^{l-1})\Big), \\
f'_5(l) &= (\mu_{22}-\mu_{11})\Big[\Big(l\mu_{22}+(l+1)\mu_{21}\Big) \\
&\quad\times\Big(\mu_{12}\mu_{21}(\mu_{22}^{s-1}-\mu_{11}^{s-1})+(\mu_{12}+\mu_{21})(\mu_{22}^{s}-\mu_{11}^{s})+\mu_{22}^{s+1}-\mu_{11}^{s+1}\Big) \\
&\quad-\mu_{22}^{s-2}(\mu_{22}-\mu_{11})(\mu_{11}+\mu_{21})(\mu_{21}+\mu_{22})\Big(\mu_{12}\sum_{k=1}^{s-1}k(\frac{\mu_{11}}{\mu_{22}})^{k-1}+\mu_{22}\sum_{k=1}^{s}k(\frac{\mu_{11}}{\mu_{22}})^{k-1}\Big)\Big].
\end{aligned}$$

Now, consider the states $i \in \{s, \ldots, B+1\}$, whenever this set is nonempty. Recall that $d_0(i) = a_{22}$ for $i \in \{s, \ldots, B+1\}$. With some algebra we have

$$\Delta(i, a_{11}) = \frac{f'_2(s) + hf'_3(i)}{(\mu_{22}-\mu_{11})(\mu_{12}+\mu_{22})\Big(\mu_{22}^{s+1}-\mu_{11}^{s+1}+(\mu_{12}+\mu_{21})(\mu_{22}^{s}-\mu_{11}^{s})+\mu_{12}\mu_{21}(\mu_{22}^{s-1}-\mu_{11}^{s-1})\Big)}.$$

The denominator of this expression is nonnegative. We observe that $f'_2(s) \geq 0$. Note that $f'_3(i) \geq 0$ is nondecreasing in i, hence we only need to show that $\Delta(s, a_{11}) = f'_2(s)+hf'_3(s) \geq 0$ (which consequently implies that $f'_2(s) + hf'_3(i) \geq 0$ for $i \in \{s, \ldots, B+1\}$). Some algebra shows that

$$\begin{aligned}
f'_3(s) &= \mu_{22}^{s-2}(\mu_{22}-\mu_{11})^2\Big[\mu_{22}(\mu_{12}+\mu_{22})\Big((\mu_{11}+\mu_{21})+s(\mu_{11}+\mu_{21})(\mu_{21}+\mu_{22})\Big) \\
&\quad+(\mu_{11}+\mu_{21})\Big((\mu_{12}+\mu_{22})^2\sum_{k=1}^{s-1}(s-k)(\frac{\mu_{11}}{\mu_{22}})^{k-1}+\mu_{12}(\mu_{11}+\mu_{21})\sum_{k=1}^{s-1}(s+1-k)(\frac{\mu_{11}}{\mu_{22}})^{k-1} \\
&\quad+\mu_{22}(\mu_{11}+\mu_{21})\sum_{k=1}^{s}(s+1-k)(\frac{\mu_{11}}{\mu_{22}})^{k-1}\Big)\Big].
\end{aligned}$$

It is clear that $f'_3(s) \geq 0$. Hence we can conclude that $\Delta(s, a_{11}) \geq 0$ and consequently $\Delta(i, a_{11}) \geq 0$ for $i \in \{s, \ldots, B+1\}$.

Similarly, we can show that

$$\Delta(i,a_{12}) = \frac{-f_1(s,s+1)+hf_1'(i)}{(\mu_{22}-\mu_{11})(\mu_{12}+\mu_{22})\Big(\mu_{22}^{s+1}-\mu_{11}^{s+1}+(\mu_{12}+\mu_{21})(\mu_{22}^{s}-\mu_{11}^{s})+\mu_{12}\mu_{21}(\mu_{22}^{s-1}-\mu_{11}^{s-1})\Big)}.$$

Note that the denominator of this expression is always positive. We observe that $f_1'(s) = f_2(s,s+1)$. Moreover, $f_1'(i)$ is nonnegative (by Lemma 3.3) and nondecreasing in i. Hence, it is sufficient to show that $-f_1(s,s+1)+hf_1'(s) = -f_1(s,s+1)+hf_2(s,s+1) \geq 0$ (which consequently implies that $-f_1(s,s+1)+hf_1'(i) \geq 0$ for $i \in \{s,\ldots,B+1\}$). This follows since $h > \frac{f_1(s,s+1)}{f_2(s,s+1)}$, and thus $\Delta(i,a_{22}) \geq 0$ for $i \in \{s,\ldots,B+1\}$.

Finally, some algebra shows that

$$\Delta(i,a_{21}) = \frac{f_4'(s)+hf_5'(i)}{(\mu_{22}-\mu_{11})(\mu_{12}+\mu_{22})\Big(\mu_{22}^{s+1}-\mu_{11}^{s+1}+(\mu_{12}+\mu_{21})(\mu_{22}^{s}-\mu_{11}^{s})+\mu_{12}\mu_{21}(\mu_{22}^{s-1}-\mu_{11}^{s-1})\Big)}.$$

The denominator of this expression is nonnegative. We observe that $f_4'(s) \geq 0$. Note that $f_5'(i) \geq 0$ is nondecreasing in i, hence we only need to show that $\Delta(s,a_{21}) = f_4'(s)+hf_5'(s) \geq 0$ (which consequently implies that $f_4'(s)+hf_5'(i) \geq 0$ for $i \in \{s,\ldots,B+1\}$). Some algebra shows that

$$\begin{aligned} f_5'(s) \;=\; & (\mu_{22}-\mu_{11})^2\Big[\mu_{21}\Big(\mu_{11}^{s-1}(\mu_{11}+\mu_{21})+\mu_{22}^{s-1}(\mu_{12}+\mu_{22})+s\mu_{22}^{s-1}(\mu_{12}+\mu_{22})(\mu_{21}+\mu_{22})\Big) \\ & +\mu_{22}^{s-1}(\mu_{11}+\mu_{21})(\mu_{12}+\mu_{22})\Big(\mu_{21}\sum_{k=1}^{s-1}(s+1-k)(\frac{\mu_{11}}{\mu_{22}})^{k-1}+\mu_{22}\sum_{k=1}^{s-1}(s-k)(\frac{\mu_{11}}{\mu_{22}})^{k-1}\Big)\Big]. \end{aligned}$$

It is clear that $f_5'(s) \geq 0$. Hence we can conclude that $\Delta(s,a_{21}) \geq 0$ and consequently $\Delta(i,a_{21}) \geq 0$ for $i \in \{s,\ldots,B+1\}$.

Note that any action that takes the process to one of the recurrent states can be used in states $\{s+1,\ldots,B+2\}$ because they are transient. The uniqueness of the optimal actions in states $\{0,\ldots,B+2\}$ when $\mu_{11}\mu_{22} > \mu_{12}\mu_{21}$ and $h \neq \frac{f_1(s-1,s)}{f_2(s-1,s)}$ follows from the proof of Theorem 3.1 of Andradóttir and Ayhan [5] and the discussion in Section 9.5.2 of Puterman [33]. □

References

[1] H.-S. Ahn and R. Righter. Dynamic load balancing with flexible workers. *Advances in Applied Probability*, 38:621–642, 2006.

[2] H.-S. Ahn, I. Duenyas, and R. Zhang. Optimal scheduling of a 2-stage tandem queue with parallel servers. *Advances in Applied Probability*, 31:1095–1117, 1999.

[3] H.-S. Ahn, I. Duenyas, and M. E. Lewis. The optimal control of a two-stage tandem queueing system with flexible servers. *Probability in the Engineering and Informational Sciences*, 16:453–469, 2002.

[4] H.-S. Ahn, I. Duenyas, and R. Zhang. Optimal control of a flexible server. *Advances in Applied Probability*, 36:139–170, 2004.

[5] S. Andradóttir and H. Ayhan. Throughput maximization for tandem lines with two stations and flexible servers. *Operations Research*, 53:516–531, 2005.

[6] S. Andradóttir, H. Ayhan, and D. G. Down. Server assignment policies for maximizing the steady-state throughput of finite queueing systems. *Management Science*, 47:1421–1439, 2001.

[7] S. Andradóttir, H. Ayhan, and D. G. Down. Dynamic server allocation for queueing networks with flexible servers. *Operations Research*, 51:952–968, 2003.

[8] S. Andradóttir, H. Ayhan, and D. G. Down. Compensating for failures with flexible servers. *Operations Research*, 55:753–768, 2007.

[9] S. Andradóttir, H. Ayhan, and D. G. Down. Dynamic assignment of dedicated and flexible servers in tandem lines. *Probability in the Engineering and Informational Sciences*, 21:497–538, 2007.

[10] S. Andradóttir, H. Ayhan, and D. G. Down. Maximizing the throughput of tandem lines with failure-prone servers and finite buffers. *Probability in the Engineering and Informational Sciences*, 22:191–211, 2008.

[11] S. Andradóttir, H. Ayhan, and E. Kırkızlar. Flexible servers in tandem lines with setups. Submitted for publication, 2010.

[12] J. J. Bartholdi, III and D. D. Eisenstein. A production line that balances itself. *Operations Research*, 44:21–34, 1996.

[13] J. J. Bartholdi, III, L. A. Bunimovich, and D. D. Eisenstein. Dynamics of two- and three-worker "bucket brigade" production lines. *Operations Research*, 47:488–491, 1999.

[14] J. J. Bartholdi, III, D. D. Eisenstein, and R. D. Foley. Performance of bucket brigades when work is stochastic. *Operations Research*, 49:710–719, 2001.

[15] S. L. Bell and R. J. Williams. Dynamic scheduling of a system with two parallel servers in heavy traffic with complete resource pooling: Asymptotic optimality of a threshold policy. *Annals of Applied Probability*, 11:608–649, 2001.

[16] S. L. Bell and R. J. Williams. Dynamic scheduling of a parallel server system in heavy traffic with complete resource pooling: Asymptotic optimality of a threshold policy. *Electronic Journal of Probability*, 10:1044–1115, 2005.

[17] P. Billingsley. *Probability and Measure*. Wiley, New York, NY, 1995.

[18] T. M. Farrar. Optimal use of an extra server in a two station tandem queueing network. *IEEE Transactions on Automatic Control*, 38:1296–1299, 1993.

[19] E. S. Gel, W. J. Hopp, and M. P. Van Oyen. Factors affecting opportunity of worksharing as a dynamic line balancing mechanism. *IIE Transactions*, 34:847–863, 2002.

[20] J. M. Harrison and M. J. López. Heavy traffic resource pooling in parallel-server systems. *Queueing Systems*, 33:339–368, 1999.

[21] W. J. Hopp, E. Tekin, and M. P. Van Oyen. Benefits of skill chaining in serial production lines with cross-trained workers. *Management Science*, 50:83–98, 2004.

[22] S. M. R. Iravani, M. J. M. Posner, and J. M. Buzacott. A two-stage tandem queue attended by a moving server with holding and switching costs. *Queueing Systems*, 26: 203–228, 1997.

[23] D. L. Kaufman, H.-S. Ahn, and M. E. Lewis. On the introduction of an agile, temporary workforce into a tandem queueing system. *Queueing Systems*, 51:135–171, 2005.

[24] E. Kırkızlar, S. Andradóttir, and H. Ayhan. Robustness of efficient server assignment policies to service time distributions in finite-buffered lines. *Naval Research Logistics*, 57:563–582, 2010.

[25] E. Kırkızlar, S. Andradóttir, and H. Ayhan. Flexible servers in understaffed tandem lines. *Production and Operations Management*, under revision, 2010.

[26] V. G. Kulkarni. *Modeling and Analysis of Stochastic Systems*. Chapman & Hall, London, UK, 1995.

[27] Y. F. Lim and K. K. Yang. Maximizing throughput of bucket brigades on discrete work stations. *Production and Operations Management*, 18:48–59, 2009.

[28] S. A. Lippman. Applying a new device in the optimization of exponential queueing system. *Operations Research*, 23:687–710, 1975.

[29] A. Mandelbaum and A. L. Stolyar. Scheduling flexible servers with convex delay costs: Heavy-traffic optimality of the generalized $c\mu$-rule. *Operations Research*, 52:836–855, 2004.

[30] M. E. Mayorga, K. M. Taaffe, and R. Arumugam. Allocating flexible servers in serial systems with switching costs. *Annals of Operations Research*, 172:231–242, 2009.

[31] J. O. McClain, L. J. Thomas, and C. Sox. On-the-fly line balancing with very little WIP. *International Journal of Production Economics*, 27:283–289, 1992.

[32] D. G. Pandelis. Optimal use of excess capacity in two interconnected queues. *Mathematical Methods of Operations Research*, 65:179–192, 2007.

[33] M. L. Puterman. *Markov Decision Processes*. John Wiley & Sons, New York, NY, 1994.

[34] Z. Rosberg, P. P. Varaiya, and J. C. Walrand. Optimal control of service in tandem queues. *IEEE Transactions on Automatic Control*, 27:600–609, 1982.

[35] L. Tassiulas and P. P. Bhattacharya. Allocation of interdependent resources for maximal throughput. *Stochastic Models*, 16:27–48, 2000.

[36] L. Tassiulas and A. Ephrimedes. Stability properties of constrained queueing systems and scheduling properties for maximum throughput in multihop radio networks. *IEEE Transactions on Automatic Control*, 37:1936–1948, 1992.

[37] M. P. Van Oyen, E. G. S. Gel, and W. J. Hopp. Performance opportunity for workforce agility in collaborative and noncollabortive work systems. *IIE Transactions*, 33:761–777, 2001.

[38] R. W. Wolff. *Stochastic Modeling and the Theory of Queues*. Prentice Hall, Englewood Cliffs, NJ, 1989.

[39] C.-H. Wu, M. E. Lewis, and M. Veatch. Dynamic allocation of reconfigurable resources in a two-stage tandem queueing system with reliability considerations. *IEEE Transactions on Automatic Control*, 51:309–314, 2006.

[40] C.-H. Wu, D. G. Down, and M. E. Lewis. Heuristics for allocation of reconfigurable resources in a serial line with reliability considerations. *IIE Transactions*, 40:595–611, 2008.

[41] E. Zavadlav, J. O. McClain, and L. J. Thomas. Self-buffering, self-balancing, self-flushing production lines. *Management Science*, 42:1151–1164, 1996.