Flexible Servers in Tandem Lines with Setups

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

Eser Kırkızlar
School of Management
State University of New York-Binghamton
Binghamton, NY 13901-6000, U.S.A.

Abstract

We study the dynamic assignment of flexible servers to stations in the presence of setup costs that are incurred when servers move between stations. We focus on tandem lines with two stations and two servers with the goal of maximizing the long-run average profit. We investigate how the optimal server assignment policy for such systems depends on the magnitude of the setup costs, as well as on the homogeneity of servers and tasks. More specifically, for systems with either homogeneous servers or homogeneous tasks, small buffer sizes, and constant setup cost, we prove the optimality of "multiple threshold" policies (where servers' movement between stations depends on both the number of jobs in the system and the locations of the servers) and determine the values of the thresholds. For systems with heterogeneous servers and tasks, small buffers, and constant setup cost, we provide results that partially characterize the optimal server assignment policy. Finally, for systems with larger buffer sizes and different service rate and setup cost configurations, we present structural results for the optimal policy and provide numerical results that strongly support the optimality of multiple threshold policies.

Keywords: Flexible servers, finite buffers, Markov decision processes, setup costs, tandem production systems, threshold policies.

2000 Mathematics Subject Classification: Primary: 90B22; Secondary 90C40, 60J28.

1 Introduction

Consider a queueing network with $N \geq 2$ stations in tandem and $M \geq 1$ servers. We assume that there is an infinite supply of jobs in front of station 1 and infinite room for completed jobs after station N. We further assume that the buffers between the successive stations are all finite and that the network operates under the manufacturing blocking mechanism. We let $B_j < \infty$ denote the size of the buffer between stations $j-1$ and j for $j \in \{2, \ldots, N\}$. At any given time, there can be at most one job at each station and each server can work on at most one job. Furthermore, the service requirements of each job at each station $j \in \{1, \ldots, N\}$ are independent and exponentially distributed with a mean $m(j)$. We assume, without loss of generality, that $m(j) = 1$ for all $j \in \{1, \ldots, N\}$.

Each server $i \in \{1, \ldots, M\}$ works at rate $0 \leq \mu_{ij} < \infty$ at station $j \in \{1, \ldots, N\}$ (and hence server i is cross-trained to work at all stations j satisfying $\mu_{ij} > 0$). We assume that $\sum_{j=1}^{N} \mu_{ij} > 0$ for $i \in \{1, \ldots, M\}$ and $\sum_{i=1}^{M} \mu_{ij} > 0$ for $j \in \{1, \ldots, N\}$ (because otherwise we have a system with a smaller number of servers or all policies have zero throughput). We also allow several servers to work together on a single job, in which case their service rates are additive. Moreover, we assume that servers can only move between stations when a service completion occurs somewhere in the network, and that the travel times required for servers to go from one station to another station (including any setup times) are negligible, but that there is a cost associated with such server movements. For all $i \in \{1, \ldots, M\}$ and $j, k \in \{1, \ldots, N\}$, let $c_i(j, k)$ be the setup cost incurred when server i moves from station j to station k. We assume that $c_i(j, j) = 0$ and $0 \leq c_i(j, k) < \infty$ for $j \neq k$. We further assume that $c_i(j, k) \leq c_i(j, l) + c_i(l, k)$ for all $i \in \{1, \ldots, M\}$ and $j, k, l \in \{1, \ldots, N\}$, so that the least costly way of moving from one station to another does not include any intermediate stations. Every time there is a service completion at station N, a revenue of v is obtained. Without loss of generality, we assume that $v = 1$.

Our goal is to find the dynamic server assignment policy that maximizes the long-run average profit in the system described above. Most of our results concentrate on systems with two stations and two flexible servers because of the complexity associated with analyzing larger finite-buffered systems. Nevertheless, we provide the problem formulation and some basic results for larger systems as well.

Flexible workforce has been the subject of a significant amount of research in recent years. Here, we provide a review of existing results about the dynamic server assignment problem with setups. We refer the interested reader to Hopp and Van Oyen [12] for a more extensive review of flexible workforce research, and to Andradóttir, Ayhan, and Down [1] or Kırkızlar, Andradóttir, and Ayhan [15] for more concentrated reviews of research on the dynamic server assignment problem.

Most existing works about systems with setups are on polling systems, where there is only one server in the system, and the customers leave after being served at one station. Related work on polling systems includes Duenyas and Van Oyen [9], Gupta and Srinivasan [10], Hofri and Ross [11], Reiman and Wein [20], and references therein.

We are only aware of a small number of papers that study systems with setups apart from polling systems. Andradóttir, Ayhan, and Down [2] consider a general queueing network with outside arrivals, infinite buffers, and random switchover times for the servers. They show that setups do not reduce the capacity in this setting, and construct policies whose performance is arbitrarily close to the maximal system capacity. Andradóttir, Ayhan, and Down [3] study a similar problem in the presence of server failures. Duenyas, Gupta, and Olsen [8] consider a tandem line with a single flexible server, infinite buffers, and positive setup times when the server switches between the stations. They partially characterize the policy that minimizes the total holding cost and develop effective heuristic assignment policies. Iravani, Posner, and Buzacott [13] study a two-stage tandem queue with a flexible server and infinite buffer between the stations, and identify the policy that minimizes the total holding and setup costs. Sennott, Van Oyen, and Iravani [21] consider a tandem line with a dedicated server at each station, one moving server, and infinite buffers between the stations. They allow positive setup costs, holding costs, and setup times, and provide recommendations on how to use the moving server more effectively when the objective is to minimize the total cost. In a more recent paper, Mayorga, Taaffe, and Arumugam [18] study the revenue maximization problem in a two server, two station system with infinite buffers in the presence of switching and holding costs. They provide results about the complexity of the optimal server assignment policy and propose three heuristic policies.

All the papers described in the previous paragraph assume that the storage spaces in the system have infinite capacity. To the best of our knowledge, our work is the first to incorporate setups for a tandem system with finite buffers. This is an important extension of prior work because real systems do not have infinite buffers due to physical constraints (and the buffers can be limited further as a way of controlling the work-in-process). Furthermore, incorporating positive switching costs is also a more realistic representation of actual systems, because server movements often cause some efficiency loss in real life. However, the analysis of a finite-buffered system is difficult because several existing analysis tools (e.g., fluid and diffusion limits) only can be used when analyzing systems with infinite buffers, and most of the time Markov decision problems (MDPs) are analytically intractable even for simple finite-buffered systems. Moreover, the inclusion of setup costs also complicates the analysis due to the necessity of keeping track of all server locations in the state space. Hence, our state space is multi-dimensional, and the Markov chain under consideration does not have a birth-and-death process structure. Consequently, it is a challenging task to calculate performance measures like gain and bias in MDP solutions, even for systems with small buffers.

In this paper, we also consider more general service rate structures compared to most of the previous work studying setups. In particular, Sennott, Van Oyen, and Iravani [21] and Mayorga, Taaffe, and Arumugam [18] study systems with homogeneous tasks (i.e., the service rate only depends on the server). Iravani, Posner, and Buzacott [13] and Duenyas, Gupta, and Olsen [8] only study systems with one server (i.e., the service rate only depends on the task). In addition to studying both of these special service rate structures, in this work we also study systems with servers whose rates can depend on both the server and the task. Andradóttir, Ayhan, and Down [2, 3] also study systems with general service rate structures, but they only consider systems with infinite buffers, and hence are able to employ fluid limits in their solutions. Given the additional complexity resulting from the finite buffers in our systems, the difficulty level of our problem is very high.

The remainder of this paper is organized as follows. In Section 2 we formulate the problem. In Section 3 we provide preliminary results about tandem lines with two stations and setups. In Section 4 we consider systems with two stations, two homogeneous servers (who are equally skilled at all tasks), and constant setup cost. In Section 5, we study systems with two homogeneous stations, two servers, and constant setup cost. In both Sections 4 and 5, we identify the optimal server assignment policies for small buffer sizes and provide our observations about the optimal policy for larger buffer sizes based on numerical and theoretical results. In Section 6 we consider systems with heterogeneous servers and tasks and present structural results about the optimal server assignment policy for constant setup costs, and also perform numerical experiments for systems with arbitrary setup costs. Finally, in Section 7 we make some concluding remarks.

2 Problem Formulation

In this section, we formulate the dynamic server assignment problem in the presence of setups, translate it into a discrete-time Markov decision problem, and finally illustrate our model for systems with two stations and two flexible servers operating under the policy known to be throughput optimal without setup costs.

For all server assignment policies π and all $t \geq 0$, let $Y_{\pi,j}(t) \in \{0, 1, \ldots, B_{j+1} + 2\}$ denote the number of jobs that have been served at station j and are either waiting for service or in service at station $j+1$ at time t under the policy π for $j \in \{1, \ldots, N-1\}$. Similarly, for all server assignment policies π, $t \geq 0$, and $i \in \{1, \ldots, M\}$, let $Z_{\pi,i}(t)$ denote the station that server i was assigned to under the policy π at the time of the most recent service completion prior to time t in the queueing network (letting $Z_{\pi,i}(t)$ be the previous location of server i, rather than the current location of the server, will facilitate the translation of the optimization problem of interest into a Markov decision problem). We will use the stochastic process $\{X_\pi(t)\}$, where $X_\pi(t) = (Y_\pi(t), Z_\pi(t))$, $Y_\pi(t) = (Y_{\pi,1}(t), \ldots, Y_{\pi,N-1}(t))$, and $Z_\pi(t) = (Z_{\pi,1}(t), \ldots, Z_{\pi,M}(t))$ for all $t \geq 0$, to model the state of the system under the policy π as a function of time.

We assume that the class Π of server assignment policies under consideration consists of all Markovian stationary deterministic policies corresponding to the state space $\mathcal{S} \subset \mathbb{R}^{N+M-1}$ of the stochastic processes $\{X_\pi(t)\}$. In other words, the policies in Π specify whether each server is idle or not, and the station that each non-idle server is assigned to, as a function of the current state $x \in \mathcal{S}$ of the stochastic process $\{X_\pi(t)\}$. Hence the server assignments may depend both on the status of the stations and buffers and also on the previous location of the servers. Note that service may be preemptive when $M \geq 2$ (i.e., there is more than one server) because a service completion at one station may trigger the movement of servers that are currently working at other stations. Without loss of generality, we do not consider actions that assign a server to another station and then keep the server idle. The reason is that by simply idling a server without any switchover, we obtain the same departure stream from the system and postpone or avoid the setup costs that could result from idling the server after a switchover (since $0 \leq c_i(j,k) \leq c_i(j,l) + c_i(l,k)$ for all $i \in \{1, \ldots, M\}$ and $j, k, l \in \{1, \ldots, N\}$).

For all $x \in \mathcal{S}$, let A_x denote the set of allowable actions at state x. We use the notation $a_{\sigma_1\sigma_2\ldots\sigma_M}$ to represent the actions, where σ_i is the station to which server $i \in \{1, \ldots, M\}$ is assigned under this action. We use the convention that $\sigma_i = 0$ when server i is voluntarily idled at its current station, and this is treated differently from the case where server i is assigned to a station but is involuntarily idle since that station is not operating. Then, we have $A_x = \mathcal{A} = \bigcup_{\sigma \in \{0,\ldots,N\}^M} \{a_\sigma\}$ for all $x \in \mathcal{S}$. However, without loss of generality, we consider a smaller action set later in this paper because some of the actions are known to be suboptimal in each state. We choose the decision rule d 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 $\pi = (d)^\infty$. Furthermore, we use the notation $d_i(x)$ to denote the assignment of server $i \in \{1, \ldots, M\}$ in state $x \in \mathcal{S}$ under decision rule d. More specifically, $d_i(x) = \sigma_i$ for $i \in \{1, \ldots, M\}$ when $d(x) = a_{\sigma_1\sigma_2\ldots\sigma_M}$. Finally, we use the vector $\delta_d(x) = (d_1(x), \ldots, d_M(x))$ to keep track of the assignments of all servers in state $x \in \mathcal{S}$ under decision rule d.

For all $\pi \in \Pi$ and $t \geq 0$, let $D_\pi(t)$ be the number of departures from the network under the server assignment policy π by time t and $C_\pi(t)$ be the (cumulative) setup cost incurred under the server assignment policy π in the period $[0, t]$. Define

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

the long-run average profit under policy $\pi \in \Pi$. Note that existence of the limit in equation (1) follows because the state space of our Markov chain $\{X_\pi(t)\}$, as well as the immediate rewards, are finite.

We are interested in solving the optimization problem:

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

We now translate the original optimization problem (2) into an equivalent (discrete-time) MDP. Note that although we derive the alternative formulation for a system of finite queues in tandem, our arguments apply to systems with general configurations.

Let $\mathcal{S}_Y \subset \mathbb{R}^{N-1}$ and $\mathcal{S}_Z = \{1, \ldots, N\}^M$ denote the state spaces of the stochastic processes $\{Y_\pi(t)\}$ and $\{Z_\pi(t)\}$, respectively, where $\pi \in \Pi$. For the remainder of this paper, we use the decomposition $x = (y, z)$ and $x' = (y', z')$, where $x, x' \in \mathcal{S}$, $y, y' \in \mathcal{S}_Y$, and $z, z' \in \mathcal{S}_Z$. For all $a \in \mathcal{A}$, let $\pi_a = (d_a)^\infty \in \Pi$ be the server assignment policy with $d_a(x) = a$ for all $x \in \mathcal{S}$. Then it is clear that under our assumptions, the stochastic process $\{Y_{\pi_a}(t)\}$ is a continuous-time Markov chain with state space $\mathcal{S}_Y$ for all $a \in \mathcal{A}$. For all $y, y' \in \mathcal{S}_Y$ and all $a \in \mathcal{A}$, let $Q_a(y, y')$ be the rate at which the continuous-time Markov chain $\{Y_{\pi_a}(t)\}$ goes from state y to state y' (under the server assignment policy π_a). Then, it is not difficult to see that for all $\pi = (d)^\infty \in \Pi$, the stochastic process $\{X_\pi(t)\}$ is a continuous-time Markov chain with state space $\mathcal{S}$ and with transition rates

$$q_d(x, x') = \begin{cases} Q_{d(x)}(y, y') & \text{if } z' = \delta_d(x) + I_z \ , \\ 0 & \text{otherwise,} \end{cases}$$

where I_z is an M-dimensional vector whose i^{th} element is equal to $\mathbf{1}_{(d_i(x)=0)} z_i$ and $\mathbf{1}$ is the identity function. Hence, even if the decision rule voluntarily idles a server, we still keep track of this server's location in the state space.

It is also clear that 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 the continuous-time Markov chain $\{X_\pi(t)\}$ satisfy $\sum_{x' \in \mathcal{S}, x' \neq x} q_d(x, x') \leq q_\pi$ for all $x \in \mathcal{S}$. This shows that $\{X_\pi(t)\}$ is uniformizable for all $\pi \in \Pi$. We let $\{X'_\pi(k)\}$ be the corresponding discrete-time Markov chain, so that $\{X'_\pi(k)\}$ 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 use the fact that $\{X_\pi(t)\}$ is uniformizable to translate the original optimization problem (2) into an equivalent (discrete-time) Markov decision problem (using uniformization to do this type of translation was proposed originally by Lippman [17]). In particular, it is well known that one can 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 at the times of the events of $\{K_\pi(t)\}$, the next state of the continuous-time Markov chain $\{X_\pi(t)\}$ is generated using the transition probabilities of the discrete-time Markov chain $\{X'_\pi(k)\}$.

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

$$R_d(x, x') = \begin{cases} 1 - \sum_{i=1}^{M} c_i(z_i, z'_i) & \text{if } y \in \mathcal{D} \text{ and } y' \in \mathcal{D}_y, \\ -\sum_{i=1}^{M} c_i(z_i, z'_i) & \text{otherwise,} \end{cases}$$

where $\mathcal{D} = \{y \in \mathcal{S}_Y : y_{N-1} > 0\}$, and $\mathcal{D}_y = \{(y_1, \ldots, y_{N-2}, y_{N-1} - 1)\}$ for all $y \in \mathcal{D}$, and $\mathcal{D}_y = \emptyset$ for all $y \notin \mathcal{D}$. Note that incurring the setup cost one transition after the setup occurs does not change the long-run average profit. Hence, it is easy to see that for all $\pi = (d)^\infty \in \Pi$,

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

By the elementary renewal theorem, it is clear that $K_\pi(t)/t \to q_\pi$ almost surely (a.s.) as $t \to \infty$ for all $\pi \in \Pi$. Moreover, it is clear from the strong law of large numbers for Markov chains (see for example Wolff [22], page 164) 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, although the limit may depend on the initial state of the Markov chain $\{X'_\pi(k)\}$ and it may be random (see also Section 3.8 of Kulkarni [16]). Since

$$\left| \frac{1}{K} \sum_{k=1}^{K} R_d(X'_\pi(k-1), X'_\pi(k)) \right| \le 1 + \sum_{i=1}^{M} \max_{1\le j,k \le N} c_i(j,k) < \infty$$

for all $K \ge 1$ and $\sup_{t\ge 0} \mathbb{E}\{[K_\pi(t)/t]^2\} < \infty$ (because $K_\pi(t)$ is a Poisson random variable with mean $q_\pi t$), 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 [7]). This shows that the optimization problem (2) has the same solution as the optimization problem

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

We define $c_i(j,0) = 0$ for all $i \in \{1,\ldots,M\}$ and $j \in \{1,\ldots,N\}$. Then, the strong law of large numbers for Markov chains also gives that for all $\pi \in \Pi$ such that $\pi = (d)^\infty$,

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

where

$$\begin{aligned} R'_d(x) &= \sum_{x'\in\mathcal{S}} p_d(x,x')R_d(x,x') \\ &= \sum_{y'\in\mathcal{S}_Y\setminus\{y\}} p_d(x,(y',\delta_d(x)+I_z))R_d(x,(y',\delta_d(x)+I_z)) \\ &= \sum_{y'\in\mathcal{D}_y} p_d(x,(y',\delta_d(x)+I_z)) - \sum_{y'\in\mathcal{S}_Y\setminus\{y\}} p_d(x,(y',\delta_d(x)+I_z))\sum_{i=1}^{M} c_i(z_i,d_i(x)) \\ &= \sum_{y'\in\mathcal{D}_y} \frac{Q_{d(x)}(y,y')}{q_\pi} - \left(\sum_{y'\in\mathcal{S}_Y\setminus\{y\}} \frac{Q_{d(x)}(y,y')}{q_\pi}\right) \times \left(\sum_{i=1}^{M} c_i(z_i,d_i(x))\right) \end{aligned}$$

for all $x \in \mathcal{S}$ (note that both limits may be random and may depend on the initial state of the Markov chain $\{X'_\pi(k)\}$). Uniform integrability now gives that

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

This shows that the optimization problem (2) has the same solution as the optimization problem

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

Therefore, if one selects $q_\pi = q$ for all $\pi \in \Pi$ (which is always possible in our setting), then the optimization problem (2) has the same set of optimal policies as the optimization problem

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

Finally, it is clear from the above that if

$$R''_d(x) = \sum_{y'\in\mathcal{D}_y} Q_{d(x)}(y,y') - \left(\sum_{y'\in\mathcal{S}_Y\setminus\{y\}} Q_{d(x)}(y,y')\right) \times \left(\sum_{i=1}^{M} c_i(z_i,d_i(x))\right)$$

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).

In order to demonstrate the problem formulation more clearly, we provide an example that employs the server assignment policy that maximizes the throughput of the system with $M = N = 2$ and no setup costs when the servers are ordered such that $\mu_{11}\mu_{22} \geq \mu_{12}\mu_{21}$ (as shown by Andradóttir, Ayhan, and Down [1]). The description of the policy was modified in order to adapt it to our state space.

Example 2.1 Suppose that $M = N = 2$ and $B_2 = B < \infty$. Then

$$\begin{aligned} \mathcal{S} \;=\; & \{(0,1,1),(1,1,1),\ldots,(B+2,1,1),(0,1,2),(1,1,2),\ldots,(B+2,1,2), \\ & (0,2,1),(1,2,1),\ldots,(B+2,2,1),(0,2,2),(1,2,2),\ldots,(B+2,2,2)\}, \end{aligned} \tag{5}$$

where in state $(l, k_1, k_2) \in S$, l refers to the number of jobs that have been processed at station 1 and are either in service or waiting for service at station 2, and k_m refers to the station that server m was previously assigned to (prior to the most recent service completion in the network) for $m = 1, 2$. Assume that for $i = 1, 2$, we have $c_i(1, 2) = c_i^{\uparrow} \geq 0$ and $c_i(2, 1) = c_i^{\downarrow} \geq 0$.

Consider the policy $\pi_0 = (d_0)^{\infty} \in \Pi$, where

$$d_0(x) = \begin{cases} a_{11} & \text{if } x \in \{(0,1,1),(0,1,2),(0,2,1),(0,2,2)\}, \\ a_{22} & \text{if } x \in \{(B+2,1,1),(B+2,1,2),(B+2,2,1),(B+2,2,2)\}, \\ a_{12} & \text{otherwise.} \end{cases} \tag{6}$$

Let $q = q_{\pi_0} = \mu_{11} + \mu_{12} + \mu_{21} + \mu_{22}$. Then

$$p_{d_0}(x, x') = \begin{cases} \frac{\mu_{12}+\mu_{22}}{q} & \text{if } y = 0,\ y' = 0, \text{ and } z' = (1,1), \\ \frac{\mu_{11}+\mu_{21}}{q} & \text{if } y = 0,\ y' = 1, \text{ and } z' = (1,1), \\ \frac{\mu_{22}}{q} & \text{if } y = l,\ y' = l-1, \text{ and } z' = (1,2),\ \forall\, 0 < l < B+2, \\ \frac{\mu_{12}+\mu_{21}}{q} & \text{if } y = l,\ y' = l, \text{ and } z' = (1,2),\ \forall\, 0 < l < B+2, \\ \frac{\mu_{11}}{q} & \text{if } y = l,\ y' = l+1, \text{ and } z' = (1,2),\ \forall\, 0 < l < B+2, \\ \frac{\mu_{12}+\mu_{22}}{q} & \text{if } y = B+2,\ y' = B+1, \text{ and } z' = (2,2), \\ \frac{\mu_{11}+\mu_{21}}{q} & \text{if } y = B+2,\ y' = B+2, \text{ and } z' = (2,2), \\ 0 & \text{otherwise,} \end{cases}$$

and

$$R''_{d_0}(x) = \begin{cases} 0 & \text{if } x = (0,1,1), \\ -(\mu_{11}+\mu_{21})c_2^{\downarrow} & \text{if } x = (0,1,2), \\ -(\mu_{11}+\mu_{21})c_1^{\downarrow} & \text{if } x = (0,2,1), \\ -(\mu_{11}+\mu_{21})(c_1^{\downarrow}+c_2^{\downarrow}) & \text{if } x = (0,2,2), \\ \mu_{22}-(\mu_{11}+\mu_{22})c_2^{\uparrow} & \text{if } x = (l,1,1),\ \forall\, 0<l<B+2, \\ \mu_{22} & \text{if } x = (l,1,2),\ \forall\, 0<l<B+2, \\ \mu_{22}-(\mu_{11}+\mu_{22})(c_1^{\downarrow}+c_2^{\uparrow}) & \text{if } x = (l,2,1),\ \forall\, 0<l<B+2, \\ \mu_{22}-(\mu_{11}+\mu_{22})c_1^{\downarrow} & \text{if } x = (l,2,2),\ \forall\, 0<l<B+2, \\ (\mu_{12}+\mu_{22})(1-c_1^{\uparrow}-c_2^{\uparrow}) & \text{if } x = (B+2,1,1), \\ (\mu_{12}+\mu_{22})(1-c_1^{\uparrow}) & \text{if } x = (B+2,1,2), \\ (\mu_{12}+\mu_{22})(1-c_2^{\uparrow}) & \text{if } x = (B+2,2,1), \\ \mu_{12}+\mu_{22} & \text{if } x = (B+2,2,2). \end{cases}$$

Note that when the policy π_0 is used and $B > 0$, then there are only $B+5$ positive recurrent states in S in the continuous-time stochastic process $\{X_{\pi_0}(t)\}$, namely $(1,1,1)$, $(B+1,2,2)$, and $(l,1,2)$, where $0 \le l \le B+2$. Similarly, when this policy π_0 is used and $B = 0$, then there are only $B+4$ positive recurrent states, namely $(0,1,2)$, $(1,1,1)$, $(1,2,2)$, and $(2,1,2)$.

3 Preliminary Results

In this section, we provide some preliminary results about tandem lines with two stations and setup costs. We first present a result about the form of the optimal server assignment policy.

Lemma 3.1 *For a tandem line with $N = 2$, $M \ge 2$, and nonnegative setup costs, there exists an optimal policy that does not idle any server voluntarily when the first station is blocked or when the second station is starved.*

Proof: When the first station is blocked, the system is in a state $s = (B+2, z_1, \ldots, z_M)$, where $(z_1, \ldots, z_M) \in S_Z$. Now compare two policies $\pi_1 = (d^1)^\infty$ and $\pi_2 = (d^2)^\infty$ that agree with each other apart from state s. Assume that $d_i^1(s) = z_i$ and $d_i^2(s) = 0$ for some $i \in \{1, \ldots, M\}$, and $d_j^1(s) = d_j^2(s)$ for $j \in \{1, \ldots, M\} \setminus \{i\}$. If $z_i = 1$, then the performance of π_1 and π_2 will be identical (since keeping a server at station 1 is equivalent to idling that server in terms of cost). If $z_i = 2$, then the next service completion under policy π_1 will never be later than the one under policy π_2, the system state will be the same after the next service completion, and no extra cost will have been incurred by keeping server i at the second station. Hence, $D_{\pi_1}(t) \ge D_{\pi_2}(t)$ for all $t \ge 0$.

Note that it is possible to have $C_{\pi_1}(t) \geq C_{\pi_2}(t)$ because of the faster transitions under π_1. We now restrict ourselves to policies with nonzero departures from the system without loss of generality (this is possible because the optimal policy must have positive number of departures under our assumptions on the service rates). Define the setup cost per item produced up to time t under policy $\pi \in \Pi$ as $u_\pi(t) = C_\pi(t)/D_\pi(t)$. Under both policies ($\pi_1$ and π_2), the system goes through the same sequence of states, and at the time of each departure the total setup cost incurred under policy π_1 is equal to the total setup cost incurred under policy π_2. Hence, we can conclude that $u_{\pi_1}(t) = u_{\pi_2}(t)$ for all $t \geq 0$. For all $\pi \in \Pi$, we have

$$P_\pi = \lim_{t\to\infty} \mathbb{E}\left\{\frac{D_\pi(t)}{t}\Big(1 - \frac{C_\pi(t)}{D_\pi(t)}\Big)\right\} = \lim_{t\to\infty} \mathbb{E}\left\{\frac{D_\pi(t)}{t}(1 - u_\pi(t))\right\}.$$

Consequently, $P_{\pi_1} \geq P_{\pi_2}$. Hence, there exists an optimal policy that never idles the servers when the first station is blocked. A similar logic follows when the second station is starved. □

For a system with two stations, consider the reversed line where station 1 is followed by station 2, and keep the labeling of the stations as in the original line (i.e., station 2 is the upstream station and station 1 is the downstream station in the reversed line). Let B denote the buffer size between the stations. Assume that the forward line operates under a policy $\pi = (d)^\infty$ and that the reversed line operates under a policy $\pi_R = (d_R)^\infty$, where $d_R(l, z) = d(B+2-l, z)$ for $0 \leq l \leq B+2$ and $z \in S_Z$ (in both the forward and reversed lines, $z_i = j$ if the previous location of server i is station j). The following reversibility result will be used to simplify our results and proofs.

Lemma 3.2 *When $N = 2$, the policy π is optimal for the forward line if and only if the policy π_R is optimal for the reversed line.*

Proof: Let $\kappa_{\pi,1}(x)$ and $\kappa_{\pi,2}(x)$ denote the sets of servers assigned to stations 1 and 2, respectively, under policy π when the original line is in state $x \in S$. Then we see that for $\{X_\pi(t)\}$, the transition rate from state $x = (l, z)$ to $(l+1, z')$ is $\sum_{i\in\kappa_{\pi,1}(x)} \mu_{i1}$ for $l \in \{0, \ldots, B+1\}$ and the transition rate from state x to $(l-1, z')$ is $\sum_{i\in\kappa_{\pi,2}(x)} \mu_{i2}$ for $l \in \{1, \ldots, B+2\}$ and $z, z' \in S_Z$ (where z' is determined by $\kappa_{\pi,1}(x)$ and $\kappa_{\pi,2}(x)$). Now, let $\{(Y_{\pi_R}(t), Z_{\pi_R}(t))\}$ be the Markov chain corresponding to the reversed line. It is easy to see that the stochastic process $\{(B+2-Y_{\pi_R}(t), Z_{\pi_R}(t))\}$ has the same transition rates as the stochastic process $\{(Y_\pi(t), Z_\pi(t))\}$. Hence these two processes are stochastically equivalent. Consequently, the long-run average profit of the forward line under policy π is equal to the long-run average profit of the reversed line under policy π_R (because the departures from one system correspond to departures from the first station of the other system, and the buffer size between the stations is finite), and the result follows. □

4 Systems with Homogeneous Servers

In this section, we consider a tandem line with two stations and two homogeneous servers that have the same service rate at each task. Hence, the service rates depend only on the station, so that $\mu_{ij} = \gamma_j > 0$ for $i, j \in \{1, 2\}$. We focus on the case where $\gamma_1 \geq \gamma_2$, so that the task at station 1 takes on average less time than the task at station 2. The results for $\gamma_1 < \gamma_2$ are similar, and can be deduced from the results for $\gamma_1 \geq \gamma_2$ using the reversibility of two-station tandem lines, as shown in Lemma 3.2. They are provided in Appendix A. Furthermore, we assume that $c_i(1, 2) = c_i(2, 1) = c \geq 0$ for $i \in \{1, 2\}$. This is a reasonable assumption if the setup costs are due to the movement of the servers or if every machine requires similar setup procedures. Our state space $\mathcal{S}$ is given in (5). More specifically, we will consider systems with $B = 0$, $B = 1$, and $B > 1$ in Sections 4.1, 4.2, and 4.3, respectively.

4.1 Systems with Homogeneous Servers and No Buffer

In this section we provide the optimal server assignment policy for a system that has a buffer of size zero between the stations. We will need the decision rule

$$d_1(x) = \begin{cases} a_{12} & \text{if } x = (0, z_1, z_2) \text{ for all } (z_1, z_2) \in \mathcal{S}_Z \text{ or} \\ & \quad x = (1, z_1, z_2) \text{ for all } (z_1, z_2) \in \mathcal{S}_Z \setminus \{(2, 2)\}, \\ a_{22} & \text{if } x = (2, z_1, z_2) \text{ for all } (z_1, z_2) \in \mathcal{S}_Z \text{ or } x = (1, 2, 2). \end{cases} \tag{7}$$

The proof of the following theorem is provided in Appendix B. The case when $\gamma_1 < \gamma_2$ is covered by Corollary A.1, provided in Appendix A.

Theorem 4.1 *For a Markovian tandem line with two stations, two flexible servers, and buffer of size zero between the stations, if $\mu_{ij} = \gamma_j$ for $i, j \in \{1, 2\}$ and $\gamma_1 \geq \gamma_2$, then the optimal server assignment policy $\pi^* = (d^*)^\infty$ is as follows:*

(i) *If $0 \leq c \leq \frac{\gamma_2}{2\gamma_1 + 4\gamma_2}$, then $d^*(x) = d_0(x)$ for all $x \in \mathcal{S}$ (see equation (6)) and the recurrent states are $(0, 1, 2), (1, 1, 1), (1, 2, 2)$, and $(2, 1, 2)$.*

(ii) *If $\frac{\gamma_2}{2\gamma_1 + 4\gamma_2} < c \leq \frac{\gamma_1^2}{2\gamma_1^2 + 2\gamma_1\gamma_2 + 2\gamma_2^2}$, then $d^*(x) = d_1(x)$ for all $x \in \mathcal{S}$ (see equation (7)) and the recurrent states are $(0, 1, 2), (0, 2, 2), (1, 1, 2), (1, 2, 2)$, and $(2, 1, 2)$.*

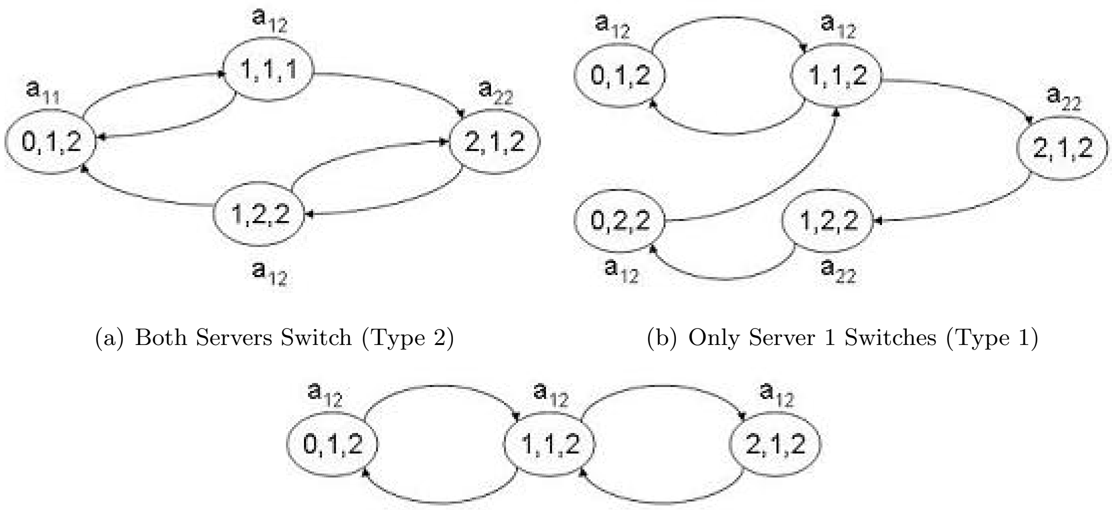
(iii) *If $c > \frac{\gamma_1^2}{2\gamma_1^2 + 2\gamma_1\gamma_2 + 2\gamma_2^2}$, then $d^*(x) = a_{12}$ for all $x \in \mathcal{S}$ and the recurrent states are $(0, 1, 2), (1, 1, 2)$, and $(2, 1, 2)$.*

Note that since we have identical servers, the policies described in Theorem 4.1 are not unique. For every specified policy, there is an alternative optimal policy where the roles of the servers are reversed. Moreover, the interval for c in part (ii) of Theorem 4.1 is non-empty because $\gamma_1 \geq \gamma_2$. Finally, the intervals considered in the theorem span all values of $c \geq 0$.

Theorem 4.1 and Corollary A.1 show that when the servers are homogeneous, the optimal policy is of one of the following three types:

- Neither server switches (Type 0);
- Only one server switches (Type 1);
- Both servers switch (Type 2).

The recurrent states, together with the actions in these states under the optimal policies of Theorem 4.1 are depicted in Figures 1(a), 1(b), and 1(c).



(a) Both Servers Switch (Type 2)

(b) Only Server 1 Switches (Type 1)

(c) Only Server 2 Switches (Type 1)

Figure 1: Recurrent States and Optimal Actions in Theorem 4.1

In the case without setups and with arbitrary service rates, both servers have a primary assignment at a station (i.e., each server works at their assigned station as long as it is neither blocked nor starved) and the servers do not idle. However, in the presence of positive setup costs, Theorem 4.1 shows that servers may have a preferred or dedicated assignment at a station, but not a primary assignment, because they may idle at their dedicated station or work at a less preferred station even when their preferred station is operating, to avoid multiple switchovers.

More specifically, for small values of c, part (i) of Theorem 4.1 shows that both servers have primary assignments and the optimal policy is of Type 2 (see Figure 1(a)). Note that this policy is the same as the optimal policy for systems with $c = 0$, as shown by Andradóttir, Ayhan, and Down [1]. For intermediate values of c, we observe that only one server switches between stations (i.e., the optimal policy is of Type 1). In particular, part (ii) of Theorem 4.1 shows that server 1 has a preferred assignment at station 1 and server 2 has a dedicated assignment at station 2 (see Figure 1(b)). Finally, for large values of c, both servers are dedicated (i.e., the optimal policy is of Type 0), as shown in part (iii) of Theorem 4.1 (see Figure 1(c)). Note that idling occurs under both Type 1 and Type 0 policies. An examination of the bounds on c in Theorem 4.1 shows that the optimal policy is not of Type 2 for any value of $c > \frac{1}{6}$, and the optimal policy is of Type 0 for all values of $c > \frac{1}{2}$.

Theorem 4.1 also introduces the notion of "multiple threshold" policies. In other words, servers move between stations when the number of jobs that are in service or waiting for service at station 2 reaches a threshold. Furthermore, the value of this threshold may depend on the location of the switching server. We use the notation $t_i(z)$ to denote the threshold where server $i \in \{1, \ldots, M\}$ switches from station z_i to the other station $3 - z_i$ when the previous locations of the servers are represented in the vector $z \in \mathcal{S}_Z$. We use the convention that server i is assigned to station $3 - z_i$ when the system is in state $(t_i(z), z)$. In Figure 1(a), server 1 switches between stations based on the thresholds $t_1(1, 2) = 2$, $t_1(2, 2) = 1$, and server 2 switches between stations based on $t_2(1, 1) = 1$, and $t_2(1, 2) = 0$. In Figure 1(b), server 1 switches based on $t_1(1, 2) = 2$ and $t_1(2, 2) = 0$, but server 2 is dedicated. Finally, in Figure 1(c), both servers are dedicated.

For a given setup cost, the efficiency loss resulting from using a policy that is optimal for another setup cost can be very high. For example, one can show that the long-run average profits associated with the Type 0 and Type 2 policies of Theorem 4.1 are $\frac{\gamma_1\gamma_2(\gamma_1+\gamma_2)}{\gamma_1^2+\gamma_1\gamma_2+\gamma_2^2}$ and $\frac{2\gamma_1\gamma_2(1-2c)}{\gamma_1+\gamma_2}$. Hence, when c is large (e.g., greater than 0.5), the efficiency loss associated with using the wrong policy is at least 100%.

4.2 Systems with Homogeneous Servers and Buffer of Size One

In this section, we provide the optimal server assignment policy for a system with a buffer of size one between the stations. We will need the decision rules

$$d_2(x) = \begin{cases} a_{12} & \text{if } x = (y, z_1, z_2) \text{ for all } y \in \{0, 1\} \text{ and } (z_1, z_2) \in \mathcal{S}_Z \text{ or} \\ & \quad x = (2, z_1, z_2) \text{ for all } (z_1, z_2) \in \mathcal{S}_Z \setminus \{(2, 2)\}, \\ a_{22} & \text{if } x = (3, z_1, z_2) \text{ for all } (z_1, z_2) \in \mathcal{S}_Z \text{ or } x = (2, 2, 2), \end{cases} \tag{8}$$

and

$$d_3(x) = \begin{cases} a_{12} & \text{if } x = (0, z_1, z_2) \text{ for all } (z_1, z_2) \in \mathcal{S}_Z \text{ or} \\ & \quad x = (y, z_1, z_2) \text{ for all } y \in \{1, 2\} \text{ and } (z_1, z_2) \in \mathcal{S}_Z \setminus \{(2, 2)\}, \\ a_{22} & \text{if } x = (y, 2, 2) \text{ for all } y \in \{1, 2\} \text{ or} \\ & \quad x = (3, z_1, z_2) \text{ for all } (z_1, z_2) \in \mathcal{S}_Z. \end{cases} \tag{9}$$

The proof of the following theorem is similar to that of Theorem 4.1. It is omitted here due to space restrictions, however it is provided in Kırkızlar [14]. The case when $\gamma_1 < \gamma_2$ is covered by

Corollary A.2, provided in Appendix A.

Theorem 4.2 *For a Markovian tandem line with two stations, two flexible servers, and buffer of size one between the stations, if $\mu_{ij} = \gamma_j$ for $i, j \in \{1, 2\}$ and $\gamma_1 \geq \gamma_2$, then the optimal server assignment policy $\pi^* = (d^*)^\infty$ is as follows:*

(i) *If $0 \leq c \leq \frac{\gamma_2}{2\gamma_1+2\gamma_2}$, then*

$$d^*(x) = \begin{cases} a_{11} & \text{if } x = (0, z_1, z_2) \text{ for all } (z_1, z_2) \in \mathcal{S}_Z \text{ or } x = (1, 1, 1), \\ a_{12} & \text{if } x = (1, z_1, z_2) \text{ for all } (z_1, z_2) \in \mathcal{S}_Z \setminus \{(1, 1, 1)\} \text{ or} \\ & \quad x = (2, z_2, z_2) \text{ for all } (z_1, z_2) \in \mathcal{S}_Z \setminus \{(2, 2, 2)\}, \\ a_{22} & \text{if } x = (3, z_1, z_2) \text{ for all } (z_1, z_2) \in \mathcal{S}_Z \text{ or } x = (2, 2, 2), \end{cases}$$

and the recurrent states are $(0, 1, 2)$, $(1, 1, 1)$, $(1, 1, 2)$, $(1, 2, 2)$, $(2, 1, 1)$, $(2, 1, 2)$, $(2, 2, 2)$, and $(3, 1, 2)$.

(ii) *If $\frac{\gamma_2}{2\gamma_1+2\gamma_2} < c \leq \min\{\frac{\gamma_1^2}{2\gamma_1^2+2\gamma_2^2}, \frac{2\gamma_1\gamma_2+\gamma_2^2}{2\gamma_1^2+4\gamma_1\gamma_2}\}$, then $d^*(x) = d_2(x)$ for all $x \in \mathcal{S}$ (see equation (8)) and the recurrent states are $(0, 1, 2)$, $(1, 1, 2)$, $(1, 2, 2)$, $(2, 1, 2)$, $(2, 2, 2)$, and $(3, 1, 2)$.*

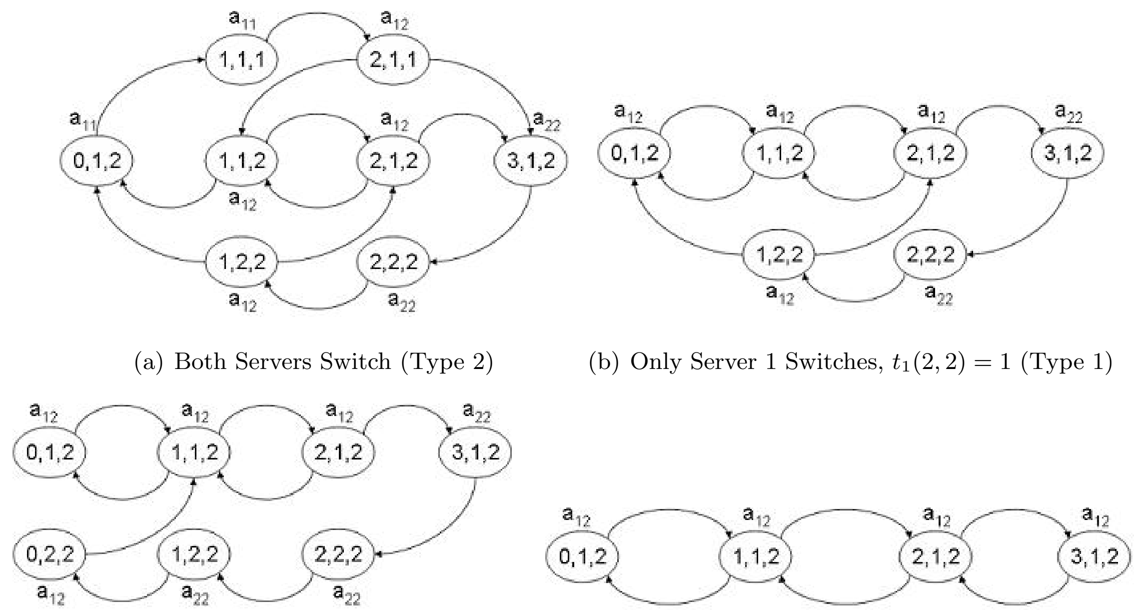
(iii) *If $\gamma_1^2 > \gamma_1\gamma_2 + \gamma_2^2$ and $\frac{2\gamma_1\gamma_2+\gamma_2^2}{2\gamma_1^2+4\gamma_1\gamma_2} < c \leq \frac{3\gamma_1^3+\gamma_1^2\gamma_2-\gamma_1\gamma_2^2}{4\gamma_1^3+4\gamma_1^2\gamma_2+4\gamma_1\gamma_2^2+4\gamma_2^3}$, then $d^*(x) = d_3(x)$ for all $x \in \mathcal{S}$ (see equation (9)) and the recurrent states are $(0, 1, 2)$, $(0, 2, 2)$, $(1, 1, 2)$, $(1, 2, 2)$, $(2, 1, 2)$, $(2, 2, 2)$, and $(3, 1, 2)$.*

(iv) *If $\gamma_1^2 > \gamma_1\gamma_2 + \gamma_2^2$ and $c > \frac{3\gamma_1^3+\gamma_1^2\gamma_2-\gamma_1\gamma_2^2}{4\gamma_1^3+4\gamma_1^2\gamma_2+4\gamma_1\gamma_2^2+4\gamma_2^3}$, then $d^*(x) = a_{12}$ for all $x \in \mathcal{S}$ and the recurrent states are $(0, 1, 2)$, $(1, 1, 2)$, $(2, 1, 2)$, and $(3, 1, 2)$.*

(v) *If $\gamma_1^2 \leq \gamma_1\gamma_2 + \gamma_2^2$ and $c > \frac{\gamma_1^2}{2\gamma_1^2+2\gamma_2^2}$, then $d^*(x) = a_{12}$ for all $x \in \mathcal{S}$ and the recurrent states are $(0, 1, 2)$, $(1, 1, 2)$, $(2, 1, 2)$, and $(3, 1, 2)$.*

Note that the policies in Theorem 4.2 are not unique (we can relabel the servers and obtain alternative optimal policies where the roles of the servers are switched). Moreover, the interval for c in part (ii) of Theorem 4.2 is non-empty because $\gamma_1 \geq \gamma_2$, and the interval in part (iii) of Theorem 4.2 is non-empty when $\gamma_1 \geq \gamma_2$ and $\gamma_1^2 > \gamma_1\gamma_2 + \gamma_2^2$. Finally, the intervals considered in the theorem span all values of $c \geq 0$.

We now depict the recurrent states and the optimal actions in Theorem 4.2. More specifically, Figure 2(a) shows the optimal policy of Type 2 corresponding to part (i) of Theorem 4.2. Figures 2(b) and 2(c) show the optimal policies of Type 1 (with different thresholds) corresponding to parts (ii) and (iii) of Theorem 4.2, respectively. Finally, Figure 2(d) shows the optimal policy of Type 0, corresponding to parts (iv) and (v) of Theorem 4.2.



(a) Both Servers Switch (Type 2)

(b) Only Server 1 Switches, $t_1(2,2) = 1$ (Type 1)

(c) Only Server 1 Switches, $t_1(2,2) = 0$ (Type 1)

(d) Neither Server Switches (Type 0)

Figure 2: Recurrent States and Optimal Actions in Theorem 4.2

As for $B = 0$, we see that server 1 has a preferred assignment at station 1 and server 2 has a preferred assignment at station 2 for all values of c. However, for $B = 0$ all the systems go through the same set of optimal policies as c increases (although the cutoffs on the value of c depend on the service rates). This is no longer correct when $B = 1$. More specifically, when $B = 1$ and the service rates satisfy $\gamma_1^2 \leq \gamma_1\gamma_2 + \gamma_2^2$, then we observe three different optimal policies for different values of the setup cost. On the other hand, if the service rates satisfy $\gamma_1^2 > \gamma_1\gamma_2 + \gamma_2^2$, then we observe four different optimal policies, depending on the value of the setup cost (in particular, as the setup cost increases, the first server completes more jobs at station 2 before switching back to station 1). Note that $\gamma_1^2 > \gamma_1\gamma_2 + \gamma_2^2$ implies that the difference between the magnitudes of γ_1 and γ_2 is guaranteed to be significant, and hence server 1 can spend more time at station 2 before switching back to station 1.

Also note that the transition from one optimal policy to another follows a similar pattern when $B = 0$ and $B = 1$. In both cases, for small values of c both servers switch and the optimal policy is of Type 2. Moreover, this policy is also optimal for systems with $c = 0$ even though it differs from π_0. This is not surprising because any non-idling policy is known to be optimal for systems with $c = 0$ and homogeneous servers, see Andradóttir, Ayhan, and Down [1]. For intermediate values of c, only one server switches (server 1 is the switching server when $\gamma_1 \geq \gamma_2$ and server 2 is the switching server when $\gamma_1 < \gamma_2$) and the optimal policy is of Type 1, and for large values of c neither server switches and the optimal policy is of Type 0. Moreover, when the optimal policy is of Type 1, we observe that the switching server is the one that has a preferred assignment at the faster station. Finally, we see that the optimal policy is not of Type 2 when $c > \frac{1}{4}$ and is of Type 0 when $c > \frac{3}{4}$.

4.3 Systems with Homogeneous Servers and Multiple Buffer Spaces

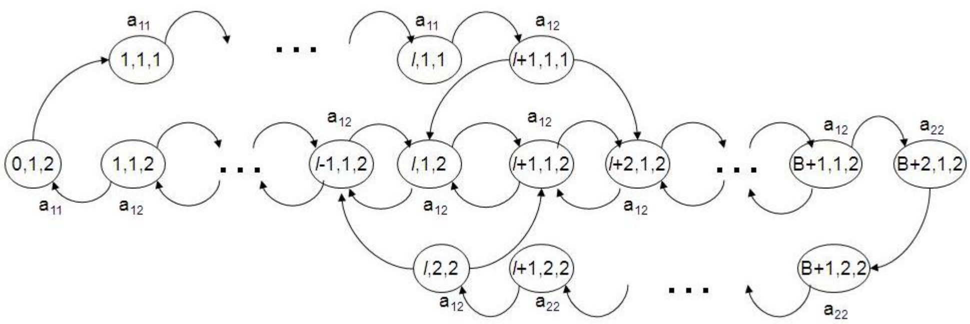
Theorems 4.1 and 4.2 provide the optimal server assignment policy for systems with two stations and two homogeneous servers when the buffer size between the stations is zero or one. In this section we provide our observations about the form of the optimal policy for systems with $B > 1$.

We randomly generate 50,000 systems with the service rate at each station independently drawn from a uniform distribution with range [0.5,2.5] and the setup cost drawn from a uniform distribution with range (0,0.5) (we have also tried a larger range for the setup cost and observed that most of the optimal policies ended up being of Type 0 with no switching). Furthermore, the buffer size B between the stations is drawn from a discrete uniform distribution with range $\{2, 3, 4, 5\}$. For each system, we determine the optimal server assignment policy using the policy iteration algorithm for weakly communicating Markov chains.

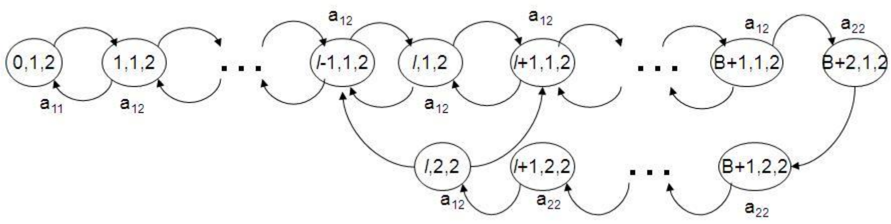
Our numerical results for systems with homogeneous servers suggest that the optimal server assignment policy is similar to that of systems with an intermediate buffer of size one, see Theorem 4.2 and Figure 2. Both servers have preferred assignments, not primary assignments, and the optimal policy is a multiple threshold policy when it is of Type 1 or 2. Moreover, if the optimal policy is of Type 1, one server is dedicated to the slower station. Furthermore, the thresholds have the special structure described below.

Consider a system where homogeneous servers 1 and 2 have preferred assignments at stations 1 and 2, respectively. If the optimal policy is of Type 2, we observe that $t_1(1, 2) = B + 2$, $t_1(2, 2) = l$, $t_2(1, 1) = l + 1$, and $t_2(1, 2) = 0$ for some $l \in \{1, \ldots, B\}$ and for both $\gamma_1 \geq \gamma_2$ and $\gamma_1 < \gamma_2$ (note that $l = 1$ in the policy of Figure 2(a)). This policy is depicted in Figure 3(a). Note that $1 \leq l \leq B$ guarantees that a server that switches to a less preferred station does not switch back immediately. We believe that for systems with homogeneous servers, the optimal policy delays the switchovers because this reduces setup costs without making the servers ineffective. The similarity of the thresholds $t_1(2, 2)$ and $t_2(1, 1)$ suggests that it is preferable to keep the number of jobs in the system close to these thresholds, probably to minimize future setup costs.

Moreover, when the servers are homogeneous, we observe that if the optimal policy is of Type 1 and server 1 is the switching server, then $t_1(1, 2) = B + 2$ and $t_1(2, 2) = l$ for some $l \in \{0, \ldots, B\}$ (note that $l = 1$ in the policy of Figure 2(b) and $l = 0$ in the policy of Figure 2(c)). This policy is depicted in Figure 3(b). Like the optimal Type 2 policy, this policy delays the switchovers because the servers do not have any special skills at any task. Similarly, if the



(a) Both Servers Switch, Primary Assignment (Type 2)



(b) Only Server 1 Switches (Type 1)

Figure 3: Recurrent States and Optimal Actions in Sections 4.3 and 5.3

optimal policy is of Type 1 and server 2 is the switching server, then $t_2(1,1) = l$ and $t_2(1,2) = 0$ for some $l \in \{2, \dots, B+2\}$. Thus, the switching server has a preferred assignment (not a primary assignment), and the optimal policy maintains a balance between avoiding idling the dedicated server (i.e., through starvation at the second station) and avoiding setups for the switching server.

Finally, we show that when $\gamma_1 \geq \gamma_2$ and the optimal policy is of Type 1, the threshold where the switching server returns to its preferred station (station 1) decreases when c increases. It then follows from Lemma 3.2 that when $\gamma_1 < \gamma_2$, the threshold where the switching server returns to its preferred station 2 increases with c. Thus the switching server helps the dedicated server with more jobs before switching back to its preferred station for larger setup costs.

Proposition 4.1 *In a tandem line with two stations, two homogeneous servers (so that $\mu_{ij} = \gamma_j$ for $i, j \in \{1, 2\}$), and buffer of size $B > 1$ between the stations, the threshold for the optimal Type 1 policy decreases as the setup cost $c > 0$ increases when $\gamma_1 \geq \gamma_2$.*

Proof: First assume that $\gamma_1 > \gamma_2$. Let $\pi_l = (d_l)^\infty$ be a Type 1 policy with $t_1(2,2) = l$, where $l \in \{0, \dots, B\}$. It is not difficult to show that

$$P_{\pi_l} = \frac{2\gamma_2\big((B+2-l-2c)\gamma_1^{B+l+4} - (B+2-l-4c)\gamma_1^{B+l+3}\gamma_2 - 2c\gamma_1^{B+l+2}\gamma_2^2 - \gamma_1^{B+2}\gamma_2^{l+2} + \gamma_1^l\gamma_2^{B+4}\big)}{(B+2-l)\gamma_1^{B+l+4} - (B+2-l)\gamma_1^{B+l+2}\gamma_2^2 - 2\gamma_1^{B+2}\gamma_2^{l+2} + 2\gamma_1^l\gamma_2^{B+4}}.$$

Some algebra shows that for $l \in \{0, \dots, B-1\}$, $P_{\pi_l} - P_{\pi_{l+1}} = (\alpha_{l,1} + c\alpha_{l,2})/\alpha_{l,3}$, where

$$\alpha_{l,1} = 2(\gamma_1 - \gamma_2)^2 \gamma_1^{B+l+2} \gamma_2^3 \Big((B+1-l)\gamma_1^{B+2}\gamma_2^l - (B+2-l)\gamma_1^{B+1}\gamma_2^{l+1} + \gamma_1^l \gamma_2^{B+2}\Big),$$

$$\begin{aligned}
\alpha_{l,2} &= 4(\gamma_1-\gamma_2)^3\gamma_1^{2B+l+3}\gamma_2(\gamma_1^{l+2}+\gamma_1^{l+1}\gamma_2-2\gamma_2^{l+2}),\\
\alpha_{l,3} &= \Big((B+2-l)\gamma_1^{B+l+4}-(B+2-l)\gamma_1^{B+l+2}\gamma_2^2-2\gamma_1^{B+2}\gamma_2^{l+2}+2\gamma_1^l\gamma_2^{B+4}\Big)\\
&\quad\times\Big((B+1-l)\gamma_1^{B+l+4}-(B+1-l)\gamma_1^{B+l+2}\gamma_2^2-2\gamma_1^{B+1}\gamma_2^{l+3}+2\gamma_1^l\gamma_2^{B+4}\Big).
\end{aligned}$$

The first term in $\alpha_{l,3}$ is positive for all $l \in \{0,\ldots,B-1\}$, because it can be rewritten as

$$\begin{aligned}
&(B+2-l)\gamma_1^{B+2+l}(\gamma_1^2-\gamma_2^2)-2\gamma_1^l\gamma_2^{l+2}(\gamma_1^{B+2-l}-\gamma_2^{B+2-l})\\
&=(\gamma_1-\gamma_2)\Big((B+2-l)\gamma_1^{B+2+l}(\gamma_1+\gamma_2)-2\gamma_1^l\gamma_2^{B+3}\sum_{i=0}^{B+1-l}(\frac{\gamma_1}{\gamma_2})^i\Big)\\
&>2(\gamma_1-\gamma_2)\Big((B+2-l)\gamma_1^{B+2+l}\gamma_2-\gamma_1^l\gamma_2^{B+3}\sum_{i=0}^{B+1-l}(\frac{\gamma_1}{\gamma_2})^i\Big)>0,
\end{aligned}$$

where the last inequality follows because $\gamma_1^l\gamma_2^{B+3}(\gamma_1/\gamma_2)^i < \gamma_1^{B+2+l}\gamma_2$ for all $i \in \{0,\ldots,B+1-l\}$. Similar calculations show that the second term in $\alpha_{l,3}$ is also positive, because it can be rewritten as

$$\begin{aligned}
&(B+1-l)\gamma_1^{B+2+l}(\gamma_1^2-\gamma_2^2)-2\gamma_1^l\gamma_2^{l+3}(\gamma_1^{B+1-l}-\gamma_2^{B+1-l})\\
&>2(\gamma_1-\gamma_2)\Big((B+1-l)\gamma_1^{B+2+l}\gamma_2-\gamma_1^l\gamma_2^{B+3}\sum_{i=0}^{B-l}(\frac{\gamma_1}{\gamma_2})^i\Big)>0.
\end{aligned}$$

Thus we have shown that $\alpha_{l,3} > 0$ for $l \in \{0,\ldots,B-1\}$. Moreover, $\alpha_{l,2} > 0$ trivially. This shows that $P_{\pi_l} > P_{\pi_{l+1}}$ for large enough c.

Next assume that $\gamma_1 = \gamma_2 = \rho$. Some algebra shows that for $l \in \{0,\ldots,B-1\}$,

$$P_{\pi_l}-P_{\pi_{l+1}}=\frac{\rho\Big(2+B^2+B(3-2l)-3l+l^2+4c(3+2l)\Big)}{(1+B-l)(2+B-l)(4+B+l)(5+B+l)}. \tag{10}$$

This expression is positive for $l \in \{0,\ldots,B-1\}$ and large enough c. This completes the proof. □

5 Systems with Homogeneous Tasks

In this section, we consider a tandem line with two homogeneous tasks and two servers. In other words, we assume the service rates depend only on the server, so that $\mu_{ij} = \mu_j > 0$ for $i,j \in \{1,2\}$. This assumption has been made frequently in other works, even for papers that study systems without setups, including the bucket brigades results of Bartholdi and Eisenstein [4] and Bartholdi, Eisenstein, and Foley [5]. Without loss of generality, assume that $\mu_1 \geq \mu_2$ because we can relabel the servers otherwise. As in Section 4, we assume that $c_i(1,2) = c_i(2,1) = c \geq 0$ for $i \in \{1,2\}$ and our state space $\mathcal{S}$ is as given in (5). More specifically, we will consider systems with $B = 0$, $B = 1$, and $B > 1$ in Sections 5.1, 5.2, and 5.3, respectively.

5.1 Systems with Homogeneous Tasks and No Buffer

In this section, we identify the optimal server assignment policy for the system with buffer of size zero between the stations. The proof of the following theorem is similar to that of Theorem 4.1. It is omitted here for reasons of brevity, however it is provided in Kırkızlar [14].

Theorem 5.1 *For a Markovian tandem line with two stations, two flexible servers, and buffer of size zero between the stations, if $\mu_{ij} = \mu_i$ for $i, j \in \{1, 2\}$, then the optimal server assignment policy $\pi^* = (d^*)^\infty$ is as follows:*

(i) *If $0 \leq c \leq \frac{\mu_2}{4\mu_1+2\mu_2}$, then then $d^*(x) = d_0(x)$ for all $x \in \mathcal{S}$ (see equation (6)) and the recurrent states are $(0, 1, 2), (1, 1, 1), (1, 2, 2)$, and $(2, 1, 2)$.*

(ii) *If $\frac{\mu_2}{4\mu_1+2\mu_2} < c \leq \frac{2\mu_1^2-\mu_1\mu_2}{2\mu_1^2+2\mu_1\mu_2+2\mu_2^2}$, then $d^*(x) = d_1(x)$ for all $x \in \mathcal{S}$ (see equation (7)) and the recurrent states are $(0, 1, 2), (0, 2, 2), (1, 1, 2), (1, 2, 2)$, and $(2, 1, 2)$.*

(iii) *If $c > \frac{2\mu_1^2-\mu_1\mu_2}{2\mu_1^2+2\mu_1\mu_2+2\mu_2^2}$, then $d^*(x) = a_{12}$ for all $x \in \mathcal{S}$ and the recurrent states are $(0, 1, 2), (1, 1, 2)$, and $(2, 1, 2)$.*

Note that these optimal policies are not unique. In particular, it follows from Lemma 3.2 that the policies where the roles of the servers are switched are also optimal. Note that the interval in part (ii) of Theorem 5.1 is non-empty when $\mu_1 \geq \mu_2$. Moreover, the intervals considered in the theorem span all values of $c \geq 0$.

We see that the optimal server assignment policy in Theorem 5.1 is similar to the optimal policy provided in Theorem 4.1 for the case where the servers are homogeneous, and has one of the forms shown in Figure 1. More specifically, both servers have preferred assignments. For small values of c they both switch to the other station when their assigned station is not operating. For intermediate values of c, the optimal policy is a multiple threshold policy with one switching server, and for large values of c, the optimal policy does not allow the servers to switch (in the recurrent states).

We also see that the policy is not of Type 2 when $c > \frac{1}{6}$, and the policy is of Type 0 when $c > 1$. Note that the switching is possible for a larger range of setup costs compared to Theorem 4.1. For example when $\mu_1 = 10$ and $\mu_2 = 1$, switching policies are optimal when $c < 0.856$. However, we saw in Section 4 that when the servers are homogeneous, no switching policy is optimal for $c > \frac{1}{2}$. When a server is extremely fast compared to the other server, it may be advantageous to move this server, even for high values of the setup cost, to benefit from the high service rate. Similarly, when the servers are homogeneous, we take advantage of the faster station by assigning servers disproportionally to the slower station. However, the impact of taking advantage of a fast server is larger than the impact of exploiting a fast station, because in the former case the throughput can be as large as $(\mu_1 + \mu_2)/2$ (the average service rate of the two servers), but in the latter case the throughput is bounded by $2\min\{\gamma_1, \gamma_2\}$ (twice the rate of the slower station). Hence, it is possible to benefit more from the differences between the service rates when the tasks are homogeneous, and this explains why the switching region is larger in this case.

5.2 Systems with Homogeneous Tasks and Buffer of Size One

In this section, we provide the optimal assignment policy when the buffer size between the stations is equal to one. The proof of the following theorem is similar to that of Theorem 4.1. It is omitted here in the interest of space, but can be found in Kırkızlar [14].

Theorem 5.2 *For a Markovian tandem line with two stations, two flexible servers, and buffer of size one between the stations, if $\mu_{ij} = \mu_i$ for $i, j \in \{1, 2\}$, then the optimal server assignment policy $\pi = (d)^\infty$ is as follows:*

(i) *If $0 \le c \le \frac{4\mu_1^2\mu_2+5\mu_1\mu_2^2+3\mu_2^3}{12\mu_1^3+20\mu_1^2\mu_2+12\mu_1\mu_2^2+4\mu_2^3}$, then*

$$d^*(x) = \begin{cases} a_{11} & \text{if } x = (0, z_1, z_2) \text{ for all } (z_1, z_2) \in \mathcal{S}_Z \text{ or } x = (1, 1, 1), \\ a_{12} & \text{if } x \in \{(1, 1, 2), (1, 2, 2), (2, 1, 2)\}, \\ a_{21} & \text{if } x \in \{(1, 2, 1), (2, 1, 1), (2, 2, 1)\}, \\ a_{22} & \text{if } x = (3, z_1, z_2) \text{ for all } (z_1, z_2) \in \mathcal{S}_Z \text{ or } x = (2, 2, 2), \end{cases}$$

and the recurrent states are $(0, 1, 2)$, $(0, 2, 1)$, $(1, 1, 1)$, $(1, 1, 2)$, $(1, 2, 1)$, $(1, 2, 2)$, $(2, 1, 1)$, $(2, 1, 2)$, $(2, 2, 1)$, $(2, 2, 2)$, $(3, 1, 2)$, and $(3, 2, 1)$.

(ii) *If $\frac{4\mu_1^2\mu_2+5\mu_1\mu_2^2+3\mu_2^3}{12\mu_1^3+20\mu_1^2\mu_2+12\mu_1\mu_2^2+4\mu_2^3} \le c < \min\left\{\frac{\mu_2}{2\mu_1}, \frac{2\mu_1^4+2\mu_1^3\mu_2+\mu_1^2\mu_2^2-\mu_1\mu_2^3}{2\mu_1^4+4\mu_1^3\mu_2+4\mu_1^2\mu_2^2+4\mu_1\mu_2^3+2\mu_2^4}\right\}$, then $d^*(x) = d_2(x)$ for all $x \in \mathcal{S}$ (see equation (8)) and the recurrent states are $(0, 1, 2)$, $(1, 1, 2)$, $(1, 2, 2)$, $(2, 1, 2)$, $(2, 2, 2)$, and $(3, 1, 2)$.*

(iii) *If $2\mu_1^5 + \mu_1^4\mu_2 \ge \mu_1^3\mu_2^2 + 3\mu_1^2\mu_2^3 + 2\mu_1\mu_2^4 + \mu_2^5$ and $\frac{\mu_2}{2\mu_1} < c \le \frac{3\mu_1^4+2\mu_1^3\mu_2-2\mu_1\mu_2^3}{2\mu_1^4+4\mu_1^3\mu_2+4\mu_1^2\mu_2^2+4\mu_1\mu_2^3+2\mu_2^4}$, then $d^*(x) = d_3(x)$ for all $x \in \mathcal{S}$ (see equation (9)) and the recurrent states are $(0, 1, 2)$, $(0, 2, 2)$, $(1, 1, 2)$, $(1, 2, 2)$, $(2, 1, 2)$, $(2, 2, 2)$, and $(3, 1, 2)$.*

(iv) *If $2\mu_1^5 + \mu_1^4\mu_2 \ge \mu_1^3\mu_2^2 + 3\mu_1^2\mu_2^3 + 2\mu_1\mu_2^4 + \mu_2^5$ and $c > \frac{3\mu_1^4+2\mu_1^3\mu_2-2\mu_1\mu_2^3}{2\mu_1^4+4\mu_1^3\mu_2+4\mu_1^2\mu_2^2+4\mu_1\mu_2^3+2\mu_2^4}$, then $d^*(x) = a_{12}$ for all $x \in \mathcal{S}$ and the recurrent states are $(0, 1, 2)$, $(1, 1, 2)$, $(2, 1, 2)$, and $(3, 1, 2)$.*

(v) *If* $2\mu_1^5 + \mu_1^4\mu_2 < \mu_1^3\mu_2^2 + 3\mu_1^2\mu_2^3 + 2\mu_1\mu_2^4 + \mu_2^5$ *and* $c > \frac{2\mu_1^4+2\mu_1^3\mu_2+\mu_1^2\mu_2^2-\mu_1\mu_2^3}{2\mu_1^4+4\mu_1^3\mu_2+4\mu_1^2\mu_2^2+4\mu_1\mu_2^3+2\mu_2^4}$*, then* $d^*(x) = a_{12}$ *for all* $x \in \mathcal{S}$ *and the recurrent states are* $(0,1,2)$*,* $(1,1,2)$*,* $(2,1,2)$*, and* $(3,1,2)$*.*

Note that these optimal policies are not unique. In particular, the policies where the roles of the servers are switched are also optimal. Moreover, the interval in part (ii) of Theorem 4.2 is non-empty when $\mu_1 \geq \mu_2$, and the interval in part (iii) of Theorem 4.2 is non-empty when $\mu_1 \geq \mu_2$ and $2\mu_1^5 + \mu_1^4\mu_2 \geq \mu_1^3\mu_2^2 + 3\mu_1^2\mu_2^3 + 2\mu_1\mu_2^4 + \mu_2^5$. Note that the intervals considered in the theorem span all values of $c \geq 0$.

As c increases, the optimal policy follows a similar pattern to the optimal policy given in Theorem 4.2 for systems with $B = 1$ and homogeneous servers. More specifically, for small values of c, the optimal policy is of Type 2; for intermediate and high values of c, the optimal policy is of Type 1 and Type 0, respectively. The recurrent states and optimal actions for the optimal policy of Type 2 is depicted in Figure 4. The other cases are omitted because they are same as the ones shown in Figures 2(b), 2(c), and 2(d).

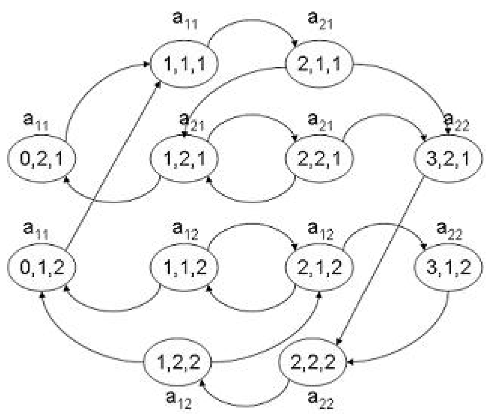


Figure 4: Recurrent States and Optimal Actions for the Type 2 Policy in Theorem 5.2

As in Theorem 4.2, we observe that there are three or four possible optimal policies for different values of the setup cost, depending on if the condition $2\mu_1^5+\mu_1^4\mu_2 \geq \mu_1^3\mu_2^2+3\mu_1^2\mu_2^3+2\mu_1\mu_2^4+\mu_2^5$ is satisfied. Note that this condition implies that the difference between the magnitudes of μ_1 and μ_2 is larger than in the other case, and hence it is advantageous for server 1 to spend more time at station 2 before switching to station 1.

When $B = 1$, we see that for small values of c, there is no preferred assignment of the servers. More specifically, both servers are at the same station when the other station is not operating. The servers remain at that station until it is one transition away from being blocked or starved, at which time the faster server switches to the other station. Furthermore, the slower server also switches between the stations for small values of c, because the increase in throughput resulting from not idling this server dominates the setup cost associated with moving him/her. More specifically, we see that $t_1(1,1) = 2$, $t_1(1,2) = 3$, $t_1(2,1) = 0$, $t_1(2,2) = 1$, $t_2(1,2) = 0$, and $t_2(2,1) = 3$. We have four different thresholds for the first server that depend on the previous locations of both servers.

The Type 2 policy in Theorem 5.2 is different from that in Theorem 4.2 because for small values of the switching cost, it is possible to take advantage of the faster server by switching him/her, whereas it is not possible to take advantage of the easier task because each job has to go through both stations. We also see that the optimal policy is not of Type 2 when $c > \frac{1}{4}$ and is of Type 0 when $c > \frac{3}{2}$. When we let $\mu_1 = 10$ and $\mu_2 = 1$, then we see that the optimal Policy is of Type 0 when $c > 1.309$. Hence, switching policies are optimal for a larger range of setup cost when $B = 1$ and the tasks are homogeneous, as compared to Theorem 4.2 where the servers are homogeneous. This conclusion is similar to the one we made for $B = 0$.

5.3 Systems with Homogeneous Tasks and Multiple Buffer Spaces

Theorems 5.1 and 5.2 provide the optimal server assignment policies for systems with homogeneous tasks when the buffer size between the stations is zero or one. In this section we provide our observations about the form of the optimal policy when $B > 1$.

We randomly generate 50,000 systems with the service rate of each server independently drawn from a uniform distribution with range [0.5,2.5] and the parameters B and c chosen as in Section 4.3. For each system, we determine the optimal server assignment policy using the policy iteration algorithm for weakly communicating Markov chains. We observe that when the optimal policy is of Type 1 or 2, it is a multiple threshold policy. Moreover, as in Theorem 5.2, servers do not have preferred assignments in the Type 2 policy, and the faster server is the switching server in the Type 1 policy. Finally, for large values of c, the optimal policy is of Type 0.

We are able to make some conclusions regarding the threshold values. For simplicity, we only provide our observations for the case where server 1 is the faster server. If the optimal policy is of Type 2, we observe that $t_1(1,1) = l$, $t_1(1,2) = B+2$, $t_1(2,1) = 0$, $t_1(2,2) = k$, $t_2(1,2) = 0$, and $t_2(2,1) = B+2$ for some $l \in \{2, \ldots, B+1\}$, where $k = B+2-l$ by Lemma 3.2, see Figure 5 (note that $l = 2$ and $k = 1 = B+2-l$ in the policy of Figure 4). This special structure of the optimal policy ensures that the faster server helps the slower server for multiple jobs before switching to the other station.

If the optimal policy is of Type 1 and server 1 is the faster server, we see that $t_1(1,2) = B+2$ and $t_1(2,2) = l$, for some $l \in \{0, \ldots, B\}$, see Figure 3(b) (note that $l \in \{0,1\}$ when $B = 1$ as shown in Theorem 5.2). In other words, there are no immediate switchovers. We believe that this results from the fact that one server is better at both tasks, and hence switching the faster server immediately back to its preferred station has a higher impact on the setup costs compared to the throughput. Moreover, this policy maintains a balance between avoiding idling the slower server (i.e., through starvation at the second station) and avoiding setups for the faster server.

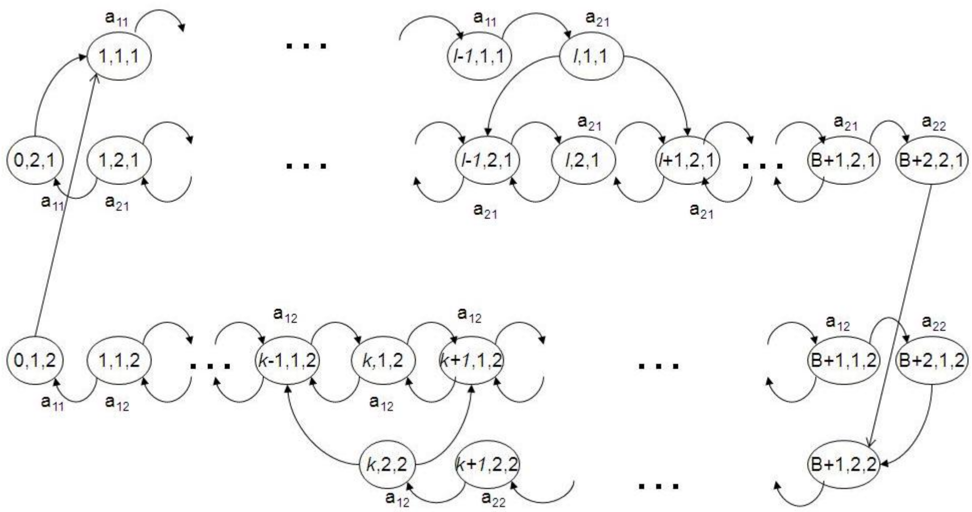


Figure 5: Recurrent States and Optimal Actions for the Type 2 Policy of Section 5.3

We conclude this section by showing that when $\mu_1 \geq \mu_2$ and the optimal policy is of Type 1, $t_1(2,2)$ decreases when c increases. If we relabel the servers, it then follows from Lemma 3.2 that when $\mu_1 < \mu_2$, $t_1(2,2)$ increases with c. Thus the switching server helps the dedicated server with more jobs before switching back to its preferred station for larger setup costs.

Proposition 5.1 *In a tandem line with two stations, two homogeneous tasks (so that $\mu_{ij} = \mu_i$ for $i, j \in \{1,2\}$), and buffer of size $B > 1$ between the stations, the threshold for the optimal Type 1 policy decreases as the setup cost $c > 0$ increases when $\mu_1 \geq \mu_2$.*

Proof: First assume that $\mu_1 > \mu_2$. Let $\pi_l = (d_l)^\infty$ be a Type 1 policy with $t_1(2,2) = l$, where $l \in \{0, \ldots, B\}$. It is not difficult to show that

$$P_{\pi_l} = \frac{(\mu_1+\mu_2)\big((B+2-l-2c)\mu_1^{B+l+4}-(B+2-l-4c)\mu_1^{B+l+3}\mu_2-2c\mu_1^{B+l+2}\mu_2^2-\mu_1^{B+2}\mu_2^{l+2}+\mu_1^l\mu_2^{B+4}\big)}{2(B+2-l)\mu_1^{B+l+4}-2(B+2-l)\mu_1^{B+l+3}\mu_2-\mu_1^{B+3}\mu_2^{l+1}-\mu_1^{B+2}\mu_2^{l+2}+\mu_1^{l+1}\mu_2^{B+3}+\mu_1^l\mu_2^{B+4}}.$$

Some algebra shows that $P_{\pi_l} - P_{\pi_{l+1}} = (\beta_{l,1} + c\beta_{l,2})/\beta_{l,3}$, where

$$\begin{aligned}
\beta_{l,1} &= (\mu_1-\mu_2)^2(\mu_1+\mu_2)\mu_1^{B+l+3}\mu_2\Big((B+1-l)\mu_1^{B+2}\mu_2^l-(B+2-l)\mu_1^{B+1}\mu_2^{l+1}+\mu_1^l\mu_2^{B+2}\Big),\\
\beta_{l,2} &= 2(\mu_1-\mu_2)^3(\mu_1+\mu_2)\mu_1^{2B+l+3}(2\mu_1^{l+2}-\mu_1\mu_2^{l+1}-\mu_2^{l+2}),\\
\beta_{l,3} &= \Big(2(B+2-l)\mu_1^{B+l+4}-2(B+2-l)\mu_1^{B+l+3}\mu_2-\mu_1^{B+3}\mu_2^{l+1}-\mu_1^{B+2}\mu_2^{l+2}+\mu_1^{l+1}\mu_2^{B+3}+\mu_1^l\mu_2^{B+4}\Big)\\
&\quad\times\Big(2(B+1-l)\mu_1^{B+l+4}-2(B+1-l)\mu_1^{B+l+3}\mu_2-\mu_1^{B+2}\mu_2^{l+2}-\mu_1^{B+1}\mu_2^{l+3}+\mu_1^{l+1}\mu_2^{B+3}+\mu_1^l\mu_2^{B+4}\Big).
\end{aligned}$$

The first term in $\beta_{l,3}$ is positive for all $l \in \{0, \ldots, B-1\}$, because it can be rewritten as

$$\begin{aligned}
&2(B+2-l)\mu_1^{B+3+l}(\mu_1-\mu_2) - \mu_1^l\mu_2^{l+1}(\mu_1+\mu_2)(\mu_1^{B+2-l} - \mu_2^{B+2-l}) \\
&= (\mu_1-\mu_2)\Big(2(B+2-l)\mu_1^{B+3+l} - \mu_1^l\mu_2^{B+2}(\mu_1+\mu_2)\sum_{i=0}^{B+1-l}(\frac{\mu_1}{\mu_2})^i\Big) \\
&> 2(\mu_1-\mu_2)\Big((B+2-l)\mu_1^{B+3+l} - \mu_1^{l+1}\mu_2^{B+2}\sum_{i=0}^{B+1-l}(\frac{\mu_1}{\mu_2})^i\Big) > 0,
\end{aligned}$$

where the last inequality follows because $\mu_1^{l+1}\mu_2^{B+2}(\mu_1/\mu_2)^i < \mu_1^{B+3+l}$ for all $i \in \{0, \ldots, B+1-l\}$. Similar calculations show that the second term in $\beta_{l,3}$ is also positive, because it can be rewritten as

$$\begin{aligned}
&2(B+1-l)\mu_1^{B+3+l}(\mu_1-\mu_2) - \mu_1^l\mu_2^{l+2}(\mu_1+\mu_2)(\mu_1^{B+1-l} - \mu_2^{B+1-l}) \\
&> 2(\mu_1-\mu_2)\Big((B+1-l)\mu_1^{B+3+l} - \mu_1^{l+1}\mu_2^{B+2}\sum_{i=0}^{B-l}(\frac{\mu_1}{\mu_2})^i\Big) > 0.
\end{aligned}$$

Thus we have shown that $\beta_{l,3} > 0$ for $l \in \{0, \ldots, B-1\}$. Moreover, $\beta_{l,2} > 0$ trivially. This shows that $P_{\pi_l} > P_{\pi_{l+1}}$ for large enough c.

Next assume that $\mu_1 = \mu_2 = \rho$. Some algebra shows that for $l \in \{0, \ldots, B-1\}$, $P_{\pi_l} - P_{\pi_{l+1}}$ is as in equation (10). Hence, $P_{\pi_l} > P_{\pi_{l+1}}$ for $l \in \{0, \ldots, B-1\}$ and large enough c. This completes the proof. □

6 Systems with Heterogeneous Servers and Tasks

In this section, we consider a tandem line with two stations and two heterogeneous servers. Our state space $\mathcal{S}$ is given in (5). We will consider systems with $B = 0$ and $B > 0$ in Sections 6.1 and 6.2, respectively. Moreover, we consider general setup costs in our numerical results.

6.1 Systems with Heterogeneous Servers and Tasks and No Buffer

In this section we consider a tandem line with two stations, two heterogeneous servers, and tasks that do not necessarily have the same average service requirement. More specifically, we study a system with $B = 0$ and specialist servers, so that the rate $\mu_{ij} \geq 0$ of server i at task j is arbitrary. This allows us, for example, to model situations where one server is more skilled at one task and the other server is more skilled at another task. Note that systems with specialist servers have the most general service rate structure, and hence their analysis is extremely difficult because the Markov decision analysis is highly complex even for systems with $B = 0$.

Without loss of generality, we assume that $\mu_{11}\mu_{22} \geq \mu_{12}\mu_{21}$ because we can relabel the servers otherwise. Under our assumptions on the service rates, this implies that $\mu_{11} > 0$ and

$\mu_{22} > 0$. Let $\xi_i \geq 0$ for $i \in \{1, \ldots, 10\}$ be as defined in Appendix C, and let $\beta_j = \frac{\xi_{2j-1}}{\xi_{2j}}$ for $j \in \{1, \ldots, 5\}$. Let T_1 (T_2) be the throughput of Type 0 policy with action a_{12} (a_{21}) used in each state. Then, we have

$$\begin{aligned} T_1 &= \frac{\mu_{11}\mu_{22}(\mu_{11}+\mu_{22})}{\mu_{11}^2+\mu_{11}\mu_{22}+\mu_{22}^2}, \\ T_2 &= \frac{\mu_{12}\mu_{21}(\mu_{12}+\mu_{21})}{\mu_{12}^2+\mu_{12}\mu_{21}+\mu_{21}^2} \end{aligned}$$

(we use the convention $0/0 = 0$). Let ϕ_1, ϕ_2, and Φ be defined as follows:

$$\begin{aligned} \phi_1 &= \min\left\{\frac{1}{\mu_{11}+\mu_{21}}, \frac{1}{\mu_{12}+\mu_{22}}\right\}, \\ \phi_2 &= \min\left\{\frac{1}{\mu_{11}+\mu_{22}}, \frac{1}{\mu_{12}+\mu_{21}}\right\}, \\ \Phi &= \min\{\phi_1, \phi_2\}. \end{aligned}$$

Let θ_1, θ_2, and Θ be defined as follows:

$$\begin{aligned} &\theta_1 = \max\left\{\frac{\mu_{11}}{\mu_{22}}, \frac{\mu_{22}}{\mu_{11}}\right\}, \\ &\theta_2 = \begin{cases} \max\{\frac{\mu_{12}}{\mu_{21}}, \frac{\mu_{21}}{\mu_{12}}\} & \text{if } \mu_{12} > 0 \text{ and } \mu_{21} > 0, \\ 0 & \text{otherwise,} \end{cases} \\ &\Theta = \max\{\theta_1, \theta_2\}. \end{aligned}$$

The theorem below characterizes the optimal server assignment policy for systems with $B = 0$ and either small or large switching costs; the sketch of its proof is provided in Appendix C.

Theorem 6.1 *For a Markovian tandem line with two stations, two flexible servers whose service rates satisfy* $\mu_{11}\mu_{22} \geq \mu_{12}\mu_{21}$*, and buffer of size zero between the stations, the optimal server assignment policy* $\pi^* = (d^*)^\infty$ *satisfies the following:*

(i) *If either* $0 \leq c \leq \min\{\beta_1, \beta_2, \beta_5\}$ *or* $\beta_5 \leq c \leq \min\{\beta_1, \beta_2, \beta_3, \beta_4\}$ *(whenever this interval is nonempty), then* $d^*(x) = d_0(x)$ *for all* $x \in \mathcal{S}$ *(see equation (6)) and the recurrent states are* $(0,1,2), (1,1,1), (1,2,2),$ *and* $(2,1,2)$.

(ii) *If* $T_1 \geq T_2$ *and* $c > \Big(1 - T_1(\Phi + \phi_1)\Big)(1 + \Theta)$*, then* $d^*(x) = a_{12}$ *for all* $x \in \mathcal{S}$ *and the recurrent states are* $(0,1,2), (1,1,2),$ *and* $(2,1,2)$.

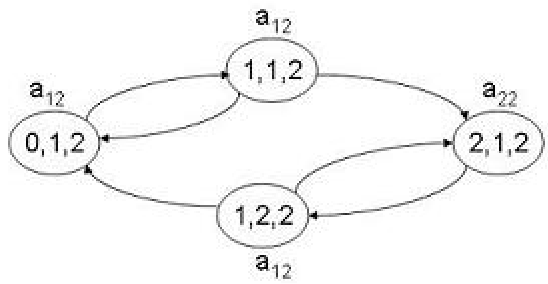
(iii) *If* $T_1 < T_2$ *and* $c > \Big(1 - T_2(\Phi + \phi_1)\Big)(1 + \Theta)$*, then* $d^*(x) = a_{21}$ *for all* $x \in \mathcal{S}$ *and the recurrent states are* $(0,2,1), (1,2,1),$ *and* $(2,2,1)$.

Similar to Theorems 4.1 and 5.1, we observe that the optimal server assignment policy is a Type 2 policy for small values of c, and it is a Type 0 policy for large values of c. Moreover, the optimal policy for small c agrees with the one for systems with $c = 0$ (as shown by Andradóttir, Ayhan, and Down [1]). Note that in parts (ii) and (iii) of Theorem 6.1, no setup cost is incurred because the servers are dedicated, and hence the optimal server assignment policy is determined by comparing the throughputs under the different dedicated assignments. The lower bounds on the setup cost in parts (ii) and (iii) of Theorem 6.1 are found by comparing the maximum possible profit under switching policies and the throughput of the best dedicated policy. Note that $1 - T_i(\Phi + \phi_1) \geq 0$ for $i \in \{1, 2\}$.

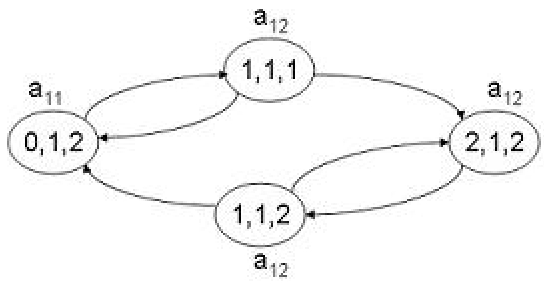
Next we provide our observations about the form of the optimal policy when the servers are specialists and the setup cost is not necessarily small or large. We randomly generate 100,000 systems with the service rates independently drawn from a uniform distribution with range $[0.5, 2.5]$, and the setup costs $c_i(1, 2)$ and $c_i(2, 1)$ for $i \in \{1, 2\}$ drawn independently from a uniform distribution with range (0,0.5). We also did the same experiment with $c = c_i(1, 2) = c_i(2, 1)$ drawn from a uniform distribution with range (0,0.5) and obtained similar results. For each system, we determine the optimal server assignment policy using the policy iteration algorithm for weakly communicating Markov chains. Note that the number of replications in this section (i.e., 100,000) is larger than for our other experiments (i.e., 50,000) because we only consider systems with $B = 0$.

As in Sections 4.1 and 4.2, we observe that both servers have preferred assignments in each experiment and the optimal policy is a multiple threshold policy. We now demonstrate these policies in more detail for a system where servers 1 and 2 have preferred assignments at stations 1 and 2, respectively. If the optimal policy is of Type 2, then we have $t_1(1, 2) = 2$, $t_1(2, 2) = 1$, $t_2(1, 1) = 1$, and $t_2(1, 2) = 0$, as in Figure 1(a). If the optimal policy is of Type 1 and server 1 is the switching server, then we have $t_1(1, 2) = 2$ and $t_1(2, 2) \in \{0, 1\}$, as in Figures 1(b) and 6(a). If the optimal policy is of Type 1 and server 2 is the switching server, we have $t_2(1, 2) = 0$ and $t_2(1, 1) \in \{1, 2\}$, as in Figures 6(b) and 6(c). Finally, if the setup cost is big, the optimal policy is of Type 0, as in Figure 1(c).

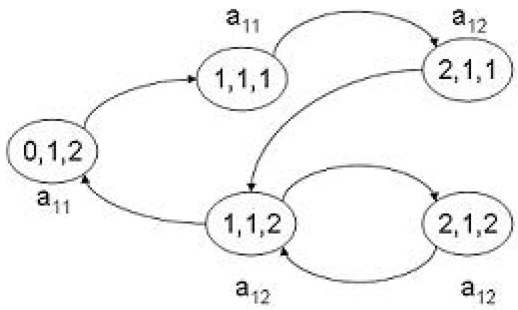
Policies of Type 0, 1, or 2 were also observed for systems with $B = 0$ and homogeneous servers or tasks, as shown in Sections 4.1 and 5.1. However, for these systems, if the optimal policy is of Type 1 and the switching server is server $i \in \{1, 2\}$, then $t_i(2, 2)$ or $t_i(1, 1)$ is never equal to one; i.e., the policies shown in Figures 6(a) and 6(b) are never optimal. Note that when the servers are specialists, it may be best to have immediate switchovers in order to exploit the special skills of the servers, although this increases the number of setups. Hence, we conclude that the form of the policy is robust to the service rates, but the thresholds can take a broader range of values in systems with specialist servers.



(a) Only Server 1 Switches, $t_1(2,2) = 1$ (Type 1)



(b) Only Server 2 Switches, $t_2(1,1) = 1$ (Type 1)



(c) Only Server 2 Switches, $t_2(1,1) = 2$ (Type 1)

Figure 6: Recurrent States and Optimal Actions in Section 6.1

6.2 Systems with Heterogeneous Servers and Tasks and Positive Buffer Size

In this section, we study systems with $B > 0$ and heterogeneous servers. We start by considering generalist servers, so that the service rate of a server at a station can be represented as the product of the server's speed and a constant reflecting the complexity level of the task at the station. Hence, we have $\mu_{ij} = \mu_i \gamma_j$ for $i, j \in \{1, 2\}$. Service rates of this form can be used to model situations where each server is equally skilled at all tasks but the tasks are of different difficulty levels. Note that systems with generalist servers are special cases of systems with specialist servers, hence their analysis is simpler and we are able to partially characterize the optimal policy for systems with arbitrary, positive buffer sizes and (constant) setup costs when the servers are generalists. Moreover, systems with either homogeneous servers or tasks are special cases of systems with generalist servers; hence this service rate structure is significantly more difficult to analyze.

We observed in Theorems 4.1 and 5.1 that the policy π_0 (as defined in Example 2.1) is optimal in systems with $B = 0$ when c is positive but small and either the servers or tasks are homogeneous. However, when $B > 0$, the optimal policies provided in Theorems 4.2 and 5.2 are different from π_0 even for small values of $c > 0$. Hence, the servers have a primary assignment when $B = 0$ and $c > 0$ is small, but this is not correct when $B > 0$. The following proposition shows that when $B > 0$, $c > 0$, and the servers are generalists, the policy π_0 (that has immediate switchovers) is dominated by policies that have preferred assignments, rather than primary assignments, including a policy that has no immediate switchovers. The fact that the optimal policy for $c = 0$ is suboptimal for all $c > 0$ when the servers are generalists shows that the best assignment of servers to tasks is highly sensitive to the presence of setups.

Proposition 6.1 *In a tandem line with two stations, two generalist servers, and a buffer of size $B > 0$ between the stations, the policy π_0 is not optimal when $c > 0$ because it is better to delay the switchovers at either end of the line, and better yet to delay the switchovers at both ends of the line.*

Proof: First, assume that $\mu_1\gamma_1 \neq \mu_2\gamma_2$. Let $\pi_0 = (d_0)^\infty$ be as described in Example 2.1. It is not difficult to show that

$$P_{\pi_0} = \frac{(\mu_1+\mu_2)\gamma_1\gamma_2}{\gamma_1+\gamma_2} - \frac{2c(\mu_1+\mu_2)\gamma_1\gamma_2(\mu_1\gamma_1-\mu_2\gamma_2)\Big((\mu_1\gamma_1)^{B+1}+(\mu_2\gamma_2)^{B+1}\Big)}{(\gamma_1+\gamma_2)\Big((\mu_1\gamma_1)^{B+2}-(\mu_2\gamma_2)^{B+2}\Big)}.$$

Now define the policies $\hat{\pi} = (\hat{d})^\infty$ and $\bar{\pi} = (\bar{d})^\infty$ such that $\hat{d}(1,1,1) = \bar{d}(1,1,1) = a_{11}$, $\bar{d}(B+1,2,2) = a_{22}$, $\hat{d}(x) = d_0(x)$ for $x \in \mathcal{S} \setminus \{(1,1,1)\}$, and $\bar{d}(x) = d_0(x)$ for $x \in \mathcal{S} \setminus \{(1,1,1),(B+1,2,2)\}$. In other words, $\hat{\pi}$ is a multiple threshold policy that delays switchovers at the beginning of the line, and $\bar{\pi}$ is a multiple threshold policy that delays switchovers at both ends of the line. One can show that

$$\begin{aligned}
P_{\hat{\pi}} &= \frac{(\mu_1+\mu_2)\gamma_1\gamma_2}{\gamma_1+\gamma_2} - \frac{2c(\mu_1+\mu_2)\gamma_1\gamma_2(\mu_1\gamma_1-\mu_2\gamma_2)\Big((\mu_1\gamma_1)^{B+1}+\mu_2\gamma_2\Big((\mu_1\gamma_1)^{B}+(\mu_2\gamma_2)^{B}\Big)\Big)}{(\gamma_1+\gamma_2)\Big((\mu_1\gamma_1)^{B+2}+\mu_2\gamma_2(\mu_1\gamma_1)^{B+1}-2(\mu_2\gamma_2)^{B+2}\Big)}, \\
P_{\bar{\pi}} &= \frac{(\mu_1+\mu_2)\gamma_1\gamma_2}{\gamma_1+\gamma_2} - \frac{c(\mu_1+\mu_2)\gamma_1\gamma_2(\mu_1\gamma_1-\mu_2\gamma_2)\Big((\mu_1\gamma_1)^{B}+(\mu_2\gamma_2)^{B}\Big)}{(\gamma_1+\gamma_2)\Big((\mu_1\gamma_1)^{B+1}-(\mu_2\gamma_2)^{B+1}\Big)}.
\end{aligned}$$

Some algebra shows that $P_{\hat{\pi}} - P_{\pi_0} = \frac{\epsilon_1}{\epsilon_2}$ and $P_{\bar{\pi}} - P_{\hat{\pi}} = \frac{\epsilon_3}{\epsilon_4}$, where

$$\begin{aligned}
\epsilon_1 &= 2c(\mu_1+\mu_2)\gamma_1\gamma_2(\mu_1\gamma_1-\mu_2\gamma_2)(\mu_2\gamma_2)^{B+3}\Big((\mu_1\gamma_1)^{B}-(\mu_2\gamma_2)^{B}\Big), \\
\epsilon_2 &= (\gamma_1+\gamma_2)\Big((\mu_1\gamma_1)^{B+2}-(\mu_2\gamma_2)^{B+2}\Big)\Big[\Big((\mu_1\gamma_1)^{B+2}-(\mu_2\gamma_2)^{B+2}\Big) \\
&\quad +\mu_2\gamma_2\Big((\mu_1\gamma_1)^{B+1}-(\mu_2\gamma_2)^{B+1}\Big)\Big], \\
\epsilon_3 &= c(\mu_1+\mu_2)\gamma_1\gamma_2(\mu_1\gamma_1-\mu_2\gamma_2)(\mu_1\gamma_1+\mu_2\gamma_2)(\mu_1\gamma_1)^{B+1}\Big((\mu_1\gamma_1)^{B}-(\mu_2\gamma_2)^{B}\Big), \\
\epsilon_4 &= (\gamma_1+\gamma_2)\Big((\mu_1\gamma_1)^{B+1}-(\mu_2\gamma_2)^{B+1}\Big)\Big((\mu_1\gamma_1)^{B+2}+\mu_2\gamma_2(\mu_1\gamma_1)^{B+1}-2(\mu_2\gamma_2)^{B+2}\Big)\Big).
\end{aligned}$$

It is easy to see that $\frac{\epsilon_1}{\epsilon_2} > 0$ and $\frac{\epsilon_3}{\epsilon_4} > 0$. Hence, $\hat{\pi}$ is a better policy than π_0 and $\bar{\pi}$ is a better policy than $\hat{\pi}$. Note that if $\tilde{d}$ is such that $\tilde{d}(B+1,2,2) = a_{22}$ and $\tilde{d}(x) = d(x)$ for $x \in \mathcal{S} \setminus \{(B+1,2,2)\}$, then the proof of Lemma 3.2 and the above calculations imply that the long-run average profit under policy $\tilde{\pi}$ satisfies $P_{\pi_0} < P_{\tilde{\pi}} < P_{\bar{\pi}}$, and hence $\tilde{\pi}$ is superior to π_0 but inferior to $\bar{\pi}$.

When $\mu_1\gamma_1 = \mu_2\gamma_2$, we can show that

$$P_{\pi_0} = \frac{\mu_1\gamma_1(2 + B - 4c)}{2 + B}, P_{\hat{\pi}} = P_{\tilde{\pi}} = \frac{\mu_1\gamma_1(3 + 2B - 6c)}{3 + 2B}, P_{\bar{\pi}} = \frac{\mu_1\gamma_1(1 + B - 2c)}{1 + B}.$$

Then $P_{\hat{\pi}} - P_{\pi_0} = \frac{2cB\mu_1\gamma_1}{6+7B+2B^2}$ and $P_{\bar{\pi}} - P_{\hat{\pi}} = \frac{2cB\mu_1\gamma_1}{3+5B+2B^2}$. Note that these quantities are strictly positive for $B > 0$. Consequently, when $c > 0$, the policies π_0, $\hat{\pi}$, and $\tilde{\pi}$ are never optimal. □

Proposition 6.1 shows that, as expected, not all nonidling policies are optimal for systems with generalist servers in the presence of small positive setup costs (unlike the case with $c = 0$ as shown in Andradóttir, Ayhan, and Down [1]). Note that Proposition 6.1 does not necessarily hold for systems with specialist servers. In particular, we will now show that the policy π_0 may be optimal for $B > 0$ and positive setup costs if the servers are specialists.

More specifically, we study systems with specialist servers, $B > 0$, and nonconstant setup costs. We randomly generate 50,000 systems with the service rates chosen as in Section 6.1, the buffer size drawn from a discrete uniform distribution with range $\{1, 2, 3, 4, 5\}$, and the setup costs chosen as in Section 6.1 (as in Section 6.1, we considered both constant setup costs c and general setup costs $c_i(1, 2)$, $c_i(2, 1)$ for $i \in \{1, 2\}$). For each system, we determine the optimal server assignment policy using the policy iteration algorithm for weakly communicating Markov chains.

In all the experiments, we observe that when the optimal policy is of Type 1 or 2, it is a multiple threshold policy. For example, consider the case where the optimal policy is of Type 2 with servers 1 and 2 having preferred assignments at stations 1 and 2, respectively. Then, we observe that $t_1(1, 2) = B + 2$, $t_1(2, 2) = k$, $t_2(1, 1) = l$, and $t_2(1, 2) = 0$, where $k, l \in \{1, \ldots, B + 1\}$. However, unlike in Section 4.3 (see Figure 3(a)), we do not observe any simple relation between the thresholds k and l. Note that we observed some cases where $l = 1$ or $k = B + 1$ in the optimal policy (i.e., π_0 was optimal). Hence, in the presence of specialist servers, the policy that has primary assignments (so that the servers switch back to their primary stations as soon as possible) can be optimal, unlike in the case with generalist servers considered in Proposition 6.1. We believe that this results from the fact that in the case of specialist servers, each server may outperform the other server at one of the tasks, and it is possible to take advantage of this difference in skills by allowing immediate switchovers.

When the optimal policy is of Type 2 and the servers do not have preferred assignments (as in Figures 4 and 5), we also do not observe any special structure in the thresholds (unlike in Section 5.3). More specifically, when server 1 is the server that switches more often, we observe that $t_1(1, 1) = k$, $t_1(1, 2) = B + 2$, $t_1(2, 1) = 0$, $t_2(2, 2) = l$, $t_2(1, 2) = 0$, and $t_2(2, 1) = B + 2$, where $k, l \in \{1, \ldots, B + 1\}$, and thus immediate switchovers are possible. Similarly, when the optimal policy is of Type 1, it is a multiple threshold policy that allows immediate switchovers (unlike the cases when the tasks or servers are homogeneous). For example, if server 1 is the switching server and (s)he has a preferred assignment at station 1, we see that $t_1(1,2) = B+2$ and $t_1(2,2) = l$, where $l \in \{0, \dots, B+1\}$ (as in Figure 3(a), except the optimal action in state $(B+1, 2, 2)$ can now be a_{12}).

The results described in this section strongly support the conjecture that the optimal policy for systems with two stations, two servers, and positive setup costs has a multiple threshold structure. More specifically, we observe that the form of the optimal policy remains the same even when the buffer size is increased, the servers have different skills at different stations, and the setup cost depends on both the server and the location of the server.

7 Conclusion

In this work, we have studied the dynamic server assignment problem in the presence of setup costs. More specifically, we have determined the optimal server assignment policy for tandem systems with two stations, two servers, constant setup costs, and small buffer sizes when either the servers or the tasks are homogeneous. We have shown that the optimal policy is of "multiple threshold" type (i.e., servers move between stations when the number of jobs in the system reaches certain thresholds that may depend on the current locations of servers). Moreover, the servers generally have preferred assignments (the only exception is when the tasks are homogeneous and the value of the setup cost is small, in which case we can take advantage of the faster server by assigning him/her to both stations). As the value of the setup cost increases, the optimal server assignment policy reduces the number of servers that move between the stations, and when there is only one switching server in the system, we have seen that the faster server or the server that is assigned to the faster station is the switching server. We have also shown that server movements are more limited when the servers are homogeneous than when the tasks are homogeneous. Finally, we have identified additional structure for the thresholds in systems with either homogeneous servers or tasks and arbitrary buffer sizes based on theoretical and numerical results.

We have also characterized the optimal server assignment policy for systems with two stations, two specialist servers (i.e., the service rates can be arbitrary), and zero buffers when the constant setup cost is small or large. Moreover, we have shown that for systems with two stations, two generalist servers, and positive buffer sizes, the optimal policy for zero setup costs (in which the servers have primary assignments and only switch to avoid idleness) is no longer optimal, even for small values of the setup cost. This discontinuity is surprising at first, but can be explained by the fact that it is possible to reduce the incurred setup cost without impacting server effectiveness by having generalist servers complete several jobs at their non-preferred stations before returning to their preferred stations. By contrast, when the servers are specialists, it may be optimal to have servers return immediately to their preferred stations in order to take advantage of their special skills. Finally, we have performed numerical experiments for systems with specialist servers, larger buffer sizes, and setup costs that are not necessarily constant that suggest that the optimal policy also has a multiple threshold structure in this setting. Consequently, the form of the optimal policy appears to be robust to the service rates, setup costs, and buffer sizes.

This work shows that the optimal server assignment policy is highly sensitive to the presence of even a small setup cost. The inclusion of positive setup times or other costs (e.g., holding costs) in finite-buffered systems with flexible servers remains an open problem.

Acknowledgments

This research was partly supported by the National Science Foundation under grant CMMI-0856600. The first author was also supported by the National Science Foundation under grant CMMI-0400260.

Appendices

A Additional Results for Systems with Homogeneous Servers

In this section we provide the optimal server assignment policy for systems with homogeneous servers and $B \in \{0, 1\}$ when $\gamma_1 < \gamma_2$. The following result follows from Theorem 4.1 and Lemma 3.2.

Corollary A.1 *For a Markovian tandem line with two stations, two flexible servers, and buffer of size zero between the stations, if $\mu_{ij} = \gamma_j$ for $i, j \in \{1, 2\}$ and $\gamma_1 < \gamma_2$, then the optimal server assignment policy $\pi^* = (d^*)^\infty$ is as follows:*

(i) *If $0 \le c \le \frac{\gamma_1}{4\gamma_1+2\gamma_2}$, then $d^*(x) = d_0(x)$ for all $x \in \mathcal{S}$ (see equation (6)) and the recurrent states are $(0,1,2), (1,1,1), (1,2,2)$, and $(2,1,2)$.*

(ii) *If $\frac{\gamma_1}{4\gamma_1+2\gamma_2} < c \le \frac{\gamma_2^2}{2\gamma_1^2+2\gamma_1\gamma_2+2\gamma_2^2}$, then*

$$d^*(x) = \begin{cases} a_{11} & \text{if } x = (0, z_1, z_2) \text{ for all } (z_1, z_2) \in \mathcal{S}_Z \text{ or } x = (1,1,1), \\ a_{12} & \text{if } x = (1, z_1, z_2) \text{ for all } (z_1, z_2) \in \mathcal{S}_Z \setminus \{(1,1)\} \text{ or} \\ & \quad x = (2, z_1, z_2) \text{ for all } (z_1, z_2) \in \mathcal{S}_Z, \end{cases}$$

and the recurrent states are $(0,1,2),(1,1,1),(1,1,2),(2,1,1),$ *and* $(2,1,2)$.

(iii) *If* $c > \frac{\gamma_2^2}{2\gamma_1^2+2\gamma_1\gamma_2+2\gamma_2^2}$, *then* $d^*(x) = a_{12}$ *for all* $x \in \mathcal{S}$ *and the recurrent states are* $(0,1,2),(1,1,2),$ *and* $(2,1,2)$.

The following result follows from Theorem 4.2 and Lemma 3.2.

Corollary A.2 *For a Markovian tandem line with two stations, two flexible servers, and buffer of size one between the stations, if* $\mu_{ij} = \gamma_j$ *for* $i, j \in \{1,2\}$ *and* $\gamma_1 < \gamma_2$, *then the optimal server assignment policy* $\pi^* = (d^*)^\infty$ *is as follows:*

(i) *If* $0 \le c \le \frac{\gamma_1}{2\gamma_1+2\gamma_2}$, *then*

$$d^*(x) = \begin{cases} a_{11} & \text{if } x = (0,z_1,z_2) \text{ for all } (z_1,z_2) \in \mathcal{S}_Z \text{ or } x = (1,1,1), \\ a_{12} & \text{if } x = (1,z_1,z_2) \text{ for all } (z_1,z_2) \in \mathcal{S}_Z \setminus \{(1,1,1)\} \text{ or} \\ & \quad x = (2,z_2,z_2) \text{ for all } (z_1,z_2) \in \mathcal{S}_Z \setminus \{(2,2,2)\}, \\ a_{22} & \text{if } x = (3,z_1,z_2) \text{ for all } (z_1,z_2) \in \mathcal{S}_Z \text{ or } x = (2,2,2), \end{cases}$$

and the recurrent states are $(0,1,2)$, $(1,1,1)$, $(1,1,2)$, $(1,2,2)$, $(2,1,1)$, $(2,1,2)$, $(2,2,2)$, *and* $(3,1,2)$.

(ii) *If* $\frac{\gamma_1}{2\gamma_1+2\gamma_2} < c \le \min\{\frac{\gamma_2^2}{2\gamma_1^2+2\gamma_2^2}, \frac{2\gamma_1\gamma_2+\gamma_1^2}{2\gamma_2^2+4\gamma_1\gamma_2}\}$, *then*

$$d^*(x) = \begin{cases} a_{11} & \text{if } x = (0,z_1,z_2) \text{ for all } (z_1,z_2) \in \mathcal{S}_Z \text{ or } x = (1,1,1), \\ a_{12} & \text{if } x = (1,z_1,z_2) \text{ for all } (z_1,z_2) \in \mathcal{S}_Z \setminus \{(1,1)\} \text{ or} \\ & \quad x = (y,z_1,z_2) \text{ for all } y \in \{2,3\} \text{ and } (z_1,z_2) \in \mathcal{S}_Z, \end{cases}$$

and the recurrent states are $(0,1,2)$, $(1,1,1)$, $(1,1,2)$, $(2,1,1)$, $(2,1,2)$, *and* $(3,1,2)$.

(iii) *If* $\gamma_2^2 > \gamma_1\gamma_2 + \gamma_1^2$ *and* $\frac{2\gamma_1\gamma_2+\gamma_1^2}{2\gamma_2^2+4\gamma_1\gamma_2} < c \le \frac{3\gamma_2^3+\gamma_1\gamma_2^2-\gamma_1^2\gamma_2}{4\gamma_1^3+4\gamma_1^2\gamma_2+4\gamma_1\gamma_2^2+4\gamma_2^3}$, *then*

$$d^*(x) = \begin{cases} a_{11} & \text{if } x = (0,z_1,z_2) \text{ for all } (z_1,z_2) \in \mathcal{S}_Z \text{ or} \\ & \quad x = (y,1,1) \text{ for all } y \in \{1,2\}, \\ a_{12} & \text{if } x = (y,z_1,z_2) \text{ for all } y \in \{1,2\} \text{ and } (z_1,z_2) \in \mathcal{S}_Z \setminus \{(1,1)\} \text{ or} \\ & \quad x = (3,z_1,z_2) \text{ for all } (z_1,z_2) \in \mathcal{S}_Z\ , \end{cases}$$

and the recurrent states are $(0,1,2)$, $(1,1,1)$, $(1,1,2)$, $(2,1,1)$, $(2,1,2)$, $(3,1,1)$, *and* $(3,1,2)$.

(iv) *If* $\gamma_2^2 > \gamma_1\gamma_2 + \gamma_1^2$ *and* $c > \frac{3\gamma_2^3+\gamma_1\gamma_2^2-\gamma_1^2\gamma_2}{4\gamma_1^3+4\gamma_1^2\gamma_2+4\gamma_1\gamma_2^2+4\gamma_2^3}$, *then* $d^*(x) = a_{12}$ *for all* $x \in \mathcal{S}$ *and the recurrent states are* $(0,1,2)$, $(1,1,2)$, $(2,1,2)$, *and* $(3,1,2)$.

(v) *If* $\gamma_2^2 \le \gamma_1\gamma_2 + \gamma_1^2$ *and* $c > \frac{\gamma_2^2}{2\gamma_1^2+2\gamma_2^2}$, *then* $d^*(x) = a_{12}$ *for all* $x \in \mathcal{S}$ *and the recurrent states are* $(0,1,2)$, $(1,1,2)$, $(2,1,2)$, *and* $(3,1,2)$.

B Proof of Theorem 4.1

Lemma 3.1 shows that servers should not be voluntarily idle when station 1 is blocked or station 2 is starved (this is different from involuntary idling due to being assigned to a station that is either blocked or starved). Furthermore, when both stations are operating, if a server is at station $j \in \{1, 2\}$ before the previous server completion, any action that idles this server and assigns the other server to station j cannot be optimal. For example, actions a_{01} and a_{20} cannot be optimal in a state $(l, 1, 2)$, where $1 \leq l \leq B + 1$, because they are strictly dominated by actions a_{11} and a_{22}, respectively (this can be shown through a sample path argument similar to that in the proof of Lemma 3.1). Moreover, the action a_{00} results in a zero long-run average profit if employed in any state, and hence is ignored. Similarly, a_{22} is never optimal in a state $(0, z)$ and a_{11} is never optimal in a state $(2, z)$, for $z \in \mathcal{S}_Z$. The states $(0, 1, 1)$ and $(2, 2, 2)$ are transient under any policy $\pi \in \Pi$, and the actions in these states do not affect the long-run average profit. Hence, they are omitted in the proof because any feasible action can be chosen in these states. Thus, we can use the following action space:

$$A_x = \begin{cases} \{a_{11}, a_{12}, a_{21}\} & \text{for } x \in \{(0,1,2), (0,2,1), (0,2,2)\}, \\ \{a_{02}, a_{11}, a_{12}, a_{20}, a_{21}, a_{22}\} & \text{for } x = (1,1,1), \\ \{a_{02}, a_{10}, a_{11}, a_{12}, a_{21}, a_{22}\} & \text{for } x = (1,1,2), \\ \{a_{01}, a_{11}, a_{12}, a_{20}, a_{21}, a_{22}\} & \text{for } x = (1,2,1), \\ \{a_{01}, a_{10}, a_{11}, a_{12}, a_{21}, a_{22}\} & \text{for } x = (1,2,2), \\ \{a_{12}, a_{21}, a_{22}\} & \text{for } x \in \{(2,1,1), (2,1,2), (2,2,1)\}. \end{cases}$$

Since the action and state spaces are finite, Theorem 9.1.8 of Puterman [19] shows the existence of an optimal Markovian stationary deterministic policy. Furthermore, since $\gamma_1, \gamma_2 > 0$, the policies described in the theorem correspond to weakly communicating Markov chains, and we can use the Linear Program (LP) approach for communicating Markov decision processes as in Sections 9.5.2 and 8.8.2 of Puterman [19].

Consider the following LP:

$$\left.\begin{array}{ll} \max & \sum_{x \in \mathcal{S}} \sum_{a \in A_x} r(x,a)\omega(x,a) \\ \text{s.t.} & \sum_{a \in A_{x'}} \omega(x',a) - \sum_{x \in \mathcal{S}} \sum_{a \in A_x} p(x'|x,a)\omega(x,a) = 0, \text{ for all } x' \in \mathcal{S}, \\ & \sum_{x \in \mathcal{S}} \sum_{a \in A_x} \omega(x,a) = 1, \\ & \omega(x,a) \geq 0, \text{ for all } x \in \mathcal{S}, a \in A_x, \end{array}\right\} \quad (12)$$

where, for all $x \in \mathcal{S}$ and $a \in A_x$, $r(x, a)$ is the immediate reward of choosing action a in state x and $p(x'|x, a)$ is the one-step transition probability from state x to x' if action a is chosen in state x. Then, in every basic feasible solution corresponding to a policy described in the theorem, we can conclude that for each $x \in \mathcal{S}$ there exists at most a single action $a_x \in A_x$ such that $\omega(x, a_x) > 0$ as a result of Corollary 8.8.7 of Puterman [19] (which can be applied because the policies we consider in the description of the theorem result in a single recurrent class). Furthermore, for every basic feasible optimal solution w^* if we define $\mathcal{S}_{w^*} = \{x \in \mathcal{S} : \sum_{a \in A_x} w^*(x, a) > 0\}$, then the optimal decision rule is as follows:

$$d_{w^*}(x) = \begin{cases} a & \text{if } w^*(x,a) > 0 \text{ for } x \in \mathcal{S}_{w^*}, \\ a' & \text{for some } a' \text{ such that there exists a state } x' \in \mathcal{S}_{w^*} \text{ for which} \\ & x' \text{ is reachable from } x \text{ under action } a' \text{ for } x \in \mathcal{S} \setminus \mathcal{S}_{w^*}. \end{cases}$$

Note that the actions in states $\mathcal{S} \setminus \mathcal{S}_{w^*}$ cannot be chosen arbitrarily as in unichain models. However, the discussion in Section 9.5.2 of Puterman [19] shows that the decision rule above results in an optimal solution. Moreover, note that an action a' that will move the process X'_π towards a recurrent state always exist. More specifically, if $x = (y, z)$ and $x' = (y', z') \in \mathcal{S}_{w^*}$, we can choose $a' = a_z$ if $y = y'$, $a' = a_{11}$ if $y < y'$, and $a' = a_{22}$ if $y > y'$.

We first prove the optimality of the policy for $0 \leq c \leq \frac{\gamma_2}{2\gamma_1 + 4\gamma_2}$ (note that this condition implies that $c \leq \frac{1}{2}$). Consider the decision rule d, where $d(x)$ is defined as follows for all $x \in \mathcal{S}$:

$$d(x) = \begin{cases} a_{11} & \text{if } x \in \{(0,1,2), (0,2,1), (0,2,2)\}, \\ a_{12} & \text{if } x \in \{(1,1,1), (1,1,2), (1,2,2)\}, \\ a_{21} & \text{if } x \in \{(1,2,1)\}, \\ a_{22} & \text{if } x \in \{(2,1,1), (2,1,2), (2,2,1)\}. \end{cases}$$

Now, consider the basic solution ω of the LP (12) corresponding to the policy $\pi = (d)^\infty$. The associated basis for the LP (12) is

$$\begin{aligned} D \;=\; & \{\omega((0,1,2), a_{11}), \omega((0,2,1), a_{11}), \omega((0,2,2), a_{11}), \\ & \omega((1,1,1), a_{12}), \omega((1,1,2), a_{12}), \omega((1,2,1), a_{21}), \omega((1,2,2), a_{12}), \\ & \omega((2,1,1), a_{22}), \omega((2,1,2), a_{22}), \omega((2,2,1), a_{22})\}. \end{aligned}$$

Let c_B be the vector of coefficients of the elements of D in the objective function, $\mathbf{B}$ be the coefficients of the elements of D in the constraint matrix, and b be the right-hand side of the constraints. Consequently, we have

$$\begin{aligned} c_B \;=\; & \{-2c\gamma_1, -2c\gamma_1, -4c\gamma_1, \gamma_2 - c(\gamma_1 + \gamma_2), \gamma_2, \gamma_2, \gamma_2 - c(\gamma_1 + \gamma_2), \\ & 2\gamma_2(1 - 2c), 2\gamma_2(1 - c), 2\gamma_2(1 - c)\}, \end{aligned}$$

and

$$\mathbf{B} = \begin{bmatrix} 2\gamma_1/q & 0 & 0 & -\gamma_2/q & \dots & 0 & 0 \\ 0 & 2\gamma_1/q & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 2\gamma_1/q & 0 & \dots & 0 & 0 \\ -2\gamma_1/q & -2\gamma_1/q & -2\gamma_1/q & (\gamma_1+\gamma_2)/q & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & -2\gamma_2/q & -2\gamma_2/q \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 2\gamma_2/q & 0 \\ 1 & 1 & 1 & 1 & \dots & 1 & 1 \end{bmatrix},$$

where q is the uniformization constant. Note that the constraint corresponding to one of the states is redundant, and hence the constraint corresponding to state $(2, 2, 1)$ is eliminated. It is easy to see that ω is also a stationary distribution for the Markov Chain X_π (since it has a finite state space and one recurrent class, the stationary distribution exists). In order to show the optimality of this basic feasible solution, we need only to show that

$$\Delta_y = c_B\mathbf{B}^{-1}v_y - c_y \geq 0 \tag{13}$$

for each nonbasic variable y, where v_y is the column in the constraint matrix of the LP (12) and c_y is the coefficient corresponding to y in the objective function (see, e.g., Theorem 3.1 of Bertsimas and Tsitsiklis [6]).

For states $(0, 1, 2), (0, 2, 1)$, and $(0, 2, 2)$, we have

$$\Delta_{w((0,1,2),a_{12})} = \Delta_{w((0,2,1),a_{21})} = \Delta_{w((0,2,2),a_{12})} = \Delta_{w((0,2,2),a_{21})} = \frac{\gamma_1(\gamma_2 - 2c\gamma_1 - 4c\gamma_2)}{\gamma_1+\gamma_2},$$
$$\Delta_{w((0,1,2),a_{21})} = \Delta_{w((0,2,1),a_{12})} = \frac{\gamma_1\gamma_2(1-2c)}{\gamma_1+\gamma_2}.$$

It is clear that these quantities are nonnegative when $0 \leq c \leq \frac{\gamma_2}{2\gamma_1+4\gamma_2}$. For state $(1, 1, 1)$ we have

$$\Delta_{w((1,1,1),a_{02})} = \Delta_{w((1,1,1),a_{20})} = \frac{\gamma_1\gamma_2(1-2c)}{\gamma_1+\gamma_2}, \quad \Delta_{w((1,1,1),a_{11})} = \Delta_{w((1,1,1),a_{21})} = 0,$$
$$\Delta_{w((1,1,1),a_{22})} = 4c\gamma_2;$$

for state $(1, 1, 2)$ we obtain

$$\Delta_{w((1,1,2),a_{02})} = \Delta_{w((1,1,2),a_{10})} = \frac{\gamma_1\gamma_2(1-2c)}{\gamma_1+\gamma_2}, \quad \Delta_{w((1,1,2),a_{11})} = 4c\gamma_1,$$
$$\Delta_{w((1,1,2),a_{21})} = 2c(\gamma_1+\gamma_2), \quad \Delta_{w((1,1,2),a_{22})} = 4c\gamma_2;$$

for state $(1,2,1)$ we obtain

$$\Delta_{w((1,2,1),a_{01})} = \Delta_{w((1,2,1),a_{20})} = \frac{\gamma_1\gamma_2(1-2c)}{\gamma_1+\gamma_2}, \quad \Delta_{w((1,2,1),a_{11})} = 4c\gamma_1,$$
$$\Delta_{w((1,2,1),a_{12})} = 2c(\gamma_1+\gamma_2), \quad \Delta_{w((1,2,1),a_{22})} = 4c\gamma_2;$$

and for state $(1,2,2)$ we have

$$\Delta_{w((1,2,2),a_{01})} = \Delta_{w((1,2,2),a_{10})} = \frac{\gamma_1\gamma_2(1-2c)}{\gamma_1+\gamma_2}, \quad \Delta_{w((1,2,2),a_{11})} = 4c\gamma_1,$$
$$\Delta_{w((1,2,2),a_{21})} = \Delta_{w((1,2,2),a_{22})} = 0.$$

Finally, for states $(2,1,1),(2,1,2)$ and $(2,2,1)$ we have

$$\Delta_{w((2,1,1),a_{12})} = \Delta_{w((2,1,2),a_{12})} = \Delta_{w((2,1,1),a_{21})} = \Delta_{w((2,2,1),a_{21})} = \frac{\gamma_2(\gamma_1-4c\gamma_1-2c\gamma_2)}{\gamma_1+\gamma_2},$$
$$\Delta_{w((2,1,2),a_{21})} = \Delta_{w((2,2,1),a_{12})} = \frac{\gamma_1\gamma_2(1-2c)}{\gamma_1+\gamma_2}.$$

These quantities are also nonnegative when c, γ_1, and γ_2 satisfy the assumptions above (note that $\frac{\gamma_2}{2\gamma_1+4\gamma_2} \leq \frac{\gamma_1}{4\gamma_1+2\gamma_2}$ because $\gamma_1 \geq \gamma_2$). Hence we have shown that the inequality (13) is satisfied for all nonbasic variables. We can conclude that D is an optimal basis for the LP (12), and consequently $\pi = (d)^\infty$ is an optimal policy when $0 \leq c \leq \frac{\gamma_2}{2\gamma_1+4\gamma_2}$. We see that the recurrent states of X_π are $(0,1,2)$, $(1,1,1)$, $(1,2,2)$, and $(2,1,2)$ under this policy. In the transient states (i.e., states in $\mathcal{S}\setminus\mathcal{S}_{w^*}$), we can select any action that will take the process to one of the recurrent states, and this shows that the policy described in the theorem is optimal when $0 \leq c \leq \frac{\gamma_2}{2\gamma_1+4\gamma_2}$.

Next, let $\frac{\gamma_2}{2\gamma_1+4\gamma_2} < c \leq \frac{\gamma_1^2}{2\gamma_1^2+2\gamma_1\gamma_2+2\gamma_2^2}$ (which also implies that $c \leq \frac{1}{2}$), and consider the decision rule d, where $d(x)$ is defined as follows for all $x \in \mathcal{S}$:

$$d(x) = \begin{cases} a_{12} & \text{if } x \in \{(0,1,2),(0,2,2),(1,1,1),(1,1,2)\}, \\ a_{21} & \text{if } x \in \{(0,2,1),(1,2,1)\}, \\ a_{22} & \text{if } x \in \{(1,2,2),(2,1,1),(2,1,2),(2,2,1)\}. \end{cases}$$

Then, the basic solution ω corresponding to the policy $\pi = (d)^\infty$ has the basis

$$\begin{aligned} D \;=\; & \{\omega((0,1,2),a_{12}), \omega((0,2,1),a_{21}), \omega((0,2,2),a_{12}), \\ & \omega((1,1,1),a_{12}), \omega((1,1,2),a_{12}), \omega((1,2,1),a_{21}), \omega((1,2,2),a_{22}), \\ & \omega((2,1,1),a_{22}), \omega((2,1,2),a_{22}), \omega((2,2,1),a_{22})\}. \end{aligned}$$

As before, we will show that inequality (13) holds for every nonbasic variable. More specifically, for states $(0,1,2),(0,2,1)$, and $(0,2,2)$, we have

$$\Delta_{w((0,1,2),a_{11})} = \Delta_{w((0,2,1),a_{11})} = \Delta_{w((0,2,2),a_{11})} = \frac{\gamma_1(2\gamma_1+\gamma_2)(4c\gamma_2+2c\gamma_1-\gamma_2)}{(\gamma_1+\gamma_2)^2},$$
$$\Delta_{w((0,1,2),a_{21})} = \Delta_{w((0,2,1),a_{12})} = 2c\gamma_1, \quad \Delta_{w((0,2,2),a_{21})} = 0.$$

These quantities are nonnegative because $c > \frac{\gamma_2}{2\gamma_1+4\gamma_2}$. For state $(1,1,1)$ we have

$$\Delta_{w((1,1,1),a_{02})} = \Delta_{w((1,1,1),a_{20})} = \frac{\gamma_1\gamma_2(1-2c)}{\gamma_1+\gamma_2}, \quad \Delta_{w((1,1,1),a_{11})} = \frac{\gamma_1\gamma_2(2c\gamma_1+4c\gamma_2-\gamma_2)}{(\gamma_1+\gamma_2)^2},$$
$$\Delta_{w((1,1,1),a_{21})} = 0, \quad \Delta_{w((1,1,1),a_{22})} = \frac{\gamma_2(2c\gamma_1^2+4c\gamma_2^2+4c\gamma_1\gamma_2+\gamma_1\gamma_2)}{(\gamma_1+\gamma_2)^2};$$

for state $(1,1,2)$ we obtain

$$\Delta_{w((1,1,2),a_{02})} = \frac{\gamma_1\gamma_2(1-2c)}{\gamma_1+\gamma_2}, \quad \Delta_{w((1,1,2),a_{10})} = \frac{\gamma_1\gamma_2(2c\gamma_2+\gamma_1)}{(\gamma_1+\gamma_2)^2},$$
$$\Delta_{w((1,1,2),a_{11})} = \frac{\gamma_1(4c\gamma_1^2+8c\gamma_2^2+10c\gamma_1\gamma_2-\gamma_2^2)}{(\gamma_1+\gamma_2)^2}, \quad \Delta_{w((1,1,2),a_{21})} = 2c(\gamma_1+\gamma_2),$$
$$\Delta_{w((1,1,2),a_{22})} = \frac{\gamma_2(2c\gamma_1^2+4c\gamma_2^2+4c\gamma_1\gamma_2+\gamma_1\gamma_2)}{(\gamma_1+\gamma_2)^2};$$

for state $(1,2,1)$ we obtain

$$\Delta_{w((1,2,1),a_{01})} = \frac{\gamma_1\gamma_2(\gamma_1+2c\gamma_2)}{(\gamma_1+\gamma_2)^2}, \quad \Delta_{w((1,2,1),a_{20})} = \frac{\gamma_1\gamma_2(1-2c)}{\gamma_1+\gamma_2},$$
$$\Delta_{w((1,2,1),a_{11})} = \frac{\gamma_1(4c\gamma_1^2+8c\gamma_2^2+10c\gamma_1\gamma_2-\gamma_2^2)}{(\gamma_1+\gamma_2)^2}, \quad \Delta_{w((1,2,1),a_{12})} = 2c(\gamma_1+\gamma_2),$$
$$\Delta_{w((1,2,1),a_{22})} = \frac{\gamma_2(2c\gamma_1^2+4c\gamma_2^2+4c\gamma_1\gamma_2+\gamma_1\gamma_2^2)}{(\gamma_1+\gamma_2)^2};$$

and for state $(1,2,2)$ we have

$$\Delta_{w((1,2,2),a_{01})} = \Delta_{w((1,2,2),a_{10})} = \frac{\gamma_2(2c\gamma_1^2+4c\gamma_2^2+4c\gamma_1\gamma_2+\gamma_1\gamma_2^2)}{(\gamma_1+\gamma_2)^2}$$
$$\Delta_{w((1,2,2),a_{11})} = \frac{\gamma_1(6c\gamma_1+8c\gamma_2-\gamma_2)}{(\gamma_1+\gamma_2)^2}, \quad \Delta_{w((1,2,2),a_{12})} = \Delta_{w((1,2,2),a_{21})} = \frac{\gamma_1(2c\gamma_1+4c\gamma_2-\gamma_2)}{2(\gamma_1+\gamma_2)}.$$

Note that

$$\frac{\gamma_2^2}{4\gamma_1^2+10\gamma_1\gamma_2+8\gamma_2^2} \le \frac{\gamma_1\gamma_2}{10\gamma_1\gamma_2+8\gamma_2^2} \le \frac{\gamma_1}{10\gamma_1+8\gamma_2} \le \frac{\gamma_1}{2\gamma_1+4\gamma_2},$$

because $\gamma_1 \ge \gamma_2 \ge 0$. Therefore, the above quantities are all nonnegative because $\frac{\gamma_2}{2\gamma_1+4\gamma_2} < c \le \frac{\gamma_1^2}{2\gamma_1^2+2\gamma_1\gamma_2+2\gamma_2^2}$. Finally, for states $(2,1,1)$, $(2,1,2)$, and $(2,2,1)$, we have

$$\Delta_{w((2,1,1),a_{12})} = \Delta_{w((2,1,2),a_{12})} = \Delta_{w((2,1,1),a_{21})} = \Delta_{w((2,2,1),a_{21})} = \frac{\gamma_2(\gamma_1^2-2c\gamma_1^2-2c\gamma_2^2-2c\gamma_1\gamma_2)}{(\gamma_1+\gamma_2)^2},$$
$$\Delta_{w((2,1,2),a_{21})} = \Delta_{w((2,2,1),a_{12})} = \frac{\gamma_1\gamma_2(2c\gamma_2+\gamma_1)}{(\gamma_1+\gamma_2)^2}.$$

These quantities are nonnegative because $c \le \frac{\gamma_1^2}{2\gamma_1^2+2\gamma_1\gamma_2+2\gamma_2^2}$. Hence, the policy $\pi = (d)^\infty$ is an optimal policy and the recurrent states of X_π under this policy are $(0,1,2)$, $(0,2,2)$, $(1,1,2)$, $(1,2,2)$, and $(2,1,2)$. In the transient states (i.e., states in $\mathcal{S} \setminus \mathcal{S}_{w^*}$),we can select any action that will take the process to one of the recurrent states, and this shows that the policy described in the theorem is optimal when $\frac{\gamma_2}{2\gamma_1+4\gamma_2} < c \le \frac{\gamma_1^2}{2\gamma_1^2+2\gamma_1\gamma_2+2\gamma_2^2}$.

Finally, let $c > \frac{\gamma_1^2}{2\gamma_1^2+2\gamma_1\gamma_2+2\gamma_2^2}$, and consider the decision rule d, where $d(x)$ is defined as follows for all $x \in \mathcal{S}$:

$$d(x) = \begin{cases} a_{12} & \text{if } x \in \{(0,1,2),(0,2,2),(1,1,1),(1,1,2),(2,1,1),(2,1,2)\}, \\ a_{21} & \text{if } x \in \{(0,2,1),(1,2,1)\}, \\ a_{22} & \text{if } x \in \{(1,2,2),(2,2,1)\}. \end{cases}$$

The basic solution ω corresponding to the policy $\pi = (d)^\infty$ has the basis

$$\begin{aligned} D &= \{\omega((0,1,2),a_{12}),\omega((0,2,1),a_{21}),\omega((0,2,2),a_{12}), \\ &\quad \omega((1,1,1),a_{12}),\omega((1,1,2),a_{12}),\omega((1,2,1),a_{21}),\omega((1,2,2),a_{22}), \\ &\quad \omega((2,1,1),a_{12}),\omega((2,1,2),a_{12}),\omega((2,2,1),a_{22})\}. \end{aligned}$$

As before, we will show that inequality (13) holds for every nonbasic variable. More specifically, for states $(0,1,2),(0,2,1)$, and $(0,2,2)$, we have

$$\begin{aligned} \Delta_{w((0,1,2),a_{11})} &= \Delta_{w((0,2,2),a_{11})} = \frac{\gamma_1(4c\gamma_1^2+4c\gamma_2^2+4c\gamma_1\gamma_2-\gamma_1\gamma_2-\gamma_1^2)}{\gamma_1^2+\gamma_1\gamma_2+\gamma_2^2}, \\ \Delta_{w((0,1,2),a_{21})} &= \frac{\gamma_1(4c\gamma_1^2+4c\gamma_2^2+4c\gamma_1\gamma_2-\gamma_1^2)}{\gamma_1^2+\gamma_1\gamma_2+\gamma_2^2}, \quad \Delta_{w((0,2,1),a_{11})} = \frac{\gamma_1(\gamma_1-\gamma_2)(2\gamma_1+\gamma_2)}{\gamma_1^2+\gamma_1\gamma_2+\gamma_2^2}, \\ \Delta_{w((0,2,1),a_{12})} &= \frac{\gamma_1^3}{\gamma_1^2+2\gamma_1\gamma_2+\gamma_2^2}, \quad \Delta_{w((0,2,2),a_{21})} = \frac{\gamma_1(2c\gamma_1^2+2c\gamma_2^2+2c\gamma_1\gamma_2)}{\gamma_1^2+\gamma_1\gamma_2+\gamma_2^2}. \end{aligned}$$

These quantities are nonnegative because $\gamma_1 \geq \gamma_2$ and $c > \frac{\gamma_1^2}{2\gamma_1^2+2\gamma_1\gamma_2+2\gamma_2^2}$. Note that $2\gamma_1^2 \geq \gamma_1^2+\gamma_1\gamma_2$, and hence $\frac{\gamma_1^2}{2\gamma_1^2+2\gamma_1\gamma_2+2\gamma_2^2} \geq \frac{\gamma_1^2+\gamma_1\gamma_1}{4\gamma_1^2+4\gamma_1\gamma_2+4\gamma_2^2}$. For state $(1,1,1)$ we have

$$\begin{aligned} \Delta_{w((1,1,1),a_{02})} &= \frac{\gamma_1\gamma_2^2}{\gamma_1^2+\gamma_1\gamma_2+\gamma_2^2}, \quad \Delta_{w((1,1,1),a_{20})} = \frac{\gamma_1(2c\gamma_1^2+2c\gamma_2^2+2c\gamma_1\gamma_2+\gamma_1\gamma_2-\gamma_1^2)}{\gamma_1^2+\gamma_1\gamma_2+\gamma_2^2}, \\ \Delta_{w((1,1,1),a_{11})} &= \frac{\gamma_1\gamma_2(\gamma_1-\gamma_2)}{\gamma_1^2+\gamma_1\gamma_2+\gamma_2^2}, \quad \Delta_{w((1,1,1),a_{21})} = \frac{(\gamma_1+\gamma_2)(2c\gamma_1^2+2c\gamma_2^2+2c\gamma_1\gamma_2-\gamma_1^2)}{\gamma_1^2+\gamma_1\gamma_2+\gamma_2^2}, \\ \Delta_{w((1,1,1),a_{22})} &= \frac{\gamma_2(4c\gamma_1^2+4c\gamma_2^2+4c\gamma_1\gamma_2+\gamma_1\gamma_2-\gamma_1^2)}{\gamma_1^2+\gamma_1\gamma_2+\gamma_2^2}; \end{aligned}$$

for state $(1,1,2)$ we obtain

$$\begin{aligned} \Delta_{w((1,1,2),a_{02})} &= \frac{\gamma_1\gamma_2^2}{\gamma_1^2+2\gamma_1\gamma_2+\gamma_2^2}, \quad \Delta_{w((1,1,2),a_{10})} = \frac{\gamma_1(4c\gamma_1^2+4c\gamma_2^2+4c\gamma_1\gamma_2+\gamma_1\gamma_2-2\gamma_1^2)}{\gamma_1^2+\gamma_1\gamma_2+\gamma_2^2}, \\ \Delta_{w((1,1,2),a_{11})} &= \frac{\gamma_1(4c\gamma_1^2+4c\gamma_2^2+4c\gamma_1\gamma_2+\gamma_1\gamma_2-\gamma_2^2)}{\gamma_1^2+\gamma_1\gamma_2+\gamma_2^2}, \\ \Delta_{w((1,1,2),a_{21})} &= \frac{(\gamma_1+\gamma_2)(4c\gamma_1^2+4c\gamma_2^2+4c\gamma_1\gamma_2-\gamma_1^2)}{\gamma_1^2+\gamma_1\gamma_2+\gamma_2^2}, \\ \Delta_{w((1,1,2),a_{22})} &= \frac{\gamma_2(4c\gamma_1^2+4c\gamma_2^2+4c\gamma_1\gamma_2+\gamma_1\gamma_2-\gamma_1^2)}{(\gamma_1+\gamma_2)^2}; \end{aligned}$$

for state $(1,2,1)$ we obtain

$$\begin{aligned}
\Delta_{w((1,2,1),a_{01})} &= \frac{\gamma_1^2\gamma_2}{\gamma_1^2+\gamma_1\gamma_2+\gamma_2^2}, \quad \Delta_{w((1,2,1),a_{20})} = \frac{\gamma_1\gamma_2^2}{\gamma_1^2+\gamma_1\gamma_2+\gamma_2^2},\\
\Delta_{w((1,2,1),a_{11})} &= \frac{\gamma_1(\gamma_1+\gamma_2(2\gamma_1-\gamma_2))}{\gamma_1^2+\gamma_1\gamma_2+\gamma_2^2}, \quad \Delta_{w((1,2,1),a_{12})} = \frac{\gamma_1^2(\gamma_1+\gamma_2)}{\gamma_1^2+\gamma_1\gamma_2+\gamma_2^2},\\
\Delta_{w((1,2,1),a_{22})} &= \frac{\gamma_1\gamma_2(\gamma_1+\gamma_2)}{\gamma_1^2+\gamma_1\gamma_2+\gamma_2^2};
\end{aligned}$$

and for state $(1,2,2)$ we have

$$\begin{aligned}
\Delta_{w((1,2,2),a_{01})} &= \frac{\gamma_1(8c\gamma_1^2+8c\gamma_2^2+8c\gamma_1\gamma_2+\gamma_1\gamma_2-3\gamma_1^2)}{(\gamma_1+\gamma_2)^2},\\
\Delta_{w((1,2,2),a_{10})} &= \frac{\gamma_2(4c\gamma_1^2+4c\gamma_2^2+4c\gamma_1\gamma_2+\gamma_1\gamma_2-\gamma_1^2)}{\gamma_1^2+\gamma_1\gamma_2+\gamma_2^2},\\
\Delta_{w((1,2,2),a_{11})} &= \frac{\gamma_1(4c\gamma_1^2+4c\gamma_2^2+4c\gamma_1\gamma_2+\gamma_1^2-\gamma_2^2)}{\gamma_1^2+\gamma_1\gamma_2+\gamma_2^2}, \quad \Delta_{w((1,2,2),a_{12})} = \frac{\gamma_1(\gamma_1-\gamma_2)(\gamma_1+\gamma_2)}{2\gamma_1^2+2\gamma_1\gamma_2+2\gamma_2^2},\\
\Delta_{w((1,2,2),a_{21})} &= \frac{(\gamma_1+\gamma_2)(4c\gamma_1^2+4c\gamma_2^2+4c\gamma_1\gamma_2-\gamma_1\gamma_2-\gamma_1^2)}{2\gamma_1^2+2\gamma_1\gamma_2+2\gamma_2^2}.
\end{aligned}$$

These quantities are nonnegative because $\gamma_1 \geq \gamma_2$ and $c > \frac{\gamma_1^2}{2\gamma_1^2+2\gamma_1\gamma_2+2\gamma_2^2} \geq \frac{\gamma_1^2+\gamma_1\gamma_1}{4\gamma_1^2+4\gamma_1\gamma_2+4\gamma_2^2}$. Finally, for states $(2,1,1), (2,1,2)$ and $(2,2,1)$ we have

$$\begin{aligned}
\Delta_{w((2,1,1),a_{21})} &= \frac{\gamma_2(2c\gamma_1^2+2c\gamma_2^2+2c\gamma_1\gamma_2-\gamma_1^2)}{\gamma_1^2+\gamma_1\gamma_2+\gamma_2^2},\\
\Delta_{w((2,1,1),a_{22})} &= \Delta_{w((2,1,2),a_{22})} = \frac{2\gamma_2(2c\gamma_1^2+2c\gamma_2^2+2c\gamma_1\gamma_2-\gamma_1^2)}{\gamma_1^2+\gamma_1\gamma_2+\gamma_2^2},\\
\Delta_{w((2,1,2),a_{21})} &= \frac{\gamma_2(4c\gamma_1^2+4c\gamma_2^2+4c\gamma_1\gamma_2-\gamma_1^2)}{2\gamma_1^2+2\gamma_1\gamma_2+2\gamma_2^2},\\
\Delta_{w((2,2,1),a_{12})} &= \frac{\gamma_1^2\gamma_2}{\gamma_1^2+\gamma_1\gamma_2+\gamma_2^2}, \quad \Delta_{w((2,2,1),a_{21})} = 0.
\end{aligned}$$

These quantities are nonnegative because $\gamma_1 \geq \gamma_2$ and $c > \frac{\gamma_1^2}{2\gamma_1^2+2\gamma_1\gamma_2+2\gamma_2^2}$. Hence, the policy $\pi = (d)^\infty$ is an optimal policy and the recurrent states under this policy are $(0,1,2)$, $(1,1,2)$, and $(2,1,2)$. In the transient states (i.e., states in $\mathcal{S} \setminus \mathcal{S}_{w^*}$), we can select any action that will take the process to one of the recurrent states, and this shows that the policy described in the theorem is optimal when $c > \frac{\gamma_1^2}{2\gamma_1^2+2\gamma_1\gamma_2+2\gamma_2^2}$. Hence the proof is complete. □

C Proof of Theorem 6.1

Let ξ_i, for $i \in \{1,\dots,10\}$, be defined as follows:

$$\begin{aligned}
\xi_1 &= \mu_{21}(\mu_{11}+\mu_{22})(\mu_{12}+\mu_{22}),\\
\xi_2 &= 2\Big((\mu_{11}^2(\mu_{11}+\mu_{12}+\mu_{21}+\mu_{22})+\mu_{21}\mu_{22}(\mu_{12}+\mu_{22})+\mu_{11}(\mu_{12}+\mu_{22})(2\mu_{21}+\mu_{22})\Big),
\end{aligned}$$

$$
\begin{aligned}
\xi_3 &= \mu_{12}(\mu_{11}+\mu_{22})(\mu_{11}+\mu_{21}),\\
\xi_4 &= 2\Big((\mu_{22}^2(\mu_{11}+\mu_{12}+\mu_{21}+\mu_{22})+\mu_{11}\mu_{12}(\mu_{11}+\mu_{12})+\mu_{22}(\mu_{11}+\mu_{21})(\mu_{11}+2\mu_{12})\Big),\\
\xi_5 &= \mu_{21}(\mu_{11}\mu_{22}-\mu_{12}\mu_{21})(\mu_{11}+\mu_{12}+\mu_{21}+\mu_{22})+\mu_{11}^2\Big(\mu_{12}^2+\mu_{21}\mu_{22}+\mu_{12}(\mu_{21}+\mu_{22})\Big)\\
&\quad +\mu_{11}\Big(\mu_{12}^2\mu_{22}+\mu_{21}\mu_{22}^2+\mu_{12}\mu_{22}(\mu_{21}+\mu_{22})\Big),\\
\xi_6 &= 2(\mu_{11}+\mu_{22})\Big[\mu_{21}\Big(\mu_{12}^2+(\mu_{12}+\mu_{21})(\mu_{21}+\mu_{22})\Big)\\
&\quad +\mu_{11}\Big(\mu_{12}^2+\mu_{12}(2\mu_{21}+\mu_{22})+\mu_{21}(\mu_{21}+2\mu_{22})\Big)\Big],\\
\xi_7 &= \mu_{12}(\mu_{11}\mu_{22}-\mu_{12}\mu_{21})(\mu_{11}+\mu_{12}+\mu_{21}+\mu_{22})+\mu_{22}^2\Big(\mu_{21}^2+\mu_{11}\mu_{12}+\mu_{21}(\mu_{11}+\mu_{12})\Big)\\
&\quad +\mu_{22}\Big(\mu_{11}\mu_{21}^2+\mu_{11}^2\mu_{12}+\mu_{11}\mu_{21}(\mu_{11}+\mu_{12})\Big),\\
\xi_8 &= 2(\mu_{11}+\mu_{22})\Big[\mu_{12}\Big(\mu_{12}^2+(\mu_{12}+\mu_{21})(\mu_{11}+\mu_{12})\Big)\\
&\quad +\mu_{22}\Big(\mu_{21}^2+\mu_{12}(2\mu_{11}+\mu_{12})+\mu_{21}(\mu_{11}+2\mu_{12})\Big)\Big],\\
\xi_9 &= (\mu_{11}\mu_{22}-\mu_{12}\mu_{21})(\mu_{11}+\mu_{12}+\mu_{21}+\mu_{22}),\\
\xi_{10} &= 2(\mu_{11}+\mu_{22})\Big(\mu_{12}^2+(\mu_{12}+\mu_{21})(\mu_{21}+\mu_{22})+\mu_{11}(\mu_{12}+\mu_{21}+\mu_{22})\Big).
\end{aligned}
$$

Note that the states $(0,1,1)$ and $(2,2,2)$ are transient under any policy $\pi \in \Pi$. Hence the actions in these states do not affect the long-run average profit and we omit these states in the rest of the proof. First assume that $0 \le c \le \min\{\beta_1, \beta_2, \beta_5\}$. Consider the decision rule d, where $d(x)$ is defined as follows for all $x \in \mathcal{S}$:

$$
d(x) = \begin{cases} a_{11} & \text{if } x \in \{(0,1,2),(0,2,1),(0,2,2)\},\\ a_{12} & \text{if } x \in \{(1,1,1),(1,1,2),(1,2,1),(1,2,2)\},\\ a_{22} & \text{if } x \in \{(2,1,1),(2,1,2),(2,2,1)\}. \end{cases}
$$

Similar calculations to those in the proof of Theorem 4.1 show that the policy $\pi = (d)^\infty$ is an optimal policy when $0 \le c \le \min\{\beta_1, \beta_2, \beta_5\}$. We see that the recurrent states of X_π are $(0,1,2)$, $(1,1,1)$, $(1,2,2)$, and $(2,1,2)$ under this policy.

Next, assume that $\beta_5 \le c \le \min\{\beta_1, \beta_2, \beta_3, \beta_4\}$ (some algebra shows that $\beta_5 \le \min\{\beta_3, \beta_4\}$, hence this interval is non-empty when $\beta_5 \le c \le \min\{\beta_1, \beta_2\}$). Consider the decision rule d, where $d(x)$ is defined as follows for all $x \in \mathcal{S}$:

$$
d(x) = \begin{cases} a_{11} & \text{if } x \in \{(0,1,2),(0,2,1),(0,2,2)\},\\ a_{12} & \text{if } x \in \{(1,1,1),(1,1,2),(1,2,2)\},\\ a_{21} & \text{if } x \in \{(1,2,1)\},\\ a_{22} & \text{if } x \in \{(2,1,1),(2,1,2),(2,2,1)\}. \end{cases}
$$

Similar calculations to those in the proof of Theorem 4.1 show that the policy $\pi = (d)^\infty$ is an optimal policy when $\beta_5 \le c \le \min\{\beta_1, \beta_2, \beta_3, \beta_4\}$. We see that the recurrent states of X_π are $(0,1,2)$, $(1,1,1)$, $(1,2,2)$, and $(2,1,2)$ under this policy. In the transient states (i.e., states in

$\mathcal{S}\backslash\mathcal{S}_{w^*}$), we can select any action that will take the process to one of the recurrent states, and this shows that the policy π^* described in the theorem is optimal when $\beta_5 \leq c \leq \min\{\beta_1, \beta_2, \beta_3, \beta_4\}$. This completes the proof of part (i) of the theorem.

Now, assume that the conditions in part (ii) of the theorem are satisfied. Let $\pi' = (d')^\infty$, where $d'(x) = a_{12}$ for all $x \in \mathcal{S}$. The condition $T_1 \geq T_2$ guarantees that π' is not worse than the policy $\pi'' = (d'')^\infty$, where $d''(x) = a_{21}$ for all $x \in \mathcal{S}$. Next, we want to show that there is no policy that allows switching of servers between stations that is better than π'. Without loss of generality, we only compare π' with policies that allow switching of servers between stations and have positive revenue (because π' is better than any policy with zero or negative revenue). We denote the set of policies that include switching and have positive revenue by Π^s, and we let $\mathcal{S}_1 = \{(1, z) : z \in \mathcal{S}_Z\}$. Under any $\pi \in \Pi^s$, there is exactly one departure from the system between two successive visits of the stochastic process X_π to a state in $\mathcal{S}_1$. We now show that for all $\pi \in \Pi^s$, there will be at least one setup with positive probability between every two visits to $\mathcal{S}_1$.

Note that under any $\pi \in \Pi^s$, every time X_π leaves the state $(1, 1, 1)$ or $(1, 2, 2)$, there has to be at least one setup before the next time the process enters a state in $\mathcal{S}_1$ (either when leaving $\mathcal{S}_1$ or when coming back to $\mathcal{S}_1$), or otherwise the long-run average profit is zero. In state $(1, 2, 1)$, if $\mu_{12} = \mu_{21} = 0$ and action a_{21} is used, then the long-run average profit is equal to zero, and if an action other than a_{21} is used, then at least one setup occurs before returning to $\mathcal{S}_1$. Furthermore, if $\mu_{12} = 0$ or $\mu_{21} = 0$, any policy that uses the action a_{21} in $\mathcal{S}_1$ results in at least one setup at an end of the line (otherwise the long-run average profit is equal to zero). Hence we can assume that $\mu_{12} > 0$ and $\mu_{21} > 0$ when X_π is in state $(1, 2, 1)$. Note that under any $\pi \in \Pi^s$, every time X_π leaves the state $(1, 1, 2)$, there has to be at least one setup with probability $p_s \geq \min\{\frac{\mu_{11}}{\mu_{11}+\mu_{22}}, \frac{\mu_{22}}{\mu_{11}+\mu_{22}}\} > 0$ before the next time the process enters a state in $\mathcal{S}_1$. Similarly, when $\mu_{12} > 0$, $\mu_{21} > 0$, and X_π leaves state $(1, 2, 1)$, there has to be at least one setup before X_π returns to $\mathcal{S}_1$ with probability $p'_s \geq \min\{\frac{\mu_{12}}{\mu_{12}+\mu_{21}}, \frac{\mu_{21}}{\mu_{12}+\mu_{21}}\} > 0$. The previous two facts follow because either an action other than a_{12} (a_{21}) is taken in state $(1, 1, 2)$ ($(1, 2, 1)$), in which case there will be at least one setup before returning to $\mathcal{S}_1$, or action a_{12} (a_{21}) is taken in state $(1, 1, 2)$ ($(1, 2, 1)$) and there has to be at least one setup at either end of the line before coming back to $\mathcal{S}_1$ (because otherwise π is not a switching policy). The four terms in the lower bounds on p_s and p'_s are equal to the probabilities of moving to $(0, 1, 2)$ or $(2, 1, 2)$ under a_{12} and the probabilities of moving to $(0, 2, 1)$ or $(2, 2, 1)$ under a_{21}. We have shown that the expected setup cost between two visits to a state in $\mathcal{S}_1$ cannot be less than cp_s or $c\min\{p_s, p'_s\}$ depending on whether $\mu_{12}\mu_{21} = 0$ or $\mu_{12}\mu_{21} > 0$.

Let v be minimum expected time between two visits to $\mathcal{S}_1$ (note that $v > 0$ because $\mu_{ij} < \infty$

for $i, j \in \{1, 2\}$). Then v is the sum of the minimum expected times for leaving $\mathcal{S}_1$ (i.e., Φ) and for returning back to $\mathcal{S}_1$ (i.e., ϕ_1), so that $v = \Phi + \phi_1$. By the renewal reward theorem, we can conclude that $P_\pi \leq \frac{1-cp_s}{v}$ ($P_\pi \leq \frac{1-c\min\{p_s, p'_s\}}{v}$) when $\mu_{12}\mu_{21} = 0$ ($\mu_{12}\mu_{21} > 0$) for all $\pi \in \Pi^s$. Hence, when $T_1 \geq \frac{1-cp_s}{v}$ (i.e., $c > \Big(1 - T_1(\Phi + \phi_1)\Big)(1 + \Theta)$), then no policy in Π^s can be optimal. Consequently, the policy that uses $d(x) = a_{12}$ for all $x \in \mathcal{S}$ is optimal. This proves part (ii) of the theorem.

Finally, assume that the conditions in part (iii) of the theorem is satisfied. Then we must have $\mu_{12} > 0$ and $\mu_{21} > 0$. Similar arguments as for part (ii) show that the policy that uses the decision rule $d(x) = a_{21}$ for all $x \in \mathcal{S}$ is optimal. □

References

[1] Andradóttir, S., H. Ayhan, and D. G. Down. 2001. Server Assignment Policies for Maximizing the Steady-State Throughput of Finite Queueing Systems. *Management Science*, 47, 1421–1439.

[2] Andradóttir, S., H. Ayhan, and D. G. Down. 2003. Dynamic Server Allocation for Queueing Networks with Flexible Servers. *Operations Research*, 51, 952–968.

[3] Andradóttir, S., H. Ayhan, and D. G. Down. 2007. Compensating for Failures with Flexible Servers. *Operations Research*, 55, 753–768.

[4] Bartholdi, III, J. J., and D. D. Eisenstein. 1996. A Production Line that Balances Itself. *Operations Research*, 44, 21–34.

[5] Bartholdi, III, J. J., D. D. Eisenstein, and R. D. Foley. 2001. Performance of Bucket Brigades when Work is Stochastic. *Operations Research*, 49, 710–719.

[6] Bertsimas, D., and J. N. Tsitsiklis. 1997. *Introduction to Linear Optimization.* Athena Scientific, Belmont, Massachusets.

[7] Billingsley, P. 1995. *Probability and Measure*, Third Edition. Wiley, New York, NY, 1995.

[8] Duenyas, I., D. Gupta, and T. L. Olsen. 1998. Control of a Single Server Queueing System with Setups. *Operations Research*, 46, 218–230.

[9] Duenyas, I., and M. P. Van Oyen. 1995. Stochastic Scheduling of Queues with Setup Costs. *Queueing Systems*, 19, 421–444.

[10] Gupta, D., and M. M. Srinivasan. 1996. Polling Systems with State Dependent Setup Times. *Queueing Systems*, 22, 403–423.

[11] Hofri, M., and K. W. Ross. 1987. On the Optimal Control of Two Queues with Server Setup Times and Its Analysis. *SIAM Journal on Computing*, 16, 399–420.

[12] Hopp, W. J., and M. P. Van Oyen. 2004. Agile Workforce Evaluation: A Framework for Cross-training and Coordination. *IIE Transactions*, 36, 919–940.

[13] Iravani, S. M. R., M. J. M. Posner, and J. M. Buzacott. 1997. A Two-Stage Tandem Queue Attended by a Moving Server with Holding and Switching Costs. *Queueing Systems*, 26, 203–228.

[14] Kırkızlar, H. E. 2008. Performance Improvements through Flexible Workforce. Ph.D. Thesis, H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia.

[15] Kırkızlar, H. E., S. Andradóttir, and H. Ayhan. 2008. Robustness of Effective Server Assignment Policies to Service Time Distributions. *Submitted for publication.*

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

[17] Lippman, S. A. 1975. Applying a New Device in the Optimization of Exponential Queueing System. *Operations Research*, 23, 687–710.

[18] Mayorga, M. E., K. M. Taaffe, and R. Arumugam. 2009. Allocating Flexible Servers in Serial Systems with Switching Costs. *Annals of Operations Research*, 172, 231–242.

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

[20] Reiman, M. I., and L. M. Wein. 1998. Dynamic Scheduling of a Two-Class Queue with Setups. *Operations Research*, 46, 532–547.

[21] Sennott, L. I., M. P. Van Oyen, and S. M. R. Iravani. 2006. Optimal Dynamic Assignment of a Flexible Server on an Open Production Line with Specialists. *European Journal of Operational Research*, 170, 541–566.

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