Online Supplement to Flexible Servers in Tandem Lines with Setup Costs

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 13902-6000, U.S.A.

Appendices

A Proof of Theorem 4.1

Let ξ_i, for $i \in \{1, \ldots, 10\}$, be defined by

$$\begin{aligned}
\xi_1 &= \mu_{21}(\mu_{11}+\mu_{22})(\mu_{12}+\mu_{22}),\\
\xi_2 &= 2\Big((\mu_{11}^2(\mu_{11}+\mu_{12}+\mu_{21}+\mu_{22})+\mu_{21}\mu_{22}(\mu_{12}+\mu_{22})+\mu_{11}(\mu_{12}+\mu_{22})(2\mu_{21}+\mu_{22})\Big),\\
\xi_3 &= \mu_{12}(\mu_{11}+\mu_{22})(\mu_{11}+\mu_{21}),\\
\xi_4 &= 2\Big((\mu_{22}^2(\mu_{11}+\mu_{12}+\mu_{21}+\mu_{22})+\mu_{11}\mu_{12}(\mu_{11}+\mu_{12})+\mu_{22}(\mu_{11}+\mu_{21})(\mu_{11}+2\mu_{12})\Big),\\
\xi_5 &= \mu_{21}(\mu_{11}\mu_{22}-\mu_{12}\mu_{21})(\mu_{11}+\mu_{12}+\mu_{21}+\mu_{22})+\mu_{11}^2\Big(\mu_{12}^2+\mu_{21}\mu_{22}+\mu_{12}(\mu_{21}+\mu_{22})\Big)\\
&\quad+\mu_{11}\Big(\mu_{12}^2\mu_{22}+\mu_{21}\mu_{22}^2+\mu_{12}\mu_{22}(\mu_{21}+\mu_{22})\Big),\\
\xi_6 &= 2(\mu_{11}+\mu_{22})\Big[\mu_{21}\Big(\mu_{12}^2+(\mu_{12}+\mu_{21})(\mu_{21}+\mu_{22})\Big)\\
&\quad+\mu_{11}\Big(\mu_{12}^2+\mu_{12}(2\mu_{21}+\mu_{22})+\mu_{21}(\mu_{21}+2\mu_{22})\Big)\Big],\\
\xi_7 &= \mu_{12}(\mu_{11}\mu_{22}-\mu_{12}\mu_{21})(\mu_{11}+\mu_{12}+\mu_{21}+\mu_{22})+\mu_{22}^2\Big(\mu_{21}^2+\mu_{11}\mu_{12}+\mu_{21}(\mu_{11}+\mu_{12})\Big)\\
&\quad+\mu_{22}\Big(\mu_{11}\mu_{21}^2+\mu_{11}^2\mu_{12}+\mu_{11}\mu_{21}(\mu_{11}+\mu_{12})\Big),\\
\xi_8 &= 2(\mu_{11}+\mu_{22})\Big[\mu_{12}\Big(\mu_{12}^2+(\mu_{12}+\mu_{21})(\mu_{11}+\mu_{12})\Big)\\
&\quad+\mu_{22}\Big(\mu_{21}^2+\mu_{12}(2\mu_{11}+\mu_{12})+\mu_{21}(\mu_{11}+2\mu_{12})\Big)\Big],\\
\xi_9 &= (\mu_{11}\mu_{22}-\mu_{12}\mu_{21})(\mu_{11}+\mu_{12}+\mu_{21}+\mu_{22}),\\
\xi_{10} &= 2(\mu_{11}+\mu_{22})\Big(\mu_{12}^2+(\mu_{12}+\mu_{21})(\mu_{21}+\mu_{22})+\mu_{11}(\mu_{12}+\mu_{21}+\mu_{22})\Big).
\end{aligned}$$

Note that the states $(0,1,1)$ and $(2,2,2)$ are transient under any policy $\pi \in \Pi$ with positive revenue. 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 $0 \leq c \leq \min\{\beta_1, \beta_2, \beta_5\}$. 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 4.2 show that the policy $\pi = (d)^\infty$ is an optimal policy when $0 \leq c \leq \min\{\beta_1, \beta_2, \beta_5\}$. 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 $\beta_5 \leq c \leq \min\{\beta_1, \beta_2, \beta_3, \beta_4\}$ (some algebra shows that $\beta_5 \leq \min\{\beta_3, \beta_4\}$, hence this interval is non-empty when $\beta_5 \leq c \leq \min\{\beta_1, \beta_2\}$). 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 4.2 show that the policy $\pi = (d)^\infty$ is an optimal policy when $\beta_5 \leq c \leq \min\{\beta_1, \beta_2, \beta_3, \beta_4\}$. 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. □

B Proof of Theorem 4.2

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 $\pi \in \Pi$ with positive revenue, 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 state and action spaces are finite, Theorem 9.1.8 of Puterman [23] 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 Programming (LP) approach for communicating Markov decision processes as in Sections 9.5.2 and 8.8.2 of Puterman [23].

Consider the following LP:

$$\left.\begin{array}{ll} \max & \sum_{x \in \mathcal{S}} \sum_{a \in A_x} r(x,a)\omega(x,a) \\ \text{s.t.} & \sum_{a \in A_{x'}} \omega(x',a) - \sum_{x \in \mathcal{S}} \sum_{a \in A_x} p(x'|x,a)\omega(x,a) = 0, \text{ for all } x' \in \mathcal{S}, \\ & \sum_{x \in \mathcal{S}} \sum_{a \in A_x} \omega(x,a) = 1, \\ & \omega(x,a) \geq 0, \text{ for all } x \in \mathcal{S}, a \in A_x, \end{array}\right\} \quad (8)$$

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 [23] (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

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

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 [23] shows that the decision rule above results in an optimal solution. Moreover, note that an action a' that will move the process X'_π toward 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 (8) corresponding to the policy $\pi = (d)^\infty$. The associated basis for the LP (8) 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

$$\Delta_y = c_B \mathbf{B}^{-1} v_y - c_y \geq 0 \tag{9}$$

for each nonbasic variable y, where v_y is the column in the constraint matrix of the LP (8) and c_y is the coefficient corresponding to y in the objective function (see, e.g., Theorem 3.1 of Bertsimas and Tsitsiklis [7]).

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

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

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

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

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

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

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 $\frac{\gamma_2}{2\gamma_1+4\gamma_2} \le \frac{\gamma_1}{4\gamma_1+2\gamma_2}$ because $\gamma_1 \ge \gamma_2$). Hence we have shown that the inequality (9) is satisfied for all nonbasic variables. We can conclude that D is an optimal basis for the LP (8), and consequently $\pi = (d)^\infty$ is an optimal policy when $0 \le c \le \frac{\gamma_2}{2\gamma_1+4\gamma_2}$. 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 $0 \le c \le \frac{\gamma_2}{2\gamma_1+4\gamma_2}$.

Next, let $\frac{\gamma_2}{2\gamma_1+4\gamma_2} < c \le \frac{\gamma_1^2}{2\gamma_1^2+2\gamma_1\gamma_2+2\gamma_2^2}$ (which also implies that $c \le \frac{1}{2}$), 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 (9) 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

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

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

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

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

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

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

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

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

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

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 (9) holds for every nonbasic variable. More specifically, for states $(0,1,2), (0,2,1)$, and $(0,2,2)$, we have

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

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

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

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

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

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

$$\Delta_{w((1,2,1),a_{01})} = \frac{\gamma_1^2\gamma_2}{\gamma_1^2+\gamma_1\gamma_2+\gamma_2^2}, \quad \Delta_{w((1,2,1),a_{20})} = \frac{\gamma_1\gamma_2^2}{\gamma_1^2+\gamma_1\gamma_2+\gamma_2^2},$$

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

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

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

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

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

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 Proofs of Propositions 4.1, 4.2, and 4.3

Proof of Proposition 4.1: First, assume that $\mu_1\gamma_1 \neq \mu_2\gamma_2$. Let $\pi_0 = (d_0)^\infty$ be as described in Section 4. 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

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

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. □

Proof of Proposition 4.2: 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

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

The first term in $\alpha_{l,3}$ is positive for all $l \in \{0, \dots, B-1\}$, 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 $i \in \{0, \dots, B+1-l\}$. 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 $l \in \{0, \dots, B-1\}$. Moreover, $\alpha_{l,2} > 0$ trivially. This shows that $P_{\pi_l} > P_{\pi_{l+1}}$ for $c \geq c_l$, where $c_l = \alpha_{l,1}/\alpha_{l,2}$. Some algebra shows that for $l \in \{0, \dots, B-1\}$,

$$c_l \geq c_{l+1} \Leftrightarrow (B+1-l)\gamma_1^{B+2}(\gamma_1+\gamma_2) \geq 2\gamma_2^{B+3}\sum_{k=0}^{B-l}(\frac{\gamma_1}{\gamma_2})^k.$$

The last inequality follows because $\gamma_1 \geq \gamma_2$, and hence the threshold decreases as the holding cost increases.

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

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

This expression is positive for $l \in \{0, \dots, B-1\}$ and $c \geq c_l$, where $c_l = \frac{2+(B-l)^2+3(B-l)}{4(3+2l)}$. Some algebra shows that

$$c_l \geq c_{l+1} \Leftrightarrow 4 + 2(B-l)^2 + 6(B-l) + 2(3+2l)(B+1-l) \geq 0. \tag{11}$$

The last inequality holds trivially for $l \in \{0, \dots, B-1\}$, and this completes the proof. □

Proof of Proposition 4.3: 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 $l \in \{0, \dots, B\}$. 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 $P_{\pi_l} - P_{\pi_{l+1}} = (-\beta_{l,1} + c\beta_{l,2})/\beta_{l,3}$ for $l \in \{0, \dots, B-1\}$, 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 $c \geq c_l$, where $c_l = \beta_{l,1}/\beta_{l,2}$. Some algebra shows that for $l \in \{0, \ldots, B-1\}$,

$$c_l \geq c_{l+1} \Leftrightarrow 2(B+1-l)\mu_1^{B+3} \geq \mu_2^{B+2}(\mu_1+\mu_2)\sum_{k=0}^{B-l}(\frac{\mu_1}{\mu_2})^k.$$

The last inequality follows because $\mu_1 \geq \mu_2$, and hence the threshold decreases as the holding cost increases.

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 $c \geq c_l$. We observe that $c_l \geq c_{l+1}$ because of (11). This completes the proof. $\square$