Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Collaboration versus Specialization in Service Systems with Impatient Customers

Bihan Chatterjee
Georgia Institute of Technology
Sigrún Andradóttir
sa@gatech.edu
Georgia Institute of Technology
Hayriye Ayhan
Georgia Institute of Technology

Research Article

Keywords: Queueing, server flexibility, server synergy, customer abandonments, Markov decision processes

Posted Date: July 15th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10121268/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Collaboration versus Specialization in Service Systems with Impatient Customers

Bihan Chatterjee, Sigrún Andradóttir*, Hayriye Ayhan

H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, USA.

*Corresponding author. E-mail: sa@gatech.edu;
Contributing authors: bchatterjee9@gatech.edu;
hayriye.ayhan@isye.gatech.edu;

Abstract

We study tandem queueing systems in which servers work more efficiently in teams than on their own and customers are impatient in that they may leave the system while waiting for service. Our goal is to determine the server assignment policy that maximizes the long-run average throughput. We show that when each server is equally skilled at all tasks, the optimal policy has all the servers working together at all times. We also provide a complete characterization of the optimal policy for Markovian systems with two stations and two servers when each server's efficiency may be task dependent. We show that the throughput is maximized under the policy which assigns one server to each station (based on their relative skill at that station) unless station 2 has no work (in which case both servers work at station 1) or the number of customers in the buffer reaches a threshold whose value we characterize (in which case both servers work at station 2). We study how the optimal policy varies with the level of server synergy (including no synergy) and also compare the optimal policy for systems with different customer abandonment rates (including no abandonments). Finally, we investigate the case where the synergy among collaborating servers can be task-dependent and provide numerical results.

Keywords: Queueing; server flexibility; server synergy; customer abandonments; Markov decision processes.

1 Introduction

In the past decade, the dynamic server assignment problem has received a lot of attention, resulting in a significant body of research addressing the question of how flexible servers should be assigned to different tasks in order to optimize system performance. In this paper we consider tandem queueing systems where collaborating servers work more efficiently than on their own and customers may abandon the system while waiting for service. Our objective is to determine the dynamic server assignment policy that maximizes the long-run average throughput. Our work is motivated by service systems in which delays can cause customers/jobs to leave the system altogether. In addition to impatient customers, such settings include deadline-sensitive jobs, perishable work, and tasks that fail, time out, or become obsolete while waiting for downstream processing. We use the terms "customers" and "abandonment" throughout, but the model should be interpreted broadly to include these settings as well.

More specifically, we consider a queueing network with $K \geq 1$ stations and $M \geq 1$ flexible servers. There is an infinite amount of raw material in front of station 1, infinite room for departing customers after station K, a finite buffer between stations j and $j+1$, for $j \in \{1, \ldots, K-1\}$, whose size is B_j, and the system is operating under manufacturing blocking. At any time there can be at most one customer at each station, and each server can work on at most one customer. The servers are flexible, in that they are cross-trained and allowed to switch between stations with negligible travel time. We assume that the service rate of server $i \in \{1, \ldots, M\}$ at station $j \in \{1, \ldots, K\}$ is μ_{ij} and, without loss of generality, we assume the mean service requirement of each customer at each station is 1. When multiple servers work on a single customer, there is synergy between them. In particular, if servers $i_1, \ldots, i_k \in \{1, \ldots, M\}$ work together on a customer at station $j \in \{1, \ldots, K\}$, the service rate is equal to $\gamma \sum_{r=1}^{k} \mu_{i_r j}$, where $\gamma \geq 1$ is the synergistic factor. Additionally, any customer may abandon the system while waiting to be processed by any of stations 2 through K;customers do not abandon while in service.

For systems where each server is equally skilled at all tasks, we show that the optimal throughput is achieved when all the servers work together in one team at all times. We also completely characterize the optimal server assignment policy for systems with exponential service requirements and abandonment times, $K = 2$ stations, and $M = 2$ servers when the service rates of each server at different stations are allowed to be heterogeneous. Further, we study some properties of the optimal policy and determine how it depends on the magnitude γ of the synergy among the servers as well as on the abandonment rate θ of the customers. Additionally, we propose a policy for systems with station (or task) dependent synergy and numerically verify its optimality. We also determine how the policy depends on the synergistic factors at the two stations.

The earliest works on server assignment problems (e.g., [1–6]) assumed the service rates to be additive, i.e., the service rate at a station is simply the sum over the rates of all servers working there (corresponding to $\gamma = 1$ in our model). However, in practice it is often observed that when multiple servers work together on a single customer, their collaboration could be synergistic and therefore the assumption that the service rates are additive is restrictive. Optimal server assignment for systems

with synergistic servers has been investigated in [7]. The authors analyze when it is better to take advantage of synergy among servers, rather than exploiting the servers' special skills, to achieve the best possible system throughput. In [8], a more general system is considered in which the synergy among collaborating servers can be task-dependent. More recently, [9] compares collaborative and non-collaborative service in multiclass multi-server systems with compatibility constraints. The authors show that collaboration is beneficial in symmetric systems, where either the servers or the job classes are stochastically indistinguishable from one another, but that the value of collaboration is limited and can depend strongly on system asymmetry.

Another factor that arises in many applications, such as call centers and hospital emergency rooms, is that customers can get impatient and leave the system while waiting for service. Queues with abandonments have been well studied in the literature (see, e.g., [10–14]). However, abandonment is a difficult phenomenon to analyze exactly and there are only a few papers that study optimal server assignment policies for queues with abandonments. For example, in [15] an optimal server assignment policy is provided for a system with two classes of customers and a single server to serve both classes when both classes have equal service rates. The authors allow the customer abandonment rates to be class dependent and derive the optimal policy under the two most common cost/reward structures, namely to maximize rewards per service or minimize holding costs per customer per unit time. This result was extended in [16] to K classes of customers with heterogeneous service rates.

Recent work has also studied dynamic scheduling in multiclass queues with abandonments. In [17], asymptotic optimality of $c\mu/\theta$ type priority rules is established for many-server queues with exponential patience times and linear holding costs, while [18] extends this line of work to general patience-time distributions and nonlinear costs. In [19], patients waiting for treatment in a hospital emergency department were modeled as parallel queues with impatient customers. In [20], the authors studied tandem queueing systems of two stations with the second station having finite buffer capacity. The authors develop a recursive algorithm to compute the throughput under various staffing strategies when the customers can get impatient and leave while waiting for service at either station. A similar system has been considered in [21], but with infinite buffer capacity at the second station. The authors consider non-collaborative homogeneous servers and rewards for each service completion. Reference [22] generalizes the setting by allowing new customers to arrive at both stations and considers holding costs and abandonment costs instead of rewards. Customers arriving at station 1 may go through both phases of service and customers arriving at station 2 may go through only one phase of service. In [21] and [22], the authors provide conditions under which there exists an optimal policy where the servers need not be split between stations. They also specify when priority rules are optimal for the single-server model. In [23], a complete characterization of the dynamic assignment policy of collaborative heterogeneous servers in tandem queues in the presence of customer abandonments was considered for the first time. The authors provide the policy that maximizes the long-run average throughput. They show that unlike the optimal policy without abandonments, here it is not necessarily optimal to use the entire buffer space, suggesting that there is a trade-off between using servers where they are effective and avoiding abandonments.

To the best of our knowledge, systems with both synergy among collaborating servers and customer abandonments have not been studied yet. The interaction between server synergy, server heterogeneity, and abandonments creates a tradeoff that is not present when these features are studied separately. Assigning servers to their relatively preferred stations uses specialization efficiently and can have more customers in service rather than leaving them waiting and exposed to abandonment. On the other hand, assigning servers to work together exploits synergy and can reduce the loss of customers or jobs waiting for downstream service. The goal of the paper is to characterize these tradeoffs.

The outline of the paper is as follows. In Section 2, we show that for systems with generalist servers, the optimal throughput is achieved when all servers work together at all times. In Section 3, we completely characterize the optimal policy for Markovian systems with two stations and two servers. In Section 4, we study how the optimal policy varies as the synergy between collaborating servers increases or when the abandonment rate of the customers changes. Section 5 provides numerical experiments to validate our theoretical findings and investigate the conditions under which the simple expedite policy serves as an effective heuristic. In Section 6 we investigate the case when the synergy can be task dependent and provide numerical results. Section 7 summarizes our findings. Finally, the proof of our main theorem and details of some other proofs and results, are included in the Appendix.

2 Systems with Generalist Servers

In this section we consider systems with generalists servers (i.e., each server is equally skilled at all tasks). Here, the service rate of each server at each station can be expressed as a product of two constants, one representing the server's speed and the other representing the difficulty of the task at the station. That is $\mu_{ij} = \mu_i \delta_j$ $\forall i \in \{1, \ldots, M\}$ and $j \in \{1, \ldots, K\}$. For each $j = 1, \ldots, K$, let the service requirement of customer $k \geq 1$ at station j be $S_{k,j}$. Assume $S_{k,j}$ are independent and identically distributed (iid) (but not necessarily exponentially distributed). And independent of the $S_{k,j}$, let $Y_{k,j}$ for $j = 2, \ldots, K$ be the patience time of customer k at station j, that is the maximum time the customer will wait for service at that station before abandoning. Let Π be the set of all server assignment policies under consideration and let $D_\pi(t)$ denote the number of departures after service completion at station K under policy $\pi \in \Pi$ by time $t \geq 0$. Define

$$g_\pi = \limsup_{t \to \infty} \frac{\mathbb{E}[D_\pi(t)]}{t} \tag{1}$$

as the long-run average throughput under policy π. Our goal is to solve the following optimization problem

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

The 'expedite policy' assigns all servers to a single team that will follow each customer from the first to the last station and only starts work on a new customer

once all work on the previous customer has been completed. The next theorem states that when the servers are generalists, the expedite policy is the optimal policy. This is not surprising because when servers are generalists, there is no tradeoff between exploiting server synergy, assigning servers to tasks at which they are more effective, and avoiding abandonments. Since each server's relative efficiency is the same across all stations, assigning the servers to different stations cannot take advantage of any task-specific skills. The expedite policy therefore fully exploits the synergy among the servers by keeping them together as a single team. At the same time, because the team follows each customer through all stations before starting a new customer, no customer waits between stations and hence no abandonment occurs.

Theorem 1 *Assume that either*

1. *Service is non-preemptive or preemptive-repeat and for all $t \geq 0$, if there is a customer in service at station j at time t, then the expected remaining service requirement at station j of that customer is bounded by a scalar $1 \leq \bar{S} < \infty$, or*
2. *Service is preemptive-resume and the service requirements are new better than used in expectation, i.e.,*

$$\mathbb{E}[S_{k,j} - x \mid S_{k,j} > x] \leq \mathbb{E}[S_{k,j}] = 1$$

for all $x \geq 0$, for each $j = 1, \ldots, K$.

When $\mu_{ij} = \mu_i \delta_j$ for all $i = 1, \ldots, M$ and $j = 1, \ldots, K$ and $\gamma \geq 1$, then the expedite policy is optimal with long-run average throughput $g_\pi = \frac{\gamma \sum_{i=1}^{M} \mu_i}{\sum_{j=1}^{K} \frac{1}{\delta_j}}$.

Proof Let $A_\pi(t)$, $Q_\pi(t)$, and $U_\pi(t)$ be the number of customers that have entered the system by time t, the number of customers in the system at time t, and the number of abandonments by time t, respectively. Then

$$A_\pi(t) = Q_\pi(t) + D_\pi(t) + U_\pi(t).$$

From (1), since $Q_\pi(t) \leq \sum_{j=1}^{K-1} B_j + K$ for all policies π, we have that

$$g_\pi = \limsup_{t\to\infty} \left[\frac{\mathbb{E}[A_\pi(t)]}{t} - \frac{\mathbb{E}[U_\pi(t)]}{t} - \frac{\mathbb{E}[Q_\pi(t)]}{t} \right] = \limsup_{t\to\infty} \left[\frac{\mathbb{E}[A_\pi(t)]}{t} - \frac{\mathbb{E}[U_\pi(t)]}{t} \right]. \quad (3)$$

Our model is equivalent to a model where the service requirements of successive customers at station $j \in \{1, \ldots, K\}$ are iid with mean $\frac{1}{\delta_j}$ and service rates depend only on the server with $\mu_{ij} = \mu_i \ \forall i \in \{1, \ldots . M\}$. Let $S_k = \sum_{j=1}^{K} \frac{S_{k,j}}{\delta_j}$ be the total service requirement of customer k and let $W_\pi(t) = \sum_{k=1}^{A_\pi(t)} S_k$. Let $W_{\pi,p}(t)$ be the total work $p$erformed by time t under policy π, $W_{\pi,a}(t)$ be the total unfinished service of the $a$bandoned customers by time t, and $W_{\pi,r}(t)$ be the total $r$emaining service requirement at time t for the customers that entered the system by time t and that have not abandoned. We have

$$E[W_\pi(t)] = \mathbb{E}[W_{\pi,p}(t)] + \mathbb{E}[W_{\pi,a}(t)] + \mathbb{E}[W_{\pi,r}(t)]. \quad (4)$$

Let

$$\bar{S}_0 = \begin{cases} \bar{S}, & \text{under condition (i),} \\ 1, & \text{under condition (ii).} \end{cases}$$

Since for any policy π, $\mathbb{E}[W_{\pi,r}(t)] \leq (K + \sum_{j=1}^{K-1} B_j)\bar{S}_0 \times \sum_{j=1}^{K} \frac{1}{\delta_j}$ we have

$$\lim_{t\to\infty} \frac{\mathbb{E}[W_{\pi,r}(t)]}{t} = 0. \tag{5}$$

Let π^* be the expedite policy. We have $W_{\pi^*,p}(t) = \gamma \sum_{i=1}^{M} \mu_i t$. The number of abandonments under this policy at any time t is $U_{\pi^*}(t) = 0$. This yields $\mathbb{E}[W_{\pi^*,a}(t)] = 0$ and hence using (3)-(5)

$$g_{\pi^*} = \limsup_{t\to\infty} \left[\frac{\mathbb{E}[A_{\pi^*}(t)]}{t} \right] \tag{6}$$

and

$$\lim_{t\to\infty} \frac{\mathbb{E}[W_{\pi^*}(t)]}{t} = \lim_{t\to\infty} \frac{\mathbb{E}[W_{\pi^*,p}(t)]}{t} = \gamma \sum_{i=1}^{M} \mu_i. \tag{7}$$

For all $n \geq 0$, let $Z_n = (S_{n,1}, \ldots, S_{n,K}, Y_{n,2}, \ldots, Y_{n,K})$. Since the event $\{A_\pi(t) = n\}$ for any policy π is completely determined by the random vectors $Z_1, \ldots, Z_n$ (and is independent of $Z_{n+1}, Z_{n+2}, \ldots$), $A_\pi(t)$ is a stopping time for the sequence of vectors $\{Z_n\}$. And $A_\pi(t) \leq C(t) + 1$ for all $t \geq 0$, where $C(t)$ is the number of customers departing station 1 by time t if all servers work at station 1 at all times and there is unlimited room for completed customers after station 1. It follows from the Elementary Renewal Theorem that $\lim_{t\to\infty} \frac{\mathbb{E}[C(t)]}{t} = \gamma \sum_{i=1}^{M} \mu_i \delta_1 < \infty$. And since $C(t)$ is a non-decreasing process we have $\mathbb{E}[A_\pi(t)] < \infty$. Therefore, from Wald's identity,

$$\mathbb{E}[W_\pi(t)] = \mathbb{E}\left[\sum_{k=1}^{A_\pi(t)} S_k \right] = \mathbb{E}[A_\pi(t)] \times \sum_{j=1}^{K} \frac{1}{\delta_j}. \tag{8}$$

In particular we have

$$\mathbb{E}[W_{\pi^*}(t)] = \mathbb{E}\left[\sum_{k=1}^{A_{\pi^*}(t)} S_k \right] = \mathbb{E}[A_{\pi^*}(t)] \times \sum_{j=1}^{K} \frac{1}{\delta_j}.$$

That is, (6)-(7) imply that

$$\gamma \sum_{i=1}^{M} \mu_i = \lim_{t\to\infty} \frac{\mathbb{E}[W_{\pi^*}(t)]}{t} = \lim_{t\to\infty} \frac{\mathbb{E}[A_{\pi^*}(t)]}{t} \times \sum_{j=1}^{K} \frac{1}{\delta_j} = g_{\pi^*} \times \sum_{j=1}^{K} \frac{1}{\delta_j}.$$

Therefore,

$$g_{\pi^*} = \frac{\gamma \sum_{i=1}^{M} \mu_i}{\sum_{j=1}^{K} \frac{1}{\delta_j}}.$$

For any other policy π we have

$$W_{\pi,p}(t) \leq \gamma \sum_{i=1}^{M} \mu_i t. \tag{9}$$

Next we show $\mathbb{E}[W_{\pi,a}(t)] \leq \mathbb{E}[U_\pi(t)] \times \sum_{j=1}^{K} \frac{1}{\delta_j}$. Let R_i be the i-th abandoning customer, and suppose this customer abandons while waiting for service at station L_i. Let $\widetilde{S}_{R_i,L_i}$ denote the unfinished service requirement of customer R_i at station L_i at the time of abandonment. Then,

$$W_{\pi,a}(t) = \sum_{i=1}^{U_\pi(t)} \left[\frac{\widetilde{S}_{R_i,L_i}}{\delta_{L_i}} + \sum_{j=L_i+1}^{K} \frac{S_{R_i,j}}{\delta_j} \right].$$

Thus,

$$\mathbb{E}[W_{\pi,a}(t)] = \sum_{n=1}^{\infty} \sum_{l_1,\ldots,l_n=2}^{K} \mathbb{E}\left[\sum_{i=1}^{n}\left(\frac{\widetilde{S}_{R_i,l_i}}{\delta_{l_i}} + \sum_{j=l_i+1}^{K}\frac{S_{R_i,j}}{\delta_j}\right)\middle| U_\pi(t)=n, L_1=l_1,\ldots,L_n=l_n\right]$$
$$\times P(U_\pi(t)=n, L_1=l_1,\ldots,L_n=l_n)$$
$$\overset{(*)}{\leq} \sum_{n=1}^{\infty}\sum_{l_1,\ldots,l_n=2}^{K}\sum_{i=1}^{n}\left(\frac{1}{\delta_{l_i}} + \sum_{j=l_i+1}^{K}\frac{1}{\delta_j}\right) P(U_\pi(t)=n, L_1=l_1,\ldots,L_n=l_n)$$
$$\leq \sum_{n=1}^{\infty}\sum_{l_1,\ldots,l_n=2}^{K} n\left(\sum_{j=1}^{K}\frac{1}{\delta_j}\right) P(U_\pi(t)=n, L_1=l_1,\ldots,L_n=l_n)$$
$$= \mathbb{E}[U_\pi(t)]\sum_{j=1}^{K}\frac{1}{\delta_j}. \tag{10}$$

The inequality $(*)$ holds because customers do not abandon while in service. Under condition (i), $\widetilde{S}_{R_i,L_i} = S_{R_i,L_i}$, since an abandoning customer either has not begun service at station L_i or would have to restart service. Hence $S_{R_i,j}$ for $1 \leq i \leq U_\pi(t)$ and $L_i \leq j \leq K$ are service requirements that have not been executed. Under condition (ii), service is preemptive-resume. Let V_i denote the amount of service already received by customer R_i at station L_i before abandonment. Then,

$$\widetilde{S}_{R_i,L_i} = S_{R_i,L_i} - V_i, \qquad S_{R_i,L_i} > V_i.$$

Under condition (ii),

$$\mathbb{E}[S_{R_i,L_i} - x \mid S_{R_i,L_i} > x] \leq 1.$$

Therefore,

$$\begin{aligned}
&\mathbb{E}\left[\widetilde{S}_{R_i,l_i} \mid U_\pi(t)=n, L_1=l_1,\ldots,L_n=l_n\right]\\
&= \mathbb{E}\left[\mathbb{E}\left[\widetilde{S}_{R_i,l_i} \mid V_i, U_\pi(t)=n, L_1=l_1,\ldots,L_n=l_n\right] \mid U_\pi(t)=n, L_1=l_1,\ldots,L_n=l_n\right]\\
&= \mathbb{E}\left[\mathbb{E}\left[S_{R_i,l_i} - V_i \mid S_{R_i,l_i} > V_i, V_i\right] \mid U_\pi(t)=n, L_1=l_1,\ldots,L_n=l_n\right]\\
&\leq \mathbb{E}\left[1 \mid U_\pi(t)=n, L_1=l_1,\ldots,L_n=l_n\right]\\
&= 1.
\end{aligned}$$

Now from (4), (8), (9), and (10)

$$\frac{\mathbb{E}[A_\pi(t)]}{t} \times \sum_{j=1}^{K}\frac{1}{\delta_j} \leq \gamma\sum_{i=1}^{M}\mu_i + \frac{\mathbb{E}[U_\pi(t)]}{t} \times \sum_{j=1}^{K}\frac{1}{\delta_j} + \frac{\mathbb{E}[W_{\pi,r}(t)]}{t}.$$

It follows from (5) that

$$\limsup_{t\to\infty}\left[\frac{\mathbb{E}[A_\pi(t)]}{t} - \frac{\mathbb{E}[U_\pi(t)]}{t}\right] \times \sum_{j=1}^{K}\frac{1}{\delta_j} \leq \gamma\sum_{i=1}^{M}\mu_i,$$

and (3) yields

$$g_\pi \times \sum_{j=1}^{K}\frac{1}{\delta_j} \leq \gamma\sum_{i=1}^{M}\mu_i.$$

Therefore π^* is the optimal policy. □

3 Optimal policy for 2 Stations and 2 Servers

In this section we completely characterize the optimal server assignment policy with $K = 2$ stations, $M = 2$ servers, and independent exponential service requirements at the stations. The customers waiting in line to be served by station 2 (after being served at station 1) may abandon the system after an exponential amount of time with parameter $\theta > 0$. We set $B_1 = B$ for notational convenience. We assume service preemption is allowed, and without loss of generality, that $\mu_{11}\mu_{22} \geq \mu_{12}\mu_{21}$. Let $\Sigma_j = \mu_{1j} + \mu_{2j}$ and $\Sigma'_j = \gamma\Sigma_j$ for $j = 1, 2$. We let

$$\lambda = \gamma - 1.$$

When both servers work in station j their combined service rate is Σ'_j. As in Section 2, our goal is to solve the optimization problem (2). In order to avoid trivial cases, we assume $\Sigma_j > 0$ for $j = 1, 2$. We also assume that if $\lambda = 0$, $\sum_{j=1}^{2} \mu_{ij} > 0$ for $i = 1, 2$. This is because $\lambda = 0$ and $\sum_{j=1}^{2} \mu_{ij} = 0$ would imply that the server cannot process customers on their own and there are no synergistic effects for collaboration. Hence we can simply exclude the server from the system. It can be easily shown (take $\mu_i = 0, \mu_j = \mu_{j1}$, $\delta_1 = 1$, and $\delta_2 = \mu_{j2}/\mu_{j1}$ in Theorem 1) that the expedite policy (that assigns all servers to the same job until completion) would be optimal in this case.

Let $X_\pi(t)$ be the number of customers that have been processed at station 1 and are waiting to be processed or being processed at station 2 at time t under policy $\pi \in \Pi$, where we have $X_\pi(t) \in \{0, 1, \ldots, B + 2\} =: S$. From now on, we let Π be the set of all stationary deterministic Markovian policies corresponding to the state space S. We define the action space $A := \{a_{\sigma_1\sigma_2} : \sigma_i \in \{1, 2\}, \forall i = 1, 2\}$, where $\sigma_i = j \in \{1, 2\}$ means that server i is assigned to station j. For any given π, $\{X_\pi(t)\}$ is a birth-death process. Since $B < \infty$ and the sum of all transition rates out of each state $s \in S$ is bounded by $\sum_{j=1}^{2} \Sigma'_j + (B + 2)\theta < \infty$, $\{X_\pi(t)\}$ is uniformizable with a finite uniformization constant, say q, satisfying $\sum_{j=1}^{2} \Sigma'_j + (B + 2)\theta \leq q < \infty$. Thus, the continuous-time optimization problem (2) can be translated into a discrete-time Markov decision problem by uniformization (i.e., modeling the state transitions at the event times of a Poisson process with rate q; see for example [4]). Note that the chosen value of q has no impact on our results.

We next provide for all $s \in S$ and $a \in A$, the immediate rewards, $r(s, a)$, and the transition probability $p(j|s, a)$ of going to state j from state s in one step when action a is chosen in state s. We have for all $a \in A$

$$r(0, a) = 0$$

and for all $s = 1, \ldots, B + 2$,

$$\left.\begin{array}{l} r(s, a_{11}) = 0, \\ r(s, a_{21}) = \mu_{12}, \\ r(s, a_{12}) = \mu_{22}, \\ r(s, a_{22}) = \Sigma'_2. \end{array}\right\} \tag{11}$$

Furthermore,

$$p(j|s,a_{12}) = \begin{cases} \frac{\mu_{11}}{q} & \text{for } s = 0,\ldots,B+1,\ j = s+1, \\ \frac{\mu_{22}+(s-1)\theta}{q} & \text{for } s = 1,\ldots,B+2,\ j = s-1, \\ 1 - \frac{\mu_{11}+\mu_{22}+(s-1)\theta}{q} & \text{for } s = 1,\ldots,B+1,\ j = s, \\ 1 - \frac{\mu_{22}+(s-1)\theta}{q} & \text{for } s = B+2,\ j = s, \\ 1 - \frac{\mu_{11}}{q} & \text{for } s = 0,\ j = s, \\ 0 & \text{otherwise;} \end{cases} \tag{12}$$

$$p(j|s,a_{21}) = \begin{cases} \frac{\mu_{21}}{q} & \text{for } s = 0,\ldots,B+1,\ j = s+1, \\ \frac{\mu_{12}+(s-1)\theta}{q} & \text{for } s = 1,\ldots,B+2,\ j = s-1, \\ 1 - \frac{\mu_{21}+\mu_{12}+(s-1)\theta}{q} & \text{for } s = 1,\ldots,B+1,\ j = s, \\ 1 - \frac{\mu_{12}+(s-1)\theta}{q} & \text{for } s = B+2,\ j = s, \\ 1 - \frac{\mu_{21}}{q} & \text{for } s = 0,\ j = s, \\ 0 & \text{otherwise;} \end{cases} \tag{13}$$

$$p(j|s,a_{11}) = \begin{cases} \frac{\Sigma_1'}{q} & \text{for } s = 0,\ldots,B+1,\ j = s+1, \\ \frac{s\theta}{q} & \text{for } s = 1,\ldots,B+2,\ j = s-1, \\ 1 - \frac{\Sigma_1'+s\theta}{q} & \text{for } s = 1,\ldots,B+1,\ j = s, \\ 1 - \frac{s\theta}{q} & \text{for } s = B+2,\ j = s, \\ 1 - \frac{\Sigma_1'}{q} & \text{for } s = 0,\ j = s, \\ 0 & \text{otherwise;} \end{cases} \tag{14}$$

and

$$p(j|s,a_{22}) = \begin{cases} \frac{\Sigma_2'+(s-1)\theta}{q} & \text{for } s = 1,\ldots,B+2,\ j = s-1, \\ 1 - \frac{\Sigma_2'+(s-1)\theta}{q} & \text{for } s = 1,\ldots,B+2,\ j = s, \\ 1 & \text{for } s = 0,\ j = s, \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$

We next characterize the dynamic server assignment policy that maximizes the long-run average throughput. For $n = 1,\ldots,B+2$, we define the decision rule d_n as follows

$$d_n(s) = \begin{cases} a_{11} \text{ for } s = 0, \\ a_{12} \text{ for } s = 1,\ldots,n-1, \\ a_{22} \text{ for } s = n,\ldots,B+2. \end{cases} \tag{16}$$

The next proposition gives an expression for the long-run average throughput g_n under policy d_n.

Proposition 2 *We have for* $n = 1, \ldots, B+2$

$$g_n = \frac{\beta(n)}{\Delta(n)}, \tag{17}$$

where

$$\beta(n) = [\Sigma_2' + (n-1)\theta]\Sigma_1'\mu_{22}\alpha(n) + \Sigma_1'\Sigma_2'\mu_{11}^{n-1} \tag{18}$$

and

$$\Delta(n) = [\Sigma_2' + (n-1)\theta][f(1,n) + \Sigma_1'\alpha(n)] + \Sigma_1'\mu_{11}^{n-1} \tag{19}$$

with

$$\alpha(n) = \sum_{k=2}^{n} \mu_{11}^{k-2} f(k,n) \text{ and}$$

$$f(k,n) = \prod_{j=k}^{n-1} [\mu_{22} + (j-1)\theta] \text{ for } k = 1, 2, \ldots, n,$$

and with the convention that summation over an empty set is 0 and the multiplication over an empty set is 1.

Proof For the policy d_n, states $n+1, \ldots, B+2$ are transient. For the recurrent states we have the corresponding birth rates given as

$$\kappa_i = \begin{cases} \Sigma_1' & \text{for } i = 0, \\ \mu_{11} & \text{for } i = 1, 2, \ldots, n-1, \\ 0 & \text{for } i = n, \end{cases}$$

and death rates given as

$$\nu_i = \begin{cases} \mu_{22} + (i-1)\theta & \text{for } i = 1, 2, \ldots, n-1, \\ \Sigma_2' + (i-1)\theta & \text{for } i = n. \end{cases}$$

We know that for birth-death processes with finite state space $\{0, 1, \ldots, n\}$, the stationary probabilities are given by

$$p_i = \frac{\prod_{j=0}^{i-1} \kappa_j}{\left(\prod_{j=1}^{i} \nu_i\right)\left(\sum_{k=0}^{n} \frac{\prod_{j=0}^{k-1} \kappa_j}{\prod_{j=1}^{k} \nu_j}\right)} \qquad 0 \le i \le n. \tag{20}$$

Therefore the long-run average throughput under policy d_n is given by

$$g_n = \sum_{i=1}^{n-1} \mu_{22} p_i + \Sigma_2' p_n.$$

Using (20) and some algebra we get

$$\begin{aligned} g_n =& \frac{[\Sigma_2' + (n-1)\theta]\Sigma_1'\mu_{22}\left[\sum_{l=1}^{n-1} \mu_{11}^{l-1}\left(\prod_{j=l+1}^{n-1}[\mu_{22} + (j-1)\theta]\right)\right] + \Sigma_1'\Sigma_2'\mu_{11}^{n-1}}{[\Sigma_2' + (n-1)\theta]\left[\prod_{j=1}^{n-1}[\mu_{22} + (j-1)\theta] + \Sigma_1'\left(\sum_{l=1}^{n-1} \mu_{11}^{l-1} \prod_{j=l+1}^{n-1}[\mu_{22} + (j-1)\theta]\right)\right] + \Sigma_1'\mu_{11}^{n-1}} \\ =& \frac{\beta(n)}{\Delta(n)}, \end{aligned}$$

where the last equality follows from the change of variables $k = l + 1$. □

The next proposition characterizes the difference $g_n - g_{n-1}$. The sign of this quantity will be essential in proving our results. Note that $g_n - g_{n-1}$ has the same sign as $\tau'(n)$.

Proposition 3 *We have*

$$g_1 = \frac{\tau'(1)}{\Sigma_1' + \Sigma_2'}$$

and for $n = 2, \ldots, B+2$,

$$g_n - g_{n-1} = \frac{\Sigma_1' \mu_{11}^{n-2} \tau'(n)}{\Delta(n)\Delta(n-1)},$$

where

$$\tau'(1) = \Sigma_1' \Sigma_2'$$

and for $n = 2, \ldots, B+2$,

$$\begin{aligned}\tau'(n) =& [\Sigma_2' + (n-1)\theta][\Sigma_2' + (n-2)\theta]\mu_{22} f(1, n-1) - [\Sigma_2' + (n-1)\theta][\Sigma_1' \alpha(n)(\Sigma_2' - \mu_{22}) + \Sigma_2' f(1,n)] \\ &+ [\Sigma_2' + (n-2)\theta]\mu_{11}[\Sigma_2' f(1, n-1) + \Sigma_1' \alpha(n-1)(\Sigma_2' - \mu_{22})]. \end{aligned} \tag{21}$$

Proof From Proposition 2 we immediately have

$$g_1 = \frac{\Sigma_1' \Sigma_2'}{\Sigma_1' + \Sigma_2'} = \frac{\tau'(1)}{\Sigma_1' + \Sigma_2'}.$$

Next we consider $n = 2, \ldots, B+2$. We will use the following equalities in deriving the expression for $\tau'(n)$:

$$f(k,n) = f(k, n-1)[\mu_{22} + (n-2)\theta], \quad k = 1, \ldots, n-1, \tag{22}$$

and

$$\begin{aligned}\alpha(n) &= \sum_{k=2}^{n-1} \mu_{11}^{k-2} f(k,n) + \mu_{11}^{n-2} f(n,n) \\ &= \sum_{k=2}^{n-1} \mu_{11}^{k-2} f(k, n-1)[\mu_{22} + (n-2)\theta] + \mu_{11}^{n-2} \\ &= [\mu_{22} + (n-2)\theta]\alpha(n-1) + \mu_{11}^{n-2}. \end{aligned} \tag{23}$$

Using these and some algebra, we have

$$\begin{aligned}&\beta(n)\Delta(n-1) - \beta(n-1)\Delta(n) \\ =&[\Sigma_2' + (n-1)\theta]\Sigma_1'\mu_{22}\alpha(n)[\Sigma_2' + (n-2)\theta][f(1,n-1) + \Sigma_1'\alpha(n-1)] \\ &- [\Sigma_2' + (n-2)\theta]\Sigma_1'\mu_{22}\alpha(n-1)[\Sigma_2' + (n-1)\theta][f(1,n) + \Sigma_1'\alpha(n)] \\ &+ [\Sigma_2' + (n-1)\theta]\Sigma_1'\mu_{22}\alpha(n)\Sigma_1'\mu_{11}^{n-2} - \Sigma_1'\Sigma_2'\mu_{11}^{n-2}[\Sigma_2' + (n-1)\theta][f(1,n) + \Sigma_1'\alpha(n)] \\ &- [\Sigma_2' + (n-2)\theta]\Sigma_1'\mu_{22}\alpha(n-1)\Sigma_1'\mu_{11}^{n-1} + \Sigma_1'\Sigma_2'\mu_{11}^{n-1}[\Sigma_2' + (n-2)\theta][f(1,n-1) + \Sigma_1'\alpha(n-1)] \\ =&\Sigma_1'[\Sigma_2' + (n-1)\theta][\Sigma_2' + (n-2)\theta]\mu_{22} f(1,n-1)\mu_{11}^{n-2} \\ &+ \Sigma_1'\mu_{11}^{n-2}[\Sigma_2' + (n-1)\theta]\Big[- \Sigma_1'\alpha(n)[\Sigma_2' - \mu_{22}] - \Sigma_2' f(1,n)\Big] \end{aligned}$$

$$+\Sigma_1'\mu_{11}^{n-2}[\Sigma_2'+(n-2)\theta]\mu_{11}\Big[\Sigma_2' f(1,n-1)+\Sigma_1'\alpha(n-1)(\Sigma_2'-\mu_{22})\Big]$$

$$=\Sigma_1'\mu_{11}^{n-2}\Big[[\Sigma_2'+(n-1)\theta][\Sigma_2'+(n-2)\theta]\mu_{22}f(1,n-1)-[\Sigma_2'+(n-1)\theta][\Sigma_1'\alpha(n)(\Sigma_2'-\mu_{22})+\Sigma_2' f(1,n)]$$
$$+[\Sigma_2'+(n-2)\theta]\mu_{11}[\Sigma_2' f(1,n-1)+\Sigma_1'\alpha(n-1)(\Sigma_2'-\mu_{22})]\Big]$$

$$=\Sigma_1'\mu_{11}^{n-2}\tau'(n).$$

Therefore, for $n=2,\ldots,B+2$

$$g_n-g_{n-1}=\frac{\beta(n)\Delta(n-1)-\beta(n-1)\Delta(n)}{\Delta(n)\Delta(n-1)}=\frac{\Sigma_1'\mu_{11}^{n-2}\tau'(n)}{\Delta(n)\Delta(n-1)}. \tag{24}$$

□

The next proposition provides a useful property of the $\tau'(\cdot)$ function.

Proposition 4 *For $n\in\{2,\ldots,B+2\}$ if $\tau'(n)\geq 0$, then $\tau'(k)\geq 0$ for all $1\leq k\leq n-1$. Similarly, if $\tau'(n)\leq 0$, then $\tau'(k)\leq 0$ for all $n+1\leq k\leq B+2$.*

Proof For $n\geq 2$,

$$\tau'(n+1)=[\Sigma_2'+n\theta][\Sigma_2'+(n-1)\theta]\mu_{22}f(1,n)-[\Sigma_2'+n\theta][\Sigma_1'\alpha(n+1)(\Sigma_2'-\mu_{22})+\Sigma_2' f(1,n+1)]$$
$$+[\Sigma_2'+(n-1)\theta]\mu_{11}[\Sigma_2' f(1,n)+\Sigma_1'\alpha(n)(\Sigma_2'-\mu_{22})]. \tag{25}$$

Using some algebra (details in Appendix A.2), we have for $n\geq 2$

$$\tau'(n+1)=[\mu_{22}+(n-1)\theta]\tau'(n)-\epsilon'(n), \tag{26}$$

where

$$\begin{aligned}\epsilon'(n)=&\theta[\Sigma_2'-\mu_{22}][\Sigma_2'-\mu_{22}+\theta]\mu_{22}f(1,n-1)\\&+\theta\Sigma_1'\mu_{11}^{n-1}[\Sigma_2'-\mu_{22}]+\theta[\mu_{22}+(n-1)\theta][\mu_{22}+(n-2)\theta]\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(n-1)]\\&+\theta[\mu_{22}+(n-1)\theta][\mu_{22}+(n-2)\theta][\Sigma_2'-\mu_{22}]f(1,n-1)\\&+\Sigma_1'[\Sigma_2'-\mu_{22}]\mu_{11}^{n-2}\theta[\mu_{22}+(n-1)\theta]\\&+\theta[\Sigma_2'-\mu_{22}]\mu_{11}[\Sigma_2' f(1,n-1)+\Sigma_1'(\Sigma_2'-\mu_{22})\alpha(n-1)]\\\geq&0.\end{aligned} \tag{27}$$

Also note that $\tau'(1)>0$. Hence, for $n\in\{2,\ldots,B+2\}$ if $\tau'(n)\geq 0$, then $\tau'(k)\geq 0$ for all $1\leq k\leq n-1$. And if $\tau'(n)\leq 0$, then $\tau'(k)\leq 0$ for all $n+1\leq k\leq B+2$. □

Define

$$N=\max\{n\in\{1,\ldots,B+2\}:\tau'(n)\geq 0\}. \tag{28}$$

This implies $g_N\geq g_{N-1}\geq\ldots\geq g_1$ and $g_N>g_{N+1}\ldots\geq g_{B+2}$.

From (11)-(16), the reward vector r_{d_N} and the transition probability matrix P_{d_N} corresponding to the decision rule d_N are given by

$$r_{d_N}(s) = \begin{cases} 0 & \text{for } s = 0, \\ \mu_{22} & \text{for } s = 1. \ldots, N-1, \\ \Sigma_2' & \text{for } s = N, \ldots, B+2; \end{cases} \tag{29}$$

$$P_{d_N}(s,s') = \begin{cases} \frac{\Sigma_1'}{q} & \text{for } s = 0, s' = 1, \\ \frac{\mu_{11}}{q} & \text{for } s = 1, \ldots, N-1, s' = s+1, \\ \frac{\mu_{22}+(s-1)\theta}{q} & \text{for } s = 1, \ldots, N-1, s' = s-1, \\ \frac{\Sigma_2'+(s-1)\theta}{q} & \text{for } s = N, \ldots, B+2, s' = s-1, \\ 1 - \frac{\Sigma_1'}{q} & \text{for } s = 0, s' = 0, \\ 1 - \frac{\mu_{11}+\mu_{22}+(s-1)\theta}{q} & \text{for } s = 1, \ldots, N-1, s' = s, \\ 1 - \frac{\Sigma_2'+(s-1)\theta}{q} & \text{for } s = N, \ldots, B+2, s' = s, \\ 0 & \text{otherwise.} \end{cases} \tag{30}$$

Due to the abandonments, the Markov decision problem is unichain and therefore we can use the unichain policy iteration algorithm. In order to use the algorithm we let h_N be such that

$$r_{d_N} - g_N e + (P_{d_N} - I)h_N = 0 \tag{31}$$

with $h_N(0) = 0$, where e is a column vector of ones and I is the identity matrix. The next two lemmas characterize this vector h_N. Lemma 5 gives an expression of $h_N(1)$ and the differences $h_N(s) - h_N(s-1)$ in terms of the gain g_N. It will also be helpful to have a closed-form expression of the difference, as given in Lemma 6. Note that Lemma 6 implies that for all $s = 1, \ldots, B+2$, $h_N(s) - h_N(s-1) \geq 0$. The proofs are included in the Appendix A.1.

Lemma 5 *We have*

- $$h_N(1) = \frac{q g_N}{\Sigma_1'} \tag{32}$$
- *For* $s = 1, \ldots, N$,
$$h_N(s) - h_N(s-1) = q\left[\sum_{i=0}^{s-2} \frac{g_N - \mu_{22}}{\mu_{11}^{i+1}} f(s-i, s) + \frac{g_N}{\Sigma_1' \mu_{11}^{s-1}} f(1, s)\right]. \tag{33}$$
- *For* $s = N, \ldots, B+2$,
$$h_N(s) - h_N(s-1) = \frac{q[\Sigma_2' - g_N]}{\Sigma_2' + (s-1)\theta}. \tag{34}$$

Lemma 6 *For $s = 1, \ldots, N$,*

$$h_N(s) - h_N(s-1)$$
$$= \frac{q}{\Delta(N)} \left[\Sigma_1' \mu_{11}^{N-s} [\Sigma_2' - \mu_{22}] \alpha(s) + [\Sigma_2' + (N-1)\theta] \mu_{22} \sum_{k=s+1}^{N} \mu_{11}^{k-s-1} f(k, N) f(1, s) + \Sigma_2' \mu_{11}^{N-s} f(1, s) \right]. \tag{35}$$

For $s = N, \ldots, B+2$,

$$h_N(s) - h_N(s-1) = \frac{q[\Sigma_2' + (N-1)\theta] \Big(\Sigma_2' f(1, N) + \Sigma_1' [\Sigma_2' - \mu_{22}] \alpha(N) \Big)}{\Delta(N)[\Sigma_2' + (s-1)\theta]}. \tag{36}$$

Another result that will be useful is given in the following lemma. The proof is included in the Appendix.

Lemma 7 *For $s = 1, 2, \ldots, N-1$, we have*

$$\alpha(N) \geq \mu_{11}^{N-s} \alpha(s). \tag{37}$$

We next state and prove the main result of this section which completely characterizes the optimal server assignment policy. While Proposition 4 implies that d_N (as defined in (28)) maximizes throughput within the restricted class of policies $\{d_n\}$, Theorem 8 establishes that it is, in fact, optimal across all policies.

Theorem 8 *For N as defined in (28), d_N is the optimal policy.*

The complete proof of the Theorem is included in the Appendix A.3. We use policy iteration for unichain models (see [24]). We set the initial decision rule ν_0 equal to d_N and then show that the next decision rule computed by the policy iteration algorithm is equal to d_N. That is we show that

$$\nu_1(s) \in \arg\max_{a \in A} \left\{ r(s, a) + \sum_{j \in S} p(j|s, a) h_N(j) \right\}$$

is equal to $d_N(s)$, $\forall s \in S$. We show this by showing the difference

$$\Gamma(s, a) = r(s, d_N(s)) + \sum_{j \in S} p(j|s, d_N(s)) h_N(j) - r(s, a) - \sum_{j \in S} p(j|s, a) h_N(j)$$

is non negative for all $s \in S$ and $a \in A$ using Lemmas 5 and 6. The result then follows from Theorem 8.6.6 in [24].

Theorem 8 shows that the optimal policy has a simple threshold structure. When there is no work at station 2, both servers work together at station 1. When the number of customers waiting to be processed or being processed at station 2 reaches

the threshold N, both servers work together at station 2. In the intermediate states $s = 1, \ldots, N-1$, the servers split between the two stations. Under our assumption that $\mu_{11}\mu_{22} \geq \mu_{12}\mu_{21}$, this assignment is a_{12}, meaning that server 1 works at station 1 and server 2 works at station 2. If the opposite inequality holds, then the roles of the two servers are reversed.

Theorem 8 highlights the tradeoff between choosing a larger threshold N which keeps the servers assigned more often to the stations at which they are relatively more effective, whereas choosing a smaller threshold N makes the system switch earlier to collaboration, reducing abandonments and exploiting synergy. Thus, the optimal policy need not use the entire available buffer. This suggests that in applications where buffer capacity is costly, the optimal threshold can also provide guidance for how much buffer space is actually useful.

4 Properties of the Optimal Policy

In this section we provide some conditions under which the expedite policy is the optimal policy. We also determine how the optimal buffer size N changes with the synergistic factor γ and the abandonment rate θ.

The next proposition gives a lower bound on γ that guarantees that the expedite policy is optimal. Note that from our assumption that $\mu_{11}\mu_{22} - \mu_{21}\mu_{12} \geq 0$, we have $\gamma \geq 1$. The result states that when the synergistic factor reaches a threshold, then irrespective of the abandonment rate, the servers should always work in teams to take advantage of the synergy.

Proposition 9 *(i) If*

$$\gamma > \frac{\left(\Sigma_2 + \frac{\mu_{11}\mu_{22}-\mu_{21}\mu_{12}}{\Sigma_1} - \theta\right) + \sqrt{\left(\Sigma_2 + \frac{\mu_{11}\mu_{22}-\mu_{21}\mu_{12}}{\Sigma_1} - \theta\right)^2 + 4\theta\mu_{22}}}{2\Sigma_2},$$

we have $N = 1$, i.e., the expedite policy is the optimal policy.

(ii) If $\gamma \geq 1 + \frac{\mu_{11}\mu_{22}-\mu_{21}\mu_{12}}{\Sigma_1\Sigma_2}$, we have $N = 1$, irrespective of the value of θ.

Proof (i) We have

$$\begin{aligned}
\tau'(2) &= [\Sigma_2' + \theta]\Sigma_2'\mu_{22} - [\Sigma_2' + \theta][\Sigma_1'(\Sigma_2' - \mu_{22}) + \Sigma_2'\mu_{22}] + \mu_{11}(\Sigma_2')^2 \\
&= -[\Sigma_2' + \theta][\Sigma_1'(\Sigma_2' - \mu_{22})] + \mu_{11}(\Sigma_2')^2 \\
&= -\Sigma_2'\Sigma_1'(\Sigma_2' - \mu_{22}) + (\Sigma_2')^2\mu_{11} - \theta\Sigma_1'(\Sigma_2' - \mu_{22}) \qquad (38) \\
&= -\Sigma_2'\Sigma_1'(\lambda\Sigma_2 + \mu_{12}) + (\Sigma_2')^2\mu_{11} - \theta\Sigma_1'(\Sigma_2' - \mu_{22}) \\
&= \Sigma_2'(1+\lambda)[-\lambda\Sigma_1\Sigma_2 - \mu_{12}\Sigma_1 + \Sigma_2\mu_{11}] - \theta\Sigma_1'(\Sigma_2' - \mu_{22}) \\
&= \Sigma_2'(1+\lambda)[-\lambda\Sigma_1\Sigma_2 + \mu_{11}\mu_{22} - \mu_{21}\mu_{12}] - \theta\Sigma_1'(\Sigma_2' - \mu_{22}). \qquad (39)
\end{aligned}$$

Hence we have $\tau'(2) < 0$ and $N = 1$ if and only if

$$-\Sigma_1(\Sigma_2)^2\gamma^3 + \Sigma_2(\Sigma_1\Sigma_2 + \mu_{11}\mu_{22} - \mu_{21}\mu_{12} - \theta\Sigma_1)\gamma^2 + \theta\Sigma_1\mu_{22}\gamma < 0.$$

Let $G(\gamma) = -\Sigma_1(\Sigma_2)^2\gamma^2 + \Sigma_2(\Sigma_1\Sigma_2 + \mu_{11}\mu_{22} - \mu_{21}\mu_{12} - \theta\Sigma_1)\gamma + \theta\Sigma_1\mu_{22}$. The solutions of $G(\gamma) = 0$ are

$$\frac{\left(\Sigma_2 + \frac{\mu_{11}\mu_{22}-\mu_{21}\mu_{12}}{\Sigma_1} - \theta\right) \pm \sqrt{\left(\Sigma_2 + \frac{\mu_{11}\mu_{22}-\mu_{21}\mu_{12}}{\Sigma_1} - \theta\right)^2 + 4\theta\mu_{22}}}{2\Sigma_2}.$$

Since the smaller solution is negative, we have that $G(\gamma) < 0$ if and only if

$$\gamma > \frac{\left(\Sigma_2 + \frac{\mu_{11}\mu_{22}-\mu_{21}\mu_{12}}{\Sigma_1} - \theta\right) + \sqrt{\left(\Sigma_2 + \frac{\mu_{11}\mu_{22}-\mu_{21}\mu_{12}}{\Sigma_1} - \theta\right)^2 + 4\theta\mu_{22}}}{2\Sigma_2}.$$

(ii) From (39), if $\lambda > \frac{\mu_{11}\mu_{22}-\mu_{21}\mu_{12}}{\Sigma_1\Sigma_2}$, then $\tau'(2) < 0$. And if $\lambda = \frac{\mu_{11}\mu_{22}-\mu_{21}\mu_{12}}{\Sigma_1\Sigma_2} > 0$, then our assumption that $\Sigma_1, \Sigma_2 > 0$ in Section 3 implies $\theta\Sigma_1'(\Sigma_2' - \mu_{22}) > 0$ and hence $\tau'(2) < 0$. Finally if $\lambda = \frac{\mu_{11}\mu_{22}-\mu_{21}\mu_{12}}{\Sigma_1\Sigma_2} = 0$, from our assumption in Section 3 that $\mu_{11}, \mu_{22} > 0$, we have that $\mu_{12} > 0$. As a result $\theta\Sigma_1'(\Sigma_2' - \mu_{22}) > 0$ and again $\tau'(2) < 0$. Hence we have $N = 1$.

□

Similarly, we also have a lower bound on the abandonment rate θ that guarantees that the expedite policy is optimal. Note that the bound in Proposition 10 will be negative, and hence the expedite policy is optimal for all $\theta > 0$, if

$$\begin{aligned}
0 >& \Sigma_2'\mu_{11} - \Sigma_1'(\Sigma_2' - \mu_{22}) \\
=& (1+\lambda)\Big[(\mu_{22} + \mu_{12})\mu_{11} - (\mu_{11} + \mu_{21})(\lambda\Sigma_2 + \mu_{12})\Big] \\
=& (1+\lambda)\Big[-\lambda\Sigma_1\Sigma_2 + \mu_{11}\mu_{22} - \mu_{21}\mu_{12}\Big].
\end{aligned}$$

That is the bound can be negative only if $\lambda > \frac{\mu_{11}\mu_{22}-\mu_{21}\mu_{12}}{\Sigma_1\Sigma_2}$, and hence Proposition 10 is consistent with Proposition 9 in this case.

Remark 1 Since $\Sigma_1\Sigma_2 = \mu_{11}\mu_{12} + \mu_{11}\mu_{22} + \mu_{21}\mu_{12} + \mu_{21}\mu_{22} \geq \mu_{11}\mu_{22} - \mu_{21}\mu_{12}$, from Proposition 9 (ii) we see that if $\gamma \geq 2$, the expedite policy is always optimal.

Proposition 10 *(i) If* $\theta > \frac{\Sigma_2'[\Sigma_2'\mu_{11} - \Sigma_1'(\Sigma_2' - \mu_{22})]}{\Sigma_1'(\Sigma_2' - \mu_{22})}$, *we have* $N = 1$, *i.e., the expedite policy is the optimal policy.*

(ii) If $\theta > \frac{\Sigma_2(\mu_{11}\mu_{22} - \mu_{21}\mu_{12})}{\Sigma_1\mu_{12}}$, *we have* $N = 1$, *irrespective of the value of* γ.

Proof (i) Follows from the proof of Proposition 14 in Section 6.2. In that proof, γ_1 and γ_2 denote the synergistic factors at stations 1 and 2, respectively. The present result corresponds to the special case $\gamma_1 = \gamma_2 = \gamma$.

(ii) Let

$$F(\gamma) = \frac{\Sigma_2'[\Sigma_2'\mu_{11} - \Sigma_1'(\Sigma_2' - \mu_{22})]}{\Sigma_1'(\Sigma_2' - \mu_{22})} = \frac{(\Sigma_2')^2\mu_{11}}{\Sigma_1'(\Sigma_2' - \mu_{22})} - \Sigma_2' = \frac{\gamma\Sigma_2^2\mu_{11}}{\Sigma_1(\gamma\Sigma_2 - \mu_{22})} - \gamma\Sigma_2.$$

Taking derivative with respect to γ yields

$$F'(\gamma) = \frac{\Sigma_1(\gamma\Sigma_2 - \mu_{22})\Sigma_2^2\mu_{11} - \gamma\Sigma_2^3\mu_{11}\Sigma_1}{[\Sigma_1(\gamma\Sigma_2 - \mu_{22})]^2} - \Sigma_2 = -\frac{\Sigma_2^2\mu_{11}\mu_{22}}{\Sigma_1(\gamma\Sigma_2 - \mu_{22})^2} - \Sigma_2 < 0$$

That is $F(\gamma)$ is decreasing and the maximum value is attained at $\gamma = 1$. Hence

$$\frac{\Sigma_2(\mu_{11}\mu_{22} - \mu_{21}\mu_{12})}{\Sigma_1\mu_{12}} \geq \frac{\Sigma_2'[\Sigma_2'\mu_{11} - \Sigma_1'(\Sigma_2' - \mu_{22})]}{\Sigma_1'(\Sigma_2' - \mu_{22})}.$$

□

Note that if $\Sigma_2' - \mu_{22} = 0$, from (21)-(22) we have

$$\begin{aligned}\tau'(n) =& [\mu_{22} + (n-1)\theta][\mu_{22} + (n-2)\theta]\mu_{22}f(1, n-1) - [\mu_{22} + (n-1)\theta]\mu_{22}f(1, n) \\ &+ [\mu_{22} + (n-2)\theta]\mu_{11}\mu_{22}f(1, n-1) \\ =& [\mu_{22} + (n-2)\theta]\mu_{11}\mu_{22}f(1, n-1). \end{aligned} \tag{40}$$

Therefore $\tau'(n) \geq 0$ and from (28) we have $N = B + 2$, consistent with the result in Proposition 10(i).

The next proposition compares the optimal buffer sizes for two systems with different synergistic factors γ and γ'. In order to state the result, we modify our notation. In addition to $\lambda = \gamma - 1$, we let

$$\lambda' = \gamma' - 1$$

Furthermore, consistent with (21) and (28), we define for $\xi \in \{\gamma, \gamma'\}$

$$\begin{aligned}\tau_\xi'(n) =& [\xi\Sigma_2 + (n-1)\theta][\xi\Sigma_2 + (n-2)\theta]\mu_{22}f(1, n-1) \\ &- [\xi\Sigma_2 + (n-1)\theta][\xi\Sigma_1\alpha(n)(\xi\Sigma_2 - \mu_{22}) + \xi\Sigma_2 f(1, n)] \\ &+ [\xi\Sigma_2 + (n-2)\theta]\mu_{11}[\xi\Sigma_2 f(1, n-1) + \xi\Sigma_1\alpha(n-1)(\xi\Sigma_2 - \mu_{22})]\end{aligned}$$

and

$$N_\xi = \max\{n \in \{1, \ldots, B+2\} : \tau_\xi'(n) \geq 0\}.$$

The result of Proposition 11 says that as the synergy between the servers increases, the buffer size used by the optimal policy decreases. This is intuitive because it implies that servers should take advantage of the greater synergistic factor by working in a team more often.

Proposition 11 *For $\gamma \geq \gamma' \geq 1$, fixed $\mu_{11}, \mu_{21}, \mu_{12}, \mu_{22}$, θ, and for $n = 2, \ldots, B+2$ we have $\tau_\gamma'(n) \geq 0$ implies $\tau_{\gamma'}'(n) \geq 0$. That is the optimal buffer size satisfies $N_\gamma \leq N_{\gamma'}$.*

Proof We first consider the case when $\gamma' = 1$.

Some algebra, as shown in Appendix A.4, yields,

$$\tau'_\gamma(n) = (1+\lambda)\tau'_1(n) + \delta_\lambda(n), \tag{41}$$

where

$$\begin{aligned}\delta_\lambda(n) =& -(1+\lambda)\lambda\Sigma_2\delta'(n) - (1+\lambda)\lambda\Sigma_2^2(n-2)\theta f(1,n-1)\\ &-(1+\lambda)\lambda\Sigma_2[\lambda\Sigma_2+(n-2)\theta]\Sigma_1[\alpha(n)-\mu_{11}\alpha(n-1)] - (1+\lambda)\lambda\Sigma_2\theta\Sigma_1\alpha(n)\\ &-\lambda(n-1)(n-2)\theta^2\mu_{22}f(1,n-1)\end{aligned} \tag{42}$$

and

$$\delta'(n) = \Sigma_1\alpha(n)\mu_{12} - \mu_{11}\Sigma_2 f(1,n-1) - \Sigma_1\alpha(n-1)\mu_{12}\mu_{11} + \Sigma_2\Sigma_1\alpha(n) - \Sigma_2\Sigma_1\alpha(n-1)\mu_{11}.$$

We want to show that $\delta_\lambda(n) \le 0$. Since from (37) we have $\alpha(n) \ge \mu_{11}\alpha(n-1)$, it is sufficient to show that $\delta'(n) \ge 0$. We show this by induction on n. For $n = 2$ we have

$$\begin{aligned}\delta'(2) =& \Sigma_1\alpha(2)\mu_{12} - \mu_{11}\Sigma_2 f(1,1) - \Sigma_1\alpha(1)\mu_{12}\mu_{11} + \Sigma_2\Sigma_1\alpha(2) - \Sigma_2\Sigma_1\alpha(1)\mu_{11}\\ =& \Sigma_1\mu_{12} - \mu_{11}\Sigma_2 + \Sigma_2\Sigma_1\\ =& \Sigma_1\mu_{12} + \Sigma_2\mu_{21} \ge 0.\end{aligned}$$

Next we assume that $\delta'(n) \ge 0$ and show that it implies $\delta'(n+1) \ge 0$. We have

$$\delta'(n+1) = \Sigma_1\alpha(n+1)\mu_{12} - \mu_{11}\Sigma_2 f(1,n) - \Sigma_1\alpha(n)\mu_{12}\mu_{11} + \Sigma_2\Sigma_1\alpha(n+1) - \Sigma_2\Sigma_1\alpha(n)\mu_{11}.$$

Now, from (22)-(23),

$$\begin{aligned}&\delta'(n+1)\\ =&\Sigma_1\mu_{12}[\mu_{22}+(n-1)\theta]\alpha(n) + \Sigma_1\mu_{12}\mu_{11}^{n-1} - \mu_{11}\Sigma_2[\mu_{22}+(n-2)\theta]f(1,n-1)\\ &-\Sigma_1\mu_{12}\mu_{11}[\mu_{22}+(n-2)\theta]\alpha(n-1) - \Sigma_1\mu_{12}\mu_{11}^{n-1}\\ &+\Sigma_2\Sigma_1[\mu_{22}+(n-1)\theta]\alpha(n) + \Sigma_2\Sigma_1\mu_{11}^{n-1} - \Sigma_2\Sigma_1\mu_{11}[\mu_{22}+(n-2)\theta]\alpha(n-1) - \Sigma_2\Sigma_1\mu_{11}^{n-1}\\ =&\Sigma_1\mu_{12}[\mu_{22}+(n-1)\theta]\alpha(n) - \mu_{11}\Sigma_2[\mu_{22}+(n-2)\theta]f(1,n-1) - \Sigma_1\mu_{12}\mu_{11}[\mu_{22}+(n-2)\theta]\alpha(n-1)\\ &+\Sigma_2\Sigma_1[\mu_{22}+(n-1)\theta]\alpha(n) - \Sigma_2\Sigma_1\mu_{11}[\mu_{22}+(n-2)\theta]\alpha(n-1)\\ =&[\mu_{22}+(n-2)\theta]\Big[\Sigma_1\mu_{12}\alpha(n) - \mu_{11}\Sigma_2 f(1,n-1) - \Sigma_1\mu_{12}\mu_{11}\alpha(n-1) + \Sigma_2\Sigma_1\alpha(n) - \Sigma_2\Sigma_1\mu_{11}\alpha(n-1)\Big]\\ &+\theta\Sigma_1\mu_{12}\alpha(n) + \theta\Sigma_2\Sigma_1\alpha(n)\\ =&[\mu_{22}+(n-2)\theta]\delta'(n) + \theta\Sigma_1\mu_{12}\alpha(n) + \theta\Sigma_2\Sigma_1\alpha(n) \ge 0,\end{aligned}$$

where the inequality follows from the induction assumption. Hence we have $\delta'(n) \ge 0$ for all $n \ge 2$ and therefore $\delta_\lambda(n) \le 0$. As a result from (41), we have shown that $\tau'_\gamma(n) \ge 0$ implies $\tau'_1(n) \ge 0$ and therefore $N_\gamma \le N_1$.

Next we consider the general case where the synergy for the first system is γ and for the second system is γ', where $\gamma \ge \gamma'$. From (41) we have

$$\tau'_{\gamma'}(n) = (1+\lambda')\tau'_1(n) + \delta_{\lambda'}(n).$$

After solving for $\tau'_1(n)$, (41) yields

$$\tau'_\gamma(n) = (1+\lambda)\left[\frac{\tau'_{\gamma'}(n) - \delta_{\lambda'}(n)}{(1+\lambda')}\right] + \delta_\lambda(n) = \frac{(1+\lambda)}{(1+\lambda')}\tau'_{\gamma'}(n) + \delta_\lambda(n) - \frac{(1+\lambda)}{(1+\lambda')}\delta_{\lambda'}(n).$$

In order to show $\tau'_\gamma(n) \ge 0$ implies $\tau'_{\gamma'}(n) \ge 0$, it is sufficient to show that $\delta_\lambda(n) - \frac{(1+\lambda)}{(1+\lambda')}\delta_{\lambda'}(n) \le 0$. We have from (42) that

$$\delta_\lambda(n) - \frac{(1+\lambda)}{(1+\lambda')}\delta_{\lambda'}(n)$$

$$
\begin{aligned}
&= -(1+\lambda)(\lambda-\lambda')\Sigma_2\delta'(n) - (1+\lambda)(\lambda-\lambda')\Sigma_2\Sigma_2(n-2)\theta f(1,n-1) \\
&\quad -(1+\lambda)\Sigma_2\left[(\lambda^2-(\lambda')^2)\Sigma_2 + (\lambda-\lambda')(n-2)\theta\right]\Sigma_1[\alpha(n)-\mu_{11}\alpha(n-1)] \\
&\quad -(1+\lambda)(\lambda-\lambda')\Sigma_2\theta\Sigma_1\alpha(n) - \left[\lambda - \frac{1+\lambda}{1+\lambda'}\lambda'\right](n-1)(n-2)\theta^2\mu_{22}f(1,n-1) \le 0,
\end{aligned}
$$

where in the last inequality we have used $\delta'(n) \ge 0$, $\alpha(n) \ge \mu_{11}\alpha(n-1)$, and $\lambda \ge \lambda'$.

Hence, $\tau'_\gamma(n) \ge 0$ implies $\tau'_{\gamma'}(n) \ge 0$ and therefore $N_\gamma \le N_{\gamma'}$. □

We conclude this section by stating a proposition which compares the optimal buffer size as the abandonment rate changes (including no abandonments). We let $N^{(\theta)}$ denote the optimal buffer size when the abandonment rate is θ. The following proposition states that the optimal buffer size decreases as the abandonment rate increases. This result is expected since with abandonments the work done in station 1 is wasted. Therefore, in order to maximize the throughput, the optimal policy will try to reduce the number of abandonments by reducing the number of customers waiting in line to start service in station 2.

Proposition 12 *For $\theta_1 \ge \theta_2 \ge 0$, fixed $\mu_{11}, \mu_{21}, \mu_{12}, \mu_{22}$, γ, we have the optimal buffer size $N^{(\theta_1)} \le N^{(\theta_2)}$.*

Proof First we consider the scenario where $\theta_1 > \theta_2 = 0$. Here we consider two cases.

- Case 1 : $\gamma - 1 = \lambda > \frac{\mu_{11}\mu_{22}-\mu_{21}\mu_{12}}{\Sigma_1\Sigma_2}$. In this case, Proposition 9 yields that $N^{(\theta_1)} = 1$. Moreover, we have

$$
\gamma = 1 + \lambda > 1 + \frac{\mu_{11}\mu_{22} - \mu_{21}\mu_{12}}{\Sigma_1\Sigma_2} = \frac{\mu_{11}}{\mu_{11}+\mu_{21}} + \frac{\mu_{22}}{\mu_{22}+\mu_{12}}.
$$

  From Theorem 3.1 in [7] we have that in the absence of abandonments a policy is optimal if and only if it is nonidling and fully collaborative, i.e., $N^{(0)} = 1$. We have shown that $N^{(\theta_1)} = N^{(0)}$.

- Case 2 : $\gamma - 1 = \lambda \le \frac{\mu_{11}\mu_{22}-\mu_{21}\mu_{12}}{\Sigma_1\Sigma_2}$. In this case

$$
\gamma = 1 + \lambda \le \frac{\mu_{11}}{\mu_{11}+\mu_{21}} + \frac{\mu_{22}}{\mu_{22}+\mu_{12}}
$$

  and hence from Theorem 3.1 in [7] we have that in the absence of abandonments $N^{(0)} = B + 2$. Hence from definition of N we have $N^{(\theta_1)} \le N^{(0)}$ for $\theta_1 > 0$.

Next we consider the scenerio $\theta_1 \ge \theta_2 > 0$. Note that if $\Sigma'_2 - \mu_{22} = 0$, we have from (27) that $\epsilon'(n) = 0$ and therefore (26) and (28) imply that $N^{(\theta_1)} = B+2$ for any θ_1 and the result holds trivially. Also, if $\mu_{11}\mu_{22} - \mu_{21}\mu_{12} = 0$ or $\lambda \ge \frac{\mu_{11}\mu_{22}-\mu_{21}\mu_{12}}{\Sigma_1\Sigma_2}$, we have from Proposition 9 that $N^{(\theta_1)} = 1$ for all $\theta_1 > 0$ and again the result holds trivially. Therefore we assume that $\Sigma'_2 - \mu_{22} > 0$, $\mu_{11}\mu_{22} - \mu_{21}\mu_{12} > 0$ and $\lambda < \frac{\mu_{11}\mu_{22}-\mu_{21}\mu_{12}}{\Sigma_1\Sigma_2}$. We write $\tau'(n,\theta)$ to specify the θ used in the expression for $\tau'(n)$ and show that $\tau'(n,\theta)$ has a single positive root $\theta(n)$. For $n \ge 2$ we have from (26), (27), and (39) that

$$
\tau'(n,0) = \mu_{22}^{n-2}\tau'(2,0) > 0, \tag{43}
$$

where the last inequality follows from our assumption that $\lambda < \frac{\mu_{11}\mu_{22}-\mu_{21}\mu_{12}}{\Sigma_1\Sigma_2}$. Since the highest power of θ in $f(1, n-1)$ and $\alpha(n-1)$ is $n-3$, we have from (27) that the highest power of θ in $\epsilon'(n-1)$ is $n-1$ and from (38)

$$\tau'(2,\theta) = -\Sigma_2'\Sigma_1'(\Sigma_2' - \mu_{22}) + (\Sigma_2')^2\mu_{11} - \theta\Sigma_1'(\Sigma_2' - \mu_{22}),$$

it follows from (26) that the highest power of θ in $\tau'(n,\theta)$ is θ^{n-1} and it has a negative coefficient. Hence,

$$\lim_{\theta\to\infty} \tau'(n,\theta) = -\infty,$$

and therefore for each $n \geq 2$, $\tau'(n,\theta)$ has a positive root. We show that it has exactly one positive root. We do this by showing that $g_n(\theta)$, the gain under policy d_n when the abandonment rate is θ, has the following two properties

(1) $g_n(\theta)$ is non-increasing in θ for $n \geq 1$.

(2) As θ increases, $g_n(\theta)$ decreases at least as rapidly as $g_{n-1}(\theta)$ for $n \geq 2$.

Assuming these properties, suppose there exist two positive roots of $\tau'(n,\theta)$, say $\theta^1(n)$ and $\theta^2(n)$, with $0 < \theta^1(n) < \theta^2(n)$, that is, $\tau'(n,\theta^1(n)) = \tau'(n,\theta^2(n)) = 0$. From Proposition 3 this implies that $g_n(\theta^1(n)) = g_{n-1}(\theta^1(n))$ and $g_n(\theta^2(n)) = g_{n-1}(\theta^2(n))$. However, from, property (2) above, if $g_n(\theta^1(n)) = g_{n-1}(\theta^1(n))$ and since $\theta^1(n) < \theta^2(n)$ we must have $g_n(\theta^2(n)) \leq g_{n-1}(\theta^2(n))$. Further, the inequality should be strict because otherwise we would have $g_n(\theta) = g_{n-1}(\theta)\ \forall\theta^1(n) \leq \theta \leq \theta^2(n)$, which cannot be true since $\tau'(n,\theta)$ is a polynomial in θ and hence cannot satisfy $\tau'(n,\theta) = 0\ \forall\theta^1(n) \leq \theta \leq \theta^2(n)$.

We have shown that there exists a single $\theta(n) > 0$ such that $\tau'(n,\theta(n)) = 0$. From (26) this implies that

$$\tau'(n+1,\theta(n)) = [\mu_{22} + (n-1)\theta]\tau'(n,\theta(n)) - \epsilon'(n,\theta(n)) < 0,$$

where the last inequality follows from (27) and $\Sigma_2' - \mu_{22} > 0$. This together with the fact that $\tau'(n,0) > 0$ from (43) implies that $\theta(n+1) < \theta(n)$. For convenience, define $\theta(1) = \infty$ and $\theta(B+3) = 0$. Therefore for $\theta_1 \geq \theta_2 > 0$ we can have the following scenerios :

(i) $\exists n \in \{1,\ldots,B+2\}$ such that $\theta(n+1) < \theta_2 \leq \theta_1 \leq \theta(n)$.
In this case we have $N^{(\theta_1)} = N^{(\theta_2)} = n$.

(ii) $\exists n \in \{2,\ldots,B+2\}$ such that $\theta_2 \leq \theta(n) < \theta_1$.
In this case we have $N^{(\theta_1)} < n$ and $N^{(\theta_2)} \geq n$.

Hence, in both cases we have $N^{(\theta_1)} \leq N^{(\theta_2)}$.

It remains to show that properties $(1)-(2)$ hold for $g_n(\theta)$. We start by showing property (1). Note that from Proposition 3, $g_1(\theta) = \frac{\Sigma_1'\Sigma_2'}{\Sigma_1'+\Sigma_2'}$, that is, $g_1(\theta)$ is independent of θ. Now consider two systems, System 1 and System 2, where both the systems are operating under the policy d_n but the abandonment rate for System 1 is θ^1 and that of System 2 is θ^2 with $\theta^1 > \theta^2$. Consider sample paths of both systems starting in the same state and generated using (the same) independent Poisson processes with rates $\Sigma_1', \Sigma_2', (n-1)\theta^1$, and thinning. Due to the higher abandonment rate of System 1, the number of customers in System 2 will be always greater than or equal to the number of customers in System 1. In particular, when both systems are in the same state, every abandonment from (service completion at station 1 in) System 2 is associated with a simultaneous abandonment from (service completion at station 1 in) System 1. Furthermore, every service completion at station 2 in System 1 is associated with a simultaneous service completion at station 2 in System 2. This shows that more customers complete service at station 2 in System 2 and $g_n(\theta)$ is non-increasing in θ.

Next we show property (2). Consider two systems, System 1 and System 2, now with the same abandonment rate θ, but System 1 operates under policy d_n and System 2 under policy d_{n-1}. Consider sample paths of both systems generated using (the same) independent Poisson processes with rates $\Sigma_1', \Sigma_2', (n-1)\theta^1$, and thinning and starting from the same state $s \leq n-1$. Both systems evolve identically until state $n-1$ is reached. If there is an abandonment in state $n-1$ both systems experience the abandonment, but it is possible that in state $n-1$, (i) System 1 (but not System 2) can transition to state n due to a service completion at station 1 or (ii) System 2 (but not necessarily System 1) will transition to state $n-2$ due to a service completion at station 2. If this happens, then from this point onward, the number of customers in System 1 will dominate the number of customers in System 2. As a result every abandonment in System 2 is associated with a simultaneous abandonment in System 1. However, when the two systems are not in the same state, there may be abandonments in System 1 that do not occur in System 2. Some of these additional abandonments in System 1 may be replaced by additional service completions at station 1 in System 1 but not in System 2 (if both systems are in state $n-1$); others may not be replaced. As θ increases, these unreplaced additional abandonments also increase and since all service completions at station 1 lead to either abandonments or service completions at station 2, this implies that as θ increases, $g_n(\theta)$ decreases at least as rapidly as $g_{n-1}(\theta)$. This proves the two properties completing the proof. □

5 Numerical Results

In this section, we provide a numerical study to illustrate the theoretical results derived in Sections 3 and 4. Specifically, we examine the behavior of the long-run average throughput g_n as a function of n, see how the optimal buffer size N changes as a function of the synergistic factor γ and abandonment rate θ, analyze the distribution of the optimal buffer sizes across a wide range of parameter sets, and evaluate the relative performance of the optimal policy compared to the simple expedite policy for different domains of the system parameters.

First, we investigate how the system throughput g_n varies with respect to n. In Proposition 4, we showed that there exists an N such that $g_N \geq g_{N-1} \geq \cdots \geq g_1$ and $g_N > g_{N+1} \geq \cdots \geq g_{B+2}$, i.e, the long-run average throughput g_n for policy d_n (as defined in (16)) first increases with respect to the buffer size, reaches a maximum, and then decreases. Figure 1 illustrates this behavior for different sets of parameters chosen to show the different shapes that g_n can take under Proposition 4. Note that N can be equal to 1, in which case the throughput always decreases in n, or N can be equal to $B+2$ in which case the throughput always increases in n (as shown in Figure 1).

We next examine how the optimal buffer size behaves as a function of the synergistic factor γ and abandonment rate θ. From Propositions 11 and 12, we know that the optimal buffer size N decreases with respect to both θ and γ. We observe this behavior in Figure 2, where $\mu_{11} = 10, \mu_{12} = 2, \mu_{21} = 2, \mu_{22} = 13$, and $B = 12$. Over the plotted range of (γ, θ), the optimal buffer size decreases from the upper boundary $B+2$ (here, 14) for small γ and θ to the lower boundary 1 for sufficiently large γ and θ. We also observe that when θ is close to 0, N exhibits a sharp drop once γ crosses a threshold. However, the drop is more gradual as θ increases.

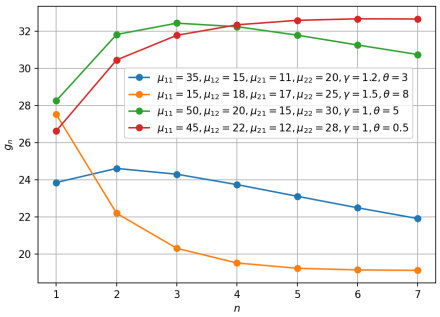


Fig. 1: Behavior of the long-run average throughput g_n as a function of n ($B = 5$).

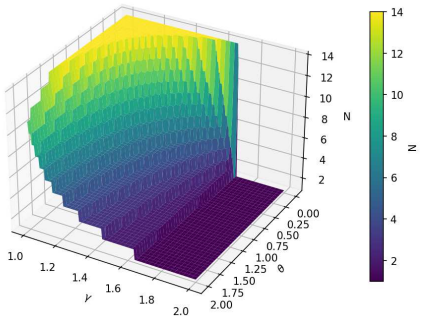


Fig. 2: Behavior of N as a function of $\gamma \in [1, 2]$ and $\theta \in [0, 2]$.

To understand the prevalence of different optimal policies in a general setting, we conducted a numerical experiment consisting of 10,000 randomly generated parameter instances from the ranges given in Table 1. For each instance, we computed the optimal buffer size N that maximizes the throughput. The frequency of different N values is given in Table 2, and a graphical representation is provided in Figure 3.

Parameter	Range
μ_{ij}	[0,400]
γ	[1,2]
θ	[0,10]
B	$\{0, \ldots, 100\}$

Table 1: Ranges of Parameters.

N value	Frequency
1	7362
2	338
3	354
4	348
5	256
6	182
7	158
8	128
9	104
10	82
11	69
12	54
13	47
14	46
15	44
16	30
17	29
18	27
19	25
20	19
21	19
22	20
23	17
24	13
25	15
26	12
27	8
28	5
29	10
30	5

N Value	Frequency
31	6
32	5
33	5
34	7
35	5
36	6
37	8
38	2
39	4
40	3
41	5
42	4
43	8
44	3
45	7
46	3
47	2
48	4
49	4
50	2
51	3
52	2
53	4
54	2
55	3
56	2
57	1
58	3
59	3
60	1

N Value	Frequency
61	2
62	5
63	1
64	1
66	1
67	3
69	1
70	4
71	4
72	1
73	1
74	2
75	1
76	1
79	3
80	5
82	2
83	3
85	1
88	1
91	3
92	3
94	2
96	2
97	1
99	2
100	1

Table 2: Frequency of N values.

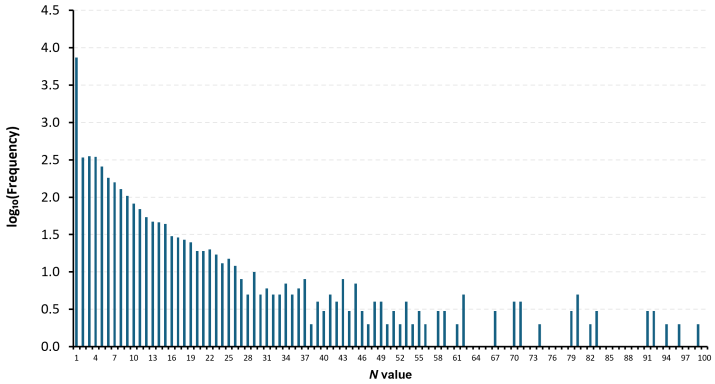


Fig. 3: Frequency distribution of N (log scale).

The results given in Table 2/Figure 3 indicate that the expedite policy ($N = 1$) is optimal in a significant proportion of the generated cases (73.62%). This could be due in part to the fact that the possible values of N for a buffer size of B are $\{1, 2, \ldots, B + 2\}$, and therefore for the system to have a high N value, the randomly chosen B should also be high.

Given the high frequency of the expedite policy's optimality observed in Table 2, it is practically useful to determine the "penalty" for using the expedite policy when it is not optimal. We compare the optimal throughput g_N to the throughput of the expedite policy g_1. We perform 10,000 random replications of the parameters from different parameter ranges as given in Table 3 and estimate the relative performance of the expedite policy via the ratio

$$\frac{\sum_{i=1}^{10,000} g_1^{(i)}}{\sum_{i=1}^{10,000} g_N^{(i)}},$$

where $g_1^{(i)}$ and $g_N^{(i)}$ correspond to the throughputs of the expedite and the optimal policies, respectively, for the i^{th} parameter instance.

Table 3 highlights the patterns in the relative performance of the expedite policy. Firstly, we observe that for each fixed choice of the service-rate ranges (i.e., within each of the three μ_{ij} range blocks), the column with $0 \leq \theta \leq 1$ is consistently lower than the corresponding column with $0 \leq \theta \leq 10$ across all values of γ. Thus, restricting the abandonment rate θ to the smaller range makes the expedite policy marginally less competitive relative to the optimal policy. This behavior is expected from Proposition 12. Further, we can observe the dependence of performance on the degree of server specialization. In parameter regimes where the service rates across stations are broadly comparable (e.g., μ_{11}, μ_{22} drawn from a similar range as μ_{12}, μ_{21}), the expedite policy performs very well, even for moderate values of the synergistic factor γ. In contrast, when servers are highly specialized (characterized by large μ_{11}, μ_{22} as compared to μ_{12}, μ_{21}), the performance of the expedite policy deteriorates substantially, especially

γ	Ranges of Parameters					
	$0 \le \mu_{11}, \mu_{22} \le 400$ $0 \le \mu_{12}, \mu_{21} \le 400$		$200 \le \mu_{11}, \mu_{22} \le 400$ $0 \le \mu_{12}, \mu_{21} \le 200$		$200 \le \mu_{11}, \mu_{22} \le 400$ $0 \le \mu_{12}, \mu_{21} \le 20$	
	$0 \le \theta \le 10$	$0 \le \theta \le 1$	$0 \le \theta \le 10$	$0 \le \theta \le 1$	$0 \le \theta \le 10$	$0 \le \theta \le 1$
1.0	0.856	0.839	0.722	0.701	0.588	0.570
1.1	0.909	0.896	0.782	0.761	0.638	0.618
1.2	0.947	0.938	0.839	0.820	0.686	0.665
1.3	0.971	0.966	0.892	0.876	0.733	0.712
1.4	0.986	0.983	0.936	0.924	0.780	0.758
1.5	0.994	0.992	0.967	0.960	0.826	0.805
1.6	0.997	0.997	0.986	0.982	0.871	0.851
1.7	0.999	0.998	0.995	0.994	0.914	0.897
1.8	0.999	0.999	0.999	0.998	0.955	0.942
1.9	0.999	0.999	0.999	0.999	0.990	0.985
2.0	1.0	1.0	1.0	1.0	1.0	1.0

Table 3: Relative performance of the expedite policy vs. the optimal policy ($B = 100$).

for low to moderate values of γ. In these regimes, the expedite policy may divert capacity away from tasks for which servers are significantly more efficient, leading to pronounced throughput losses. This reinforces the practical relevance of computing or approximating the optimal policy in such systems. While the expedite policy offers simplicity and performs well in many cases, Table 3 shows that its use can entail substantial throughput penalties in structurally asymmetric systems.

6 Task Dependent Synergy

In this section, we generalize the models in Sections 2, 3, and 4 by allowing different synergistic factors for tasks processed at different stations. That is, collaboration is more effective at some stations than others. We first show in Section 6.1 that in the case of systems with K stations and M generalist servers, the expedite policy is the optimal policy. Next for systems with 2 servers and 2 stations and non-generalist servers, we propose a dynamic server assignment policy in Section 6.2 and provide numerical results that suggest that the proposed policy is indeed the policy that achieves the maximum throughput in Section 6.3.

6.1 Generalist Servers with K Stations and M Servers

We consider the same setting as in Section 2 but now the synergy among the servers is allowed to be task (station) dependent. We show that even under this generalization, the expedite policy continues to be optimal.

Theorem 13 *Assume that either*

1. *Service is non-preemptive or preemptive-repeat and for all $t \geq 0$, if there is a customer in service at station j at time t, then the expected remaining service requirement at station j of that customer is bounded by a scalar $1 \leq \bar{S} < \infty$, or*
2. *Service is preemptive-resume and the service requirements are new better than used in expectation, i.e.,*

$$\mathbb{E}[S_{k,j} - x \mid S_{k,j} > x] \leq \mathbb{E}[S_{k,j}] = 1$$

for all $x \geq 0$, for each $j = 1, \ldots, K$.

When $\mu_{ij} = \mu_i \delta_j$ for all $i = 1, \ldots, M$ and $j = 1, \ldots, K$ and $\gamma_j \geq 1$ is the synergistic factor at station $j = 1, \ldots, K$, then the expedite policy is optimal with long-run average throughput $g_{\pi^*} = \frac{\sum_{i=1}^{M} \mu_i}{\sum_{j=1}^{K} \frac{1}{\delta_j \gamma_j}}$.

Proof Consider a model $\mathcal{H}$ where the service rate of server i at station j is $\mu_{ij} = \mu_i \delta_j \gamma_j$ and there is no synergy among the servers. This is equivalent to a model where the service requirements of successive customers at station $j \in \{1, \ldots, K\}$ are iid with mean $\frac{1}{\delta_j \gamma_j}$, the service rates depend only on the server with $\mu_{ij} = \mu_i \ \forall i \in \{1, \ldots .M\}$, and $\gamma = 1$. Now suppose the servers can also choose to operate at full efficiency at station j with $\mu_{ij} = \mu_i$ or partial efficiency with $\mu_{ij} = \mu_i / \gamma_j$. Let π^*_+ be the policy where all servers always work in a single team at full efficiency, and the team follows each customer from the first to the last station and only starts work on a new customer once all work on the previous customer has been completed. Since (9) is still satisfied (with $\gamma = 1$) for $\mathcal{H}$ (because partial efficiency can only reduce the total work performed), we have from the proof of Theorem 1 with $\gamma = 1$ and δ_j replaced by $\delta_j \gamma_j$ for $j \in \{1, \ldots, K\}$ that π^*_+ is the optimal policy in model $\mathcal{H}$ with the long-run average throughput $g^{\mathcal{H}}_{\pi^*_+} = \frac{\sum_{i=1}^{M} \mu_i}{\sum_{j=1}^{K} \frac{1}{\delta_j \gamma_j}}$. That is, if π' is any other policy in $\mathcal{H}$ and $g^{\mathcal{H}}_{\pi'}$ is the long-run average throughput under that policy we have $g^{\mathcal{H}}_{\pi^*_+} \geq g^{\mathcal{H}}_{\pi'}$.

Now consider our original model where the service rate of server i at station j is $\mu_{ij} = \mu_i \delta_j$ and the synergistic factor at station j is γ_j. Let π^* be the expedite policy. Since all servers work in a single team at all times, the long-run average throughput for the expedite policy in this model will be the same as the long-run average throughput for the π^*_+ policy in model $\mathcal{H}$. That is we have $g_{\pi^*} = g^{\mathcal{H}}_{\pi^*_+}$.

Now for any policy π in the original model we can get a policy π' in our modified model $\mathcal{H}$ such that $g^{\mathcal{H}}_{\pi'} = g_{\pi}$. This is because when a server i is working on his own at a station j in π, we can make the server i work at station j with partial efficiency in the modified model $\mathcal{H}$. And since we have shown that for all π'

$$g_{\pi^*} = g^{\mathcal{H}}_{\pi^*_+} \geq g^{\mathcal{H}}_{\pi'} = g_{\pi},$$

we have that the expedite policy is the optimal policy for our original model. □

6.2 Proposed Policy for 2 Stations and 2 Servers

We consider the same settings as in Section 3 except that the synergy is task (station) dependent, that is, the synergistic factor for station 1 is γ_1 and that for station 2 is γ_2. For $j = 1, 2$, let

$$\bar{\Sigma}'_j = \gamma_j(\mu_{1j} + \mu_{2j}) = \gamma_j \Sigma_j$$

and adapt (21) and (28) as follows

$$\begin{aligned}\bar{\tau}'(n) =&[\bar{\Sigma}_2' + (n-1)\theta][\bar{\Sigma}_2' + (n-2)\theta]\mu_{22} f(1, n-1) \\ &- [\bar{\Sigma}_2' + (n-1)\theta][\bar{\Sigma}_1'\alpha(n)(\bar{\Sigma}_2' - \mu_{22}) + \bar{\Sigma}_2' f(1,n)] \\ &+ [\bar{\Sigma}_2' + (n-2)\theta]\mu_{11}[\bar{\Sigma}_2' f(1, n-1) + \bar{\Sigma}_1'\alpha(n-1)(\bar{\Sigma}_2' - \mu_{22})]; \end{aligned} \tag{44}$$

$$\bar{N} = \max\{n \in \{1, 2, \ldots, B+2\} : \bar{\tau}'(n) \geq 0\}.$$

Our proposed policy is very similar to the optimal policy d_N defined in Section 3. In particular, we define decision rule $d_{\bar{N}}$ to be d_n as in (16) with $n = \bar{N}$. The immediate reward $\bar{r}_{d_{\bar{N}}}$, transition probability matrix $\bar{P}_{d_{\bar{N}}}$ and the gain $\bar{g}_{\bar{N}}$ under this policy are given as in (29), (30), and (17), respectively, with Σ_i' replaced by $\bar{\Sigma}_i'$ for $i = 1, 2$ and N replaced by $\bar{N}$. The next proposition gives a lower bound on θ, that guarantees $\bar{N} = 1$.

Proposition 14 *If* $\theta > \frac{\bar{\Sigma}_2'[\bar{\Sigma}_2'\mu_{11} - \bar{\Sigma}_1'(\bar{\Sigma}_2' - \mu_{22})]}{\bar{\Sigma}_1'(\bar{\Sigma}_2' - \mu_{22})}$, *we have* $\bar{N} = 1$.

Proof From (44) we have

$$\begin{aligned}\bar{\tau}'(2) =&[\bar{\Sigma}_2' + \theta]\bar{\Sigma}_2'\mu_{22} - [\bar{\Sigma}_2' + \theta][\bar{\Sigma}_1'(\bar{\Sigma}_2' - \mu_{22}) + \bar{\Sigma}_2'\mu_{22}] + \mu_{11}(\bar{\Sigma}_2')^2 \\ =& - [\bar{\Sigma}_2' + \theta]\bar{\Sigma}_1'(\bar{\Sigma}_2' - \mu_{22}) + \mu_{11}(\bar{\Sigma}_2')^2 \\ =& - \bar{\Sigma}_2'\bar{\Sigma}_1'(\bar{\Sigma}_2' - \mu_{22}) + (\bar{\Sigma}_2')^2\mu_{11} - \theta\bar{\Sigma}_1'(\bar{\Sigma}_2' - \mu_{22}).\end{aligned}$$

When $\theta > \frac{\bar{\Sigma}_2'[\bar{\Sigma}_2'\mu_{11} - \bar{\Sigma}_1'(\bar{\Sigma}_2' - \mu_{22})]}{\bar{\Sigma}_1'(\bar{\Sigma}_2' - \mu_{22})}$, we have $\bar{\tau}'(2) < 0$. Exactly as in the proof of Proposition 4, it can be shown that if $\bar{\tau}'(n) \geq 0$, then $\bar{\tau}'(k) \geq 0$ for all $1 \leq k \leq n-1$. Similarly if $\bar{\tau}'(n) \leq 0$, then $\bar{\tau}'(k) \leq 0$ for all $n+1 \leq k \leq B+2$. Therefore, $\bar{N} = 1$. □

Note that if $\bar{\Sigma}_2' - \mu_{22} = 0$ (i.e. $\mu_{12} = 0$ and $\gamma = 1$), from (44) we have

$$\begin{aligned}\bar{\tau}'(n) =&[\mu_{22} + (n-1)\theta][\mu_{22} + (n-2)\theta]\mu_{22} f(1, n-1) - [\mu_{22} + (n-1)\theta]\mu_{22} f(1,n) \\ &+ [\mu_{22} + (n-2)\theta]\mu_{11}\mu_{22} f(1, n-1) \\ \geq& 0,\end{aligned}$$

where the inequality follows from (40), and hence we have $\bar{N} = B+2$, consistent with the result of Proposition 14.

6.3 Numerical Results for 2 Stations and 2 Servers

In this section, we present numerical results that suggest that the server assignment policy $d_{\bar{N}}$ appears to be optimal with task dependent synergy. Let $h_{\bar{N}}$ be the vector solving

$$r_{d_{\bar{N}}} - \bar{g}_{\bar{N}} e + (\bar{P}_{d_{\bar{N}}} - I)h_{\bar{N}} = 0$$

with $h_{\bar{N}}(0) = 0$. As in Section 3, in order to show that $d_{\bar{N}}$ is the optimal policy using policy iteration for unichain models, we set the initial decision rule ν_0 equal to $d_{\bar{N}}$ and then show that the next decision rule computed by the policy iteration algorithm is equal to $d_{\bar{N}}$. That is we show that

$$\nu_1(s) \in \arg\max_{a \in A} \left\{ r(s,a) + \sum_{j \in S} p(j|s,a) h_{\bar{N}}(j) \right\}$$

is equal to $d_{\bar{N}}(s)$ $\forall s \in S$. We show this by showing the difference

$$\bar{\Gamma}(s,a) = r(s, d_{\bar{N}}(s)) + \sum_{j \in S} p(j|s, d_{\bar{N}}(s)) h_{\bar{N}}(j) - r(s,a) - \sum_{j \in S} p(j|s,a) h_{\bar{N}}(j) \qquad (45)$$

is non negative, $\forall s \in S$ and $a \in A$. The difference (45) was calculated for all states s and actions a for 10,000 replications where the parameters have been chosen independently and uniformly at random from the ranges given in Table 4. All the numerically computed values of $\bar{\Gamma}$ are non-negative for all $s \in S$ and $a \in A$ and for all 10,000 replications suggesting that $d_{\bar{N}}$ is the optimal policy.

Parameter	Range
$\mu_{11}, \mu_{22}, \mu_{12}, \mu_{21}$	[0,400]
γ_1, γ_2	[1,2]
θ	[0,10]
B	$\{0, 1, \ldots, 100\}$

Table 4: Ranges of Parameters.

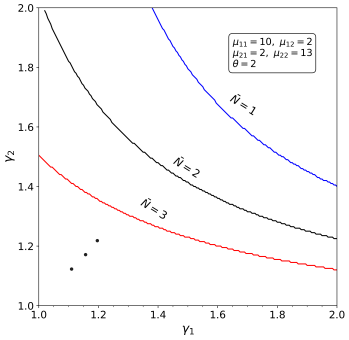


Fig. 4: Domains of the optimal buffer size $\bar{N}$ as functions of γ_1 and γ_2.

The frequency of the observed $\bar{N}$ values for the parameter ranges in Table 4 is given in Appendix A.5. We see that for more than 77.02% of the cases, the expedite policy is optimal and the frequency of the $\bar{N}$ value decreases for higher $\bar{N}$ values, just as in Table 2/Figure 3.

Figure 4 shows the domains of $\bar{N}$ as functions of γ_1 and γ_2 for parameter values $\mu_{11} = 10, \mu_{12} = 2, \mu_{21} = 2, \mu_{22} = 13$ and $\theta = 2$. It partitions the plane $\gamma_1 \geq 1, \gamma_2 \geq 1$ into regions based on the value of $\bar{N}$. The figure suggests that as the values of the synergistic factors γ_1 and γ_2 increase, the buffer size used by the proposed optimal policy decreases, as expected from Proposition 11. This extends Theorem (8), which characterizes the behavior along the line $\gamma_1 = \gamma_2$, to the region $\gamma_1, \gamma_2 \geq 1$.

7 Conclusion

In this paper, we considered tandem queueing lines with flexible synergistic servers, finite buffers, and impatient customers. We showed that for systems where each server is equally skilled at all tasks, the expedite policy achieves the maximum long-run average throughput. For two-station, two-server Markovian systems, we show that the optimal policy does not necessarily use the full buffer. Instead, there exists a threshold depending on the synergistic factor, the abandonment rate, and the service rates of the servers, such that when the number of customers waiting to be processed or being processed at station 2 reaches the threshold the optimal policy switches from having the servers working at their preferred stations to both of them working at station 2. We further showed that as the synergistic factor and abandonment rate increase, the buffer space used by the optimal policy decreases and ultimately the expedite policy becomes the optimal policy. Finally, we investigate systems where the synergy can be task dependent, propose an optimal policy, and numerically verify its optimality.

There are several natural directions for future work. First, one could allow $\gamma < 1$, which would capture settings in which collaboration creates coordination losses rather than synergistic gains. Second, one could relax the linear multiplicative structure used here, under which the service rate of a team is a constant multiple of the sum of the individual service rates. A more general model could allow the service rate generated by a team to be an arbitrary function of the servers assigned to a station, thereby capturing diminishing returns and other forms of collaboration.

Funding Declaration

This work was supported by NSF under grant 2127778. The second author was also supported by NSF under grant 2348409.

References

[1] Andradóttir, S., Ayhan, H.: Throughput maximization for tandem lines with two stations and flexible servers. Operations Research **53**(3), 516–531 (2005) https://doi.org/10.1287/opre.1040.0200

[2] Wu, C.-H., Lewis, M.E., Veatch, M.: Dynamic allocation of reconfigurable resources in a two-stage tandem queueing system with reliability considerations. IEEE Transactions on Automatic Control **51**(2), 309–314 (2006) https://doi.org/10.1109/TAC.2005.863518

[3] Ahn, H.-S., Righter, R.: Dynamic load balancing with flexible workers. Advances in Applied Probability **38**(3), 621–642 (2006) https://doi.org/10.1239/aap/1158684995

[4] Andradóttir, S., Ayhan, H., Down, D.G.: Server assignment policies for maximizing the steady-state throughput of finite queueing systems. Management Science **47**(10), 1421–1439 (2001) https://doi.org/10.1287/mnsc.47.10.1421.10262

[5] Andradóttir, S., Ayhan, H., Down, D.G.: Dynamic assignment of dedicated and flexible servers in tandem lines. Probability in the Engineering and Informational Sciences **21**(4), 497–538 (2007) https://doi.org/10.1017/S0269964807000290

[6] Kaufman, D.L., Ahn, H.-S., Lewis, M.E.: On the introduction of an agile, temporary workforce into a tandem queueing system. Queueing Systems **51**(1-2), 135–171 (2005) https://doi.org/10.1007/s11134-005-2441-x

[7] Andradóttir, S., Ayhan, H., Down, D.G.: Queueing systems with synergistic servers. Operations Research **59**(3), 772–780 (2011) https://doi.org/10.1287/opre.1110.0934

[8] Wang, X., Andradóttir, S., Ayhan, H.: Dynamic server assignment with task-dependent server synergy. IEEE Transactions on Automatic Control **60**(2), 570–575 (2015) https://doi.org/10.1109/TAC.2014.2328951

[9] Gardner, K., Righter, R.: The cost of collaboration. Queueing Systems: Theory and Applications **100**(1), 7–40 (2022) https://doi.org/10.1007/s11134-021-09721-y

[10] Fralix, B.H.: On the time-dependent moments of Markovian queues with reneging. Queueing Systems **75**(2-4), 149–168 (2013) https://doi.org/10.1007/s11134-012-9281-2

[11] Brandt, A., Brandt, M.: Workload and busy period for $M/GI/1$ with a general impatience mechanism. Queueing Systems **75**(2-4), 189–209 (2013) https://doi.org/10.1007/s11134-013-9373-7

[12] Jouini, O., Dallery, Y.: Monotonicity properties for multiserver queues with reneging and finite waiting lines. Probability in the Engineering and Informational Sciences **21**(3), 335–360 (2007) https://doi.org/10.1017/S0269964807000010

[13] Barrer, D.Y.: Queuing with impatient customers and ordered service. Operations Research **5**(5), 650–656 (1957) https://doi.org/10.1287/opre.5.5.650

[14] Baccelli, F., Boyer, P.E., Hébuterne, G.: Single-server queues with impatient customers. Advances in Applied Probability **16**(4), 887–905 (1984) https://doi.org/10.2307/1427345

[15] Down, D.G., Koole, G., Lewis, M.E.: Dynamic control of a single-server system with abandonments. Queueing Systems **67**(1), 63–90 (2011) https://doi.org/10.1007/s11134-010-9201-2

[16] Bhulai, S., Blok, H.J., Spieksma, F.M.: K competing queues with customer abandonment: Optimality of a generalised $c\mu$-rule by the smoothed rate truncation method. Annals of Operations Research **317**, 387–416 (2022) https://doi.org/10.1007/s10479-019-03131-3

[17] Atar, R., Giat, C., Shimkin, N.: The $c\mu/\theta$ rule for many-server queues with abandonment. Operations Research **58**(5), 1427–1439 (2010) https://doi.org/10.1287/opre.1100.0826

[18] Long, Z., Shimkin, N., Zhang, H., Zhang, J.: Dynamic scheduling of multiclass many-server queues with abandonment: The generalized $c\mu/h$ rule. Operations Research **68**(4), 1218–1230 (2020) https://doi.org/10.1287/opre.2019.1908

[19] Batt, R.J., Terwiesch, C.: Waiting patiently: An empirical study of queue abandonment in an emergency department. Management Science **61**(1), 39–59 (2015) https://doi.org/10.1287/mnsc.2014.2058

[20] Wang, J., Abouee-Mehrizi, H., Baron, O., Berman, O.: Tandem queues with impatient customers. Performance Evaluation **135** (2019) https://doi.org/10.1016/j.peva.2019.102011 . Article no. 102011

[21] Zayas-Cabán, G., Xie, J., Green, L.V., Lewis, M.E.: Dynamic control of a tandem system with abandonments. Queueing Systems **84**(3-4), 279–293 (2016) https://doi.org/10.1007/s11134-016-9489-7

[22] Zayas-Cabán, G., Cochran, A.L.: Scheduling servers in a two-stage queue with abandonments and costs. Probability in the Engineering and Informational Sciences **37**(3), 833–851 (2023) https://doi.org/10.1017/S0269964822000213

[23] Andradóttir, S., Ayhan, H.: Optimal server assignment in queues with flexible servers and abandonments. Stochastic Systems **13**(3), 360–376 (2023) https://doi.org/10.1287/stsy.2023.0109

[24] Puterman, M.L.: Markov Decision Processes: Discrete Stochastic Dynamic Programming, 1st edn. John Wiley & Sons, Inc., USA (1994)

Appendix

A.1 Details of Proof of Lemma 5, Lemma 6 and Lemma 7

Proof of Lemma 5 From (31), we have for $s = 0$,

$$0 - g_N - \frac{\Sigma_1'}{q} h_N(0) + \frac{\Sigma_1'}{q} h_N(1) = 0. \tag{46}$$

Next, if $N \geq 2$, then for $s = 1, \ldots N-1$,

$$\mu_{22} - g_N + \frac{\mu_{22} + (s-1)\theta}{q} h_N(s-1) - \frac{\mu_{11} + \mu_{22} + (s-1)\theta}{q} h_N(s) + \frac{\mu_{11}}{q} h_N(s+1) = 0. \tag{47}$$

And for $s = N, \ldots, B+2$,

$$\Sigma_2' - g_N + \frac{\Sigma_2' + (s-1)\theta}{q} h_N(s-1) - \frac{\Sigma_2' + (s-1)\theta}{q} h_N(s) = 0. \tag{48}$$

From (46) and $h_N(0) = 0$, we have

$$h_N(1) = \frac{g_N q}{\Sigma_1'}. \tag{49}$$

Therefore, (32) and (33) hold for $s = 1 \leq N$. Next, if $N \geq 2$, we prove by induction that (33) holds for $s = 2, \ldots, N$. Plugging $s = 1$ into (47) yields

$$\frac{\mu_{11}}{q}[h_N(2) - h_N(1)] - \frac{\mu_{22}}{q} h_N(1) = g_N - \mu_{22}.$$

That is

$$h_N(2) - h_N(1) = q\left[\frac{g_N - \mu_{22}}{\mu_{11}} + \frac{\mu_{22} g_N}{\Sigma_1' \mu_{11}}\right] = q\left[\frac{g_N - \mu_{22}}{\mu_{11}} + \frac{g_N f(1,2)}{\Sigma_1' \mu_{11}}\right].$$

This proves (33) for $s = 2 \leq N$. Now assume that (33) holds for $k = 2, \ldots, s$ and that $s + 1 \leq N$. We show that this implies (33) for $k = s + 1$. From (47)

$$\frac{\mu_{11}}{q}[h_N(s+1) - h_N(s)] = g_N - \mu_{22} + \frac{\mu_{22} + (s-1)\theta}{q}[h_N(s) - h_N(s-1)].$$

Using the induction hypothesis, we obtain

$$\begin{aligned} h_N(s+1) - h_N(s) &= \frac{q}{\mu_{11}}\left[g_N - \mu_{22} + [\mu_{22} + (s-1)\theta]\left[\sum_{i=0}^{s-2} \frac{g_N - \mu_{22}}{\mu_{11}^{i+1}} f(s-i, s) + \frac{g_N}{\Sigma_1' \mu_{11}^{s-1}} f(1, s)\right]\right] \\ &= q\left[\frac{g_N - \mu_{22}}{\mu_{11}} + \left[\sum_{i=0}^{s-2} \frac{g_N - \mu_{22}}{\mu_{11}^{i+2}} f(s-i, s+1) + \frac{g_N}{\Sigma_1' \mu_{11}^{s}} f(1, s+1)\right]\right]. \end{aligned}$$

Letting $i' = i + 1$, we have

$$\begin{aligned} h_N(s+1) - h_N(s) &= q\left[\frac{g_N - \mu_{22}}{\mu_{11}} + \sum_{i'=1}^{s-1} \frac{g_N - \mu_{22}}{\mu_{11}^{i'+1}} f(s - i' + 1, s+1) + \frac{g_N}{\Sigma_1' \mu_{11}^{s}} f(1, s+1)\right] \\ &= q\left[\sum_{i'=0}^{s-1} \frac{g_N - \mu_{22}}{\mu_{11}^{i'+1}} f(s - i' + 1, s+1) + \frac{g_N}{\Sigma_1' \mu_{11}^{s}} f(1, s+1)\right]. \end{aligned}$$

Thus (33) holds for $s + 1$. This completes the proof of (33) for $s = 1, \ldots, N$.

Finally from (48), we have for $s = N, \ldots, B + 2$,

$$h_N(s) - h_N(s-1) = \frac{q[\Sigma_2' - g_N]}{\Sigma_2' + (s-1)\theta}.$$

□

Proof of Lemma 6 For $s = 1$, we have using (17), (18), (49), and $h_N(0) = 0$ that

$$h_N(1) - h_N(0) = \frac{g_N q}{\Sigma_1'} = \frac{\beta(N) q}{\Delta(N) \Sigma_1'} = \frac{q}{\Delta(N)}\left[[\Sigma_2' + (N-1)\theta]\mu_{22} \sum_{k=2}^{N} \mu_{11}^{k-2} f(k, N) + \Sigma_2' \mu_{11}^{N-1}\right].$$

Hence (35) holds for $s = 1 \leq N$.

For $s = 2, \ldots, N$, we have using (17) that

$$\begin{aligned} g_N - \mu_{22} &= \frac{[\Sigma_2' + (N-1)\theta]\Sigma_1' \mu_{22} \alpha(N) + \Sigma_1' \Sigma_2' \mu_{11}^{N-1}}{[\Sigma_2' + (N-1)\theta][f(1,N) + \Sigma_1' \alpha(N)] + \Sigma_1' \mu_{11}^{N-1}} - \mu_{22} \\ &= \frac{-[\Sigma_2' + (N-1)\theta]\mu_{22} f(1,N) + \Sigma_1' \mu_{11}^{N-1}[\Sigma_2' - \mu_{22}]}{\Delta(N)} \end{aligned}$$

and

$$\frac{g_N}{\Sigma_1'} = \frac{[\Sigma_2' + (N-1)\theta]\mu_{22} \sum_{k=2}^{N} \mu_{11}^{k-2} f(k,N) + \Sigma_2' \mu_{11}^{N-1}}{\Delta(N)}.$$

Therefore the first term in (33) satisfies

$$\begin{aligned}
&\sum_{i=0}^{s-2} \frac{g_N - \mu_{22}}{\mu_{11}^{i+1}} f(s-i,s) \\
=&\frac{\Sigma_1' \mu_{11}^{N-1}[\Sigma_2' - \mu_{22}]}{\Delta(N)} \sum_{i=0}^{s-2} \frac{1}{\mu_{11}^{i+1}} f(s-i,s) - \frac{[\Sigma_2' + (N-1)\theta]\mu_{22} f(1,N)}{\Delta(N)} \sum_{i=0}^{s-2} \frac{1}{\mu_{11}^{i+1}} f(s-i,s).
\end{aligned}$$

The first term in the above satisfies

$$\frac{\Sigma_1' \mu_{11}^{N-1}[\Sigma_2' - \mu_{22}]}{\Delta(N)} \sum_{i=0}^{s-2} \frac{1}{\mu_{11}^{i+1}} f(s-i,s) = \frac{\Sigma_1' \mu_{11}^{N-s}[\Sigma_2' - \mu_{22}]}{\Delta(N)} \sum_{i=0}^{s-2} \mu_{11}^{s-2-i} f(s-i,s).$$

Letting $s - i = i'$ yields

$$\frac{\Sigma_1' \mu_{11}^{N-1}[\Sigma_2' - \mu_{22}]}{\Delta(N)} \sum_{i=0}^{s-2} \frac{1}{\mu_{11}^{i+1}} f(s-i,s) = \frac{\Sigma_1' \mu_{11}^{N-s}[\Sigma_2' - \mu_{22}]}{\Delta(N)} \alpha(s).$$

Therefore,

$$\sum_{i=0}^{s-2} \frac{g_N - \mu_{22}}{\mu_{11}^{i+1}} f(s-i,s) = \frac{\Sigma_1' \mu_{11}^{N-s}[\Sigma_2' - \mu_{22}]}{\Delta(N)} \alpha(s) - \frac{[\Sigma_2' + (N-1)\theta]\mu_{22} f(1,N)}{\Delta(N)} \sum_{i=0}^{s-2} \frac{1}{\mu_{11}^{i+1}} f(s-i,s). \tag{50}$$

From (17)- (18), the second term in (33) satisfies

$$\begin{aligned}
\frac{g_N}{\Sigma_1' \mu_{11}^{s-1}} f(1,s) =& \frac{[\Sigma_2' + (N-1)\theta]\mu_{22}}{\Delta(N)} \sum_{k=2}^{s} \mu_{11}^{k-1-s} f(k,N) f(1,s) \\
&+ \frac{[\Sigma_2' + (N-1)\theta]\mu_{22} \sum_{k=s+1}^{N} \mu_{11}^{k-2} f(k,N) + \Sigma_2' \mu_{11}^{N-1}}{\Delta(N) \mu_{11}^{s-1}} f(1,s).
\end{aligned}$$

For the first term in the above, letting $k' = s - k$ yields

$$\begin{aligned}
\frac{[\Sigma_2' + (N-1)\theta]\mu_{22}}{\Delta(N)} \sum_{k=2}^{s} \mu_{11}^{k-1-s} f(k,N) f(1,s) =& \frac{[\Sigma_2' + (N-1)\theta]\mu_{22}}{\Delta(N)} \sum_{k'=0}^{s-2} \frac{1}{\mu_{11}^{k'+1}} f(s-k',N) f(1,s) \\
=& \frac{[\Sigma_2' + (N-1)\theta]\mu_{22} f(1,N)}{\Delta(N)} \sum_{k'=0}^{s-2} \frac{1}{\mu_{11}^{k'+1}} f(s-k',s),
\end{aligned}$$

where in the last equality we have used $f(s-k',N)f(1,s) = f(1,N)f(s-k',s)$. Hence

$$\begin{aligned}
\frac{g_N}{\Sigma_1' \mu_{11}^{s-1}} f(1,s) =& \frac{[\Sigma_2' + (N-1)\theta]\mu_{22} f(1,N)}{\Delta(N)} \sum_{k'=0}^{s-2} \frac{1}{\mu_{11}^{k'+1}} f(s-k',s) \\
&+ \frac{[\Sigma_2' + (N-1)\theta]\mu_{22} \sum_{k=s+1}^{N} \mu_{11}^{k-s-1} f(k,N) + \Sigma_2' \mu_{11}^{N-s}}{\Delta(N)} f(1,s).
\end{aligned} \tag{51}$$

Putting (50) and (51) in (33) we get for $s = 2, \ldots, N$,

$h_N(s) - h_N(s-1)$

$$
\begin{aligned}
=&q\Bigg[\frac{\Sigma_1'\mu_{11}^{N-s}[\Sigma_2'-\mu_{22}]}{\Delta(N)}\alpha(s)-\frac{[\Sigma_2'+(N-1)\theta]\mu_{22}f(1,N)}{\Delta(N)}\sum_{i=0}^{s-2}\frac{1}{\mu_{11}^{i+1}}f(s-i,s)\\
&+\frac{[\Sigma_2'+(N-1)\theta]\mu_{22}f(1,N)}{\Delta(N)}\sum_{k'=0}^{s-2}\frac{1}{\mu_{11}^{k'+1}}f(s-k',s)\\
&+\frac{[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)+\Sigma_2'\mu_{11}^{N-s}}{\Delta(N)}f(1,s)\Bigg]\\
=&\frac{q}{\Delta(N)}\Bigg[\Sigma_1'\mu_{11}^{N-s}[\Sigma_2'-\mu_{22}]\alpha(s)+[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)+\Sigma_2'\mu_{11}^{N-s}f(1,s)\Bigg].
\end{aligned}
$$

We have shown that (35) holds for $s=1,\ldots,N$.

Finally for $s=N,\ldots,B+2$ we have from (34) that

$$
h_N(s)-h_N(s-1)=\frac{q[\Sigma_2'-g_N]}{\Sigma_2'+(s-1)\theta}.
$$

Now, from (17)-(19), we have

$$
\begin{aligned}
\Sigma_2'-g_N=\Sigma_2'-\frac{\beta(N)}{\Delta(N)}=&\Sigma_2'-\frac{[\Sigma_2'+(N-1)\theta]\Sigma_1'\mu_{22}\alpha(N)+\Sigma_1'\Sigma_2'\mu_{11}^{N-1}}{[\Sigma_2'+(N-1)\theta][f(1,N)+\Sigma_1'\alpha(N)]+\Sigma_1'\mu_{11}^{N-1}}\\
=&\frac{[\Sigma_2'+(N-1)\theta]\Big(\Sigma_2'f(1,N)+\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(N)\Big)}{\Delta(N)}.
\end{aligned}
$$

Hence,

$$
h_N(s)-h_N(s-1)=\frac{q[\Sigma_2'+(N-1)\theta]\Big(\Sigma_2'f(1,N)+\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(N)\Big)}{\Delta(N)[\Sigma_2'+(s-1)\theta]}.
$$

□

Proof of Lemma 7 We have to show that for $s=1,\ldots,N-1$, $\sum_{k=2}^{N}\mu_{11}^{k-2}f(k,N)\geq\mu_{11}^{N-s}\sum_{k=2}^{s}\mu_{11}^{k-2}f(k,s)$. That is

$$
\sum_{k=2}^{N}\mu_{11}^{k}f(k,N)\geq\sum_{k=2}^{s}\mu_{11}^{N-s+k}f(k,s)\iff\sum_{k=N-s+2}^{N}\mu_{11}^{k}f(k,N)-\sum_{k=2}^{s}\mu_{11}^{N-s+k}f(k,s)+\sum_{k=2}^{N-s+1}\mu_{11}^{k}f(k,N)\geq 0.
$$

Setting $i=N-k$ in the first term and $i=s-k$ in the second term yields

$$
\sum_{k=2}^{N}\mu_{11}^{k}f(k,N)\geq\sum_{k=2}^{s}\mu_{11}^{N-s+k}f(k,s)\iff\sum_{i=0}^{s-2}\mu_{11}^{N-i}\Big[f(N-i,N)-f(s-i,s)\Big]+\sum_{k=2}^{N-s+1}\mu_{11}^{k}f(k,N)\geq 0.
$$

Therefore it is sufficient to show that $f(N-i,N)\geq f(s-i,s)$, which is true since for $s=1,\ldots,N-1$, we have

$$
f(N-i,N)=\prod_{j=N-i}^{N-1}[\mu_{22}+(j-1)\theta]\geq\prod_{j=s-i}^{s-1}[\mu_{22}+(j-1)\theta]=f(s-i,s).
$$

□

A.2 Details of the Proof of Proposition 4

Eq. (25) states that for $n \geq 2$,

$$\begin{aligned}\tau'(n+1) =&[\Sigma_2' + n\theta][\Sigma_2' + (n-1)\theta]\mu_{22} f(1,n) - [\Sigma_2' + n\theta][\Sigma_1'\alpha(n+1)(\Sigma_2' - \mu_{22}) + \Sigma_2' f(1,n+1)] \\ &+ [\Sigma_2' + (n-1)\theta]\mu_{11}[\Sigma_2' f(1,n) + \Sigma_1'\alpha(n)(\Sigma_2' - \mu_{22})].\end{aligned}$$

The second term satisfies

$$\begin{aligned}&[\Sigma_2' + n\theta][\Sigma_1'\alpha(n+1)(\Sigma_2' - \mu_{22}) + \Sigma_2' f(1,n+1)] \\ =&[\Sigma_2' + n\theta][\Sigma_1'(\Sigma_2' - \mu_{22})(\mu_{22} + (n-1)\theta)\alpha(n) + \Sigma_1'(\Sigma_2' - \mu_{22})\mu_{11}^{n-1} + \Sigma_2' f(1,n)[\mu_{22} + (n-1)\theta]] \\ =&[\mu_{22} + (n-1)\theta][\Sigma_2' + n\theta][\Sigma_1'(\Sigma_2' - \mu_{22})\alpha(n) + \Sigma_2' f(1,n)] + \Sigma_1'\mu_{11}^{n-1}(\Sigma_2' - \mu_{22})[\Sigma_2' + n\theta] \\ =&[\mu_{22} + (n-1)\theta][\Sigma_2' + (n-1)\theta][\Sigma_1'(\Sigma_2' - \mu_{22})\alpha(n) + \Sigma_2' f(1,n)] + \Sigma_1'\mu_{11}^{n-1}(\Sigma_2' - \mu_{22})[\Sigma_2' + n\theta] \\ &+ \theta[\mu_{22} + (n-1)\theta][\Sigma_1'(\Sigma_2' - \mu_{22})\alpha(n) + \Sigma_2' f(1,n)]. \qquad (52)\end{aligned}$$

The third term satisfies

$$\begin{aligned}&[\Sigma_2' + (n-1)\theta]\mu_{11}[\Sigma_2' f(1,n) + \Sigma_1'\alpha(n)(\Sigma_2' - \mu_{22})] \\ =&[\Sigma_2' + (n-2)\theta]\mu_{11}[\Sigma_2' f(1,n) + \Sigma_1'\alpha(n)(\Sigma_2' - \mu_{22})] + \theta\mu_{11}[\Sigma_2' f(1,n) + \Sigma_1'\alpha(n)(\Sigma_2' - \mu_{22})] \\ =&[\Sigma_2' + (n-2)\theta]\mu_{11}[\Sigma_2' f(1,n-1)[\mu_{22} + (n-2)\theta] + \Sigma_1'(\Sigma_2' - \mu_{22})[\mu_{22} + (n-2)\theta]\alpha(n-1) \\ &+ \Sigma_1'(\Sigma_2' - \mu_{22})\mu_{11}^{n-2}] + \theta\mu_{11}[\Sigma_2' f(1,n) + \Sigma_1'\alpha(n)(\Sigma_2' - \mu_{22})] \\ =&[\Sigma_2' + (n-2)\theta]\mu_{11}[\Sigma_2' f(1,n-1)[\mu_{22} + (n-2)\theta] + \Sigma_1'(\Sigma_2' - \mu_{22})[\mu_{22} + (n-2)\theta]\alpha(n-1)] \\ &+ \Sigma_1'(\Sigma_2' - \mu_{22})\mu_{11}^{n-2}[\Sigma_2' + (n-2)\theta]\mu_{11} + \theta\mu_{11}[\Sigma_2' f(1,n) + \Sigma_1'\alpha(n)(\Sigma_2' - \mu_{22})] \\ =&[\Sigma_2' + (n-2)\theta]\mu_{11}[\mu_{22} + (n-2)\theta][\Sigma_2' f(1,n-1) + \Sigma_1'(\Sigma_2' - \mu_{22})\alpha(n-1)] \\ &+ \Sigma_1'(\Sigma_2' - \mu_{22})\mu_{11}^{n-1}[\Sigma_2' + (n-2)\theta] + \theta\mu_{11}[\Sigma_2' f(1,n) + \Sigma_1'\alpha(n)(\Sigma_2' - \mu_{22})] \\ =&[\Sigma_2' + (n-2)\theta]\mu_{11}[\mu_{22} + (n-1)\theta][\Sigma_2' f(1,n-1) + \Sigma_1'(\Sigma_2' - \mu_{22})\alpha(n-1)] \\ &+ \Sigma_1'(\Sigma_2' - \mu_{22})\mu_{11}^{n-1}[\Sigma_2' + (n-2)\theta] + \theta\mu_{11}[\Sigma_2' f(1,n) + \Sigma_1'\alpha(n)(\Sigma_2' - \mu_{22})] \\ &- \theta[\Sigma_2' + (n-2)\theta]\mu_{11}[\Sigma_2' f(1,n-1) + \Sigma_1'(\Sigma_2' - \mu_{22})\alpha(n-1)]. \qquad (53)\end{aligned}$$

Finally, the first term satisfies

$$\begin{aligned}&[\Sigma_2' + n\theta][\Sigma_2' + (n-1)\theta]\mu_{22} f(1,n) \\ =&[\Sigma_2' + (n-2)\theta][\Sigma_2' + (n-1)\theta]\mu_{22} f(1,n) + 2\theta[\Sigma_2' + (n-1)\theta]\mu_{22} f(1,n) \\ =&[\Sigma_2' + (n-2)\theta][\Sigma_2' + (n-1)\theta]\mu_{22} f(1,n-1)[\mu_{22} + (n-2)\theta] + 2\theta[\Sigma_2' + (n-1)\theta]\mu_{22} f(1,n) \\ =&[\Sigma_2' + (n-2)\theta][\Sigma_2' + (n-1)\theta]\mu_{22} f(1,n-1)[\mu_{22} + (n-1)\theta] + 2\theta[\Sigma_2' + (n-1)\theta]\mu_{22} f(1,n) \\ &- \theta[\Sigma_2' + (n-2)\theta][\Sigma_2' + (n-1)\theta]\mu_{22} f(1,n-1). \qquad (54)\end{aligned}$$

Combining the first terms in (52)-(54) yields $[\mu_{22} + (n-1)\theta]\tau'(n)$. Therefore, writing

$$\tau'(n+1) = [\mu_{22} + (n-1)\theta]\tau'(n) - \epsilon'(n)$$

leads to

$$
\begin{aligned}
&\epsilon'(n)\\
=&-2\theta[\Sigma_2'+(n-1)\theta]\mu_{22}f(1,n)+\theta[\Sigma_2'+(n-2)\theta][\Sigma_2'+(n-1)\theta]\mu_{22}f(1,n-1)\\
&+\Sigma_1'\mu_{11}^{n-1}(\Sigma_2'-\mu_{22})[\Sigma_2'+n\theta]+\theta[\mu_{22}+(n-1)\theta][\Sigma_1'(\Sigma_2'-\mu_{22})\alpha(n)+\Sigma_2'f(1,n)]\\
&-\Sigma_1'(\Sigma_2'-\mu_{22})\mu_{11}^{n-1}[\Sigma_2'+(n-2)\theta]-\theta\mu_{11}[\Sigma_2'f(1,n)+\Sigma_1'\alpha(n)(\Sigma_2'-\mu_{22})]\\
&+\theta[\Sigma_2'+(n-2)\theta]\mu_{11}[\Sigma_2'f(1,n-1)+\Sigma_1'(\Sigma_2'-\mu_{22})\alpha(n-1)]\\
=&-2\theta[\Sigma_2'+(n-1)\theta]\mu_{22}f(1,n)+\theta[\Sigma_2'+(n-2)\theta][\Sigma_2'+(n-1)\theta]\mu_{22}f(1,n-1)\\
&+2\theta\Sigma_1'\mu_{11}^{n-1}(\Sigma_2'-\mu_{22})+\theta[\mu_{22}+(n-1)\theta][\Sigma_1'(\Sigma_2'-\mu_{22})\alpha(n)+\Sigma_2'f(1,n)]\\
&-\theta\mu_{11}[\Sigma_2'f(1,n)+\Sigma_1'\alpha(n)(\Sigma_2'-\mu_{22})]\\
&+\theta[\Sigma_2'+(n-2)\theta]\mu_{11}[\Sigma_2'f(1,n-1)+\Sigma_1'(\Sigma_2'-\mu_{22})\alpha(n-1)]\\
=&-2\theta[\Sigma_2'+(n-1)\theta]\mu_{22}f(1,n)+\theta[\Sigma_2'+(n-2)\theta][\Sigma_2'+(n-1)\theta]\mu_{22}f(1,n-1)\\
&+2\theta\Sigma_1'\mu_{11}^{n-1}(\Sigma_2'-\mu_{22})+\theta[\mu_{22}+(n-1)\theta][\Sigma_1'(\Sigma_2'-\mu_{22})\alpha(n)+\Sigma_2'f(1,n)]\\
&-\theta\mu_{11}[\mu_{22}+(n-2)\theta][\Sigma_2'f(1,n-1)+\Sigma_1'\alpha(n-1)(\Sigma_2'-\mu_{22})]-\theta\mu_{11}^{n-1}\Sigma_1'(\Sigma_2'-\mu_{22})\\
&+\theta[\Sigma_2'+(n-2)\theta]\mu_{11}[\Sigma_2'f(1,n-1)+\Sigma_1'(\Sigma_2'-\mu_{22})\alpha(n-1)]\\
=&-2\theta[\Sigma_2'+(n-1)\theta]\mu_{22}f(1,n-1)[\mu_{22}+(n-2)\theta]+\theta[\Sigma_2'+(n-2)\theta][\Sigma_2'+(n-1)\theta]\mu_{22}f(1,n-1)\\
&+\theta\Sigma_1'\mu_{11}^{n-1}(\Sigma_2'-\mu_{22})+\theta[\mu_{22}+(n-1)\theta][\Sigma_1'(\Sigma_2'-\mu_{22})\alpha(n)+\Sigma_2'f(1,n)]\\
&+\theta(\Sigma_2'-\mu_{22})\mu_{11}[\Sigma_2'f(1,n-1)+\Sigma_1'(\Sigma_2'-\mu_{22})\alpha(n-1)]\\
=&-\theta[\Sigma_2'+(n-1)\theta]\mu_{22}f(1,n-1)[\mu_{22}+(n-2)\theta]+\theta(\Sigma_2'-\mu_{22})[\Sigma_2'+(n-1)\theta]\mu_{22}f(1,n-1)\\
&+\theta\Sigma_1'\mu_{11}^{n-1}(\Sigma_2'-\mu_{22})+\theta[\mu_{22}+(n-1)\theta][\Sigma_1'(\Sigma_2'-\mu_{22})\alpha(n)+\Sigma_2'f(1,n)]\\
&+\theta(\Sigma_2'-\mu_{22})\mu_{11}[\Sigma_2'f(1,n-1)+\Sigma_1'(\Sigma_2'-\mu_{22})\alpha(n-1)]\\
=&-\theta[\Sigma_2'+(n-1)\theta]\mu_{22}f(1,n-1)[\mu_{22}+(n-2)\theta]+\theta(\Sigma_2'-\mu_{22})[\Sigma_2'+(n-1)\theta]\mu_{22}f(1,n-1)\\
&+\theta\Sigma_1'\mu_{11}^{n-1}(\Sigma_2'-\mu_{22})+\theta[\mu_{22}+(n-1)\theta][\mu_{22}+(n-2)\theta][\Sigma_1'(\Sigma_2'-\mu_{22})\alpha(n-1)]\\
&+\theta[\mu_{22}+(n-1)\theta][\mu_{22}+(n-2)\theta]\Sigma_2'f(1,n-1)+\Sigma_1'(\Sigma_2'-\mu_{22})\mu_{11}^{n-2}\theta[\mu_{22}+(n-1)\theta]\\
&+\theta(\Sigma_2'-\mu_{22})\mu_{11}[\Sigma_2'f(1,n-1)+\Sigma_1'(\Sigma_2'-\mu_{22})\alpha(n-1)]\\
=&-\theta[\Sigma_2'+(n-1)\theta]\mu_{22}f(1,n-1)[\mu_{22}+(n-2)\theta]+\theta(\Sigma_2'-\mu_{22})[\Sigma_2'+(n-1)\theta]\mu_{22}f(1,n-1)\\
&+\theta\Sigma_1'\mu_{11}^{n-1}(\Sigma_2'-\mu_{22})+\theta[\mu_{22}+(n-1)\theta][\mu_{22}+(n-2)\theta]\Sigma_1'(\Sigma_2'-\mu_{22})\alpha(n-1)\\
&+\theta[\mu_{22}+(n-1)\theta][\mu_{22}+(n-2)\theta](\Sigma_2'-\mu_{22})f(1,n-1)\\
&+\theta[\mu_{22}+(n-1)\theta][\mu_{22}+(n-2)\theta]\mu_{22}f(1,n-1)+\Sigma_1'(\Sigma_2'-\mu_{22})\mu_{11}^{n-2}\theta[\mu_{22}+(n-1)\theta]\\
&+\theta(\Sigma_2'-\mu_{22})\mu_{11}[\Sigma_2'f(1,n-1)+\Sigma_1'(\Sigma_2'-\mu_{22})\alpha(n-1)]\\
=&-\theta(\Sigma_2'-\mu_{22})\mu_{22}f(1,n-1)[\mu_{22}+(n-2)\theta]+\theta(\Sigma_2'-\mu_{22})[\Sigma_2'+(n-1)\theta]\mu_{22}f(1,n-1)\\
&+\theta\Sigma_1'\mu_{11}^{n-1}(\Sigma_2'-\mu_{22})+\theta[\mu_{22}+(n-1)\theta][\mu_{22}+(n-2)\theta]\Sigma_1'(\Sigma_2'-\mu_{22})\alpha(n-1)\\
&+\theta[\mu_{22}+(n-1)\theta][\mu_{22}+(n-2)\theta](\Sigma_2'-\mu_{22})f(1,n-1)\\
&+\Sigma_1'(\Sigma_2'-\mu_{22})\mu_{11}^{n-2}\theta[\mu_{22}+(n-1)\theta]\\
&+\theta(\Sigma_2'-\mu_{22})\mu_{11}[\Sigma_2'f(1,n-1)+\Sigma_1'(\Sigma_2'-\mu_{22})\alpha(n-1)]\\
=&\theta(\Sigma_2'-\mu_{22})[\Sigma_2'-\mu_{22}+\theta]\mu_{22}f(1,n-1)
\end{aligned}
$$

$$+\theta\Sigma_1'\mu_{11}^{n-1}(\Sigma_2'-\mu_{22})+\theta[\mu_{22}+(n-1)\theta][\mu_{22}+(n-2)\theta]\Sigma_1'(\Sigma_2'-\mu_{22})\alpha(n-1)$$
$$+\theta[\mu_{22}+(n-1)\theta][\mu_{22}+(n-2)\theta](\Sigma_2'-\mu_{22})f(1,n-1)$$
$$+\Sigma_1'(\Sigma_2'-\mu_{22})\mu_{11}^{n-2}\theta[\mu_{22}+(n-1)\theta]$$
$$+\theta(\Sigma_2'-\mu_{22})\mu_{11}[\Sigma_2'f(1,n-1)+\Sigma_1'(\Sigma_2'-\mu_{22})\alpha(n-1)].$$

A.3 Proof of Theorem 8

Proof As mentioned at the end of Section 3, we employ policy iteration for unichain models, as detailed in [24]. By initializing the decision rule as $\nu_0 = d_N$, we demonstrate that the subsequent decision rule generated by the algorithm, remains equal to d_N. Specifically, we prove that

$$\nu_1(s) \in \arg\max_{a\in A}\left\{r(s,a)+\sum_{j\in S}p(j|s,a)h_N(j)\right\}$$

yields $d_N(s)$ for all $s \in S$. To achieve this, we establish that the difference function

$$\Gamma(s,a) = r(s,d_N(s))+\sum_{j\in S}p(j|s,d_N(s))h_N(j) - r(s,a) - \sum_{j\in S}p(j|s,a)h_N(j)$$

is non-negative for all $s \in S$ and $a \in A$. The conclusion then follows directly from Theorem 8.6.6 in [24]. We use policy iteration for unichain models (see [24]). We set the initial decision rule ν_0 equal to d_N and then show that the next decision rule computed by the policy iteration algorithm is equal to d_N. That is we show that

$$\nu_1(s) \in \arg\max_{a\in A}\left\{r(s,a)+\sum_{j\in S}p(j|s,a)h_N(j)\right\}$$

is equal to $d_N(s)$, $\forall s \in S$. We show this by showing the difference

$$\Gamma(s,a) = r(s,d_N(s))+\sum_{j\in S}p(j|s,d_N(s))h_N(j) - r(s,a) - \sum_{j\in S}p(j|s,a)h_N(j)$$

is non negative for all $s \in S$ and $a \in A$. The result then follows from Theorem 8.6.6 in [24] .

We consider three separate cases; Case 1: $s = 0$; Case 2: $s = 1,\ldots,N-1$; and Case 3: $s = N,\ldots,B+2$.

Case 1: $s = 0$.

In this case we have $d_N(s) = a_{11}$.

Case 1, Part (i) : $a = a_{12}$.

From Lemma 5, we have

$$\begin{aligned}\Gamma(0,a_{12}) =& \frac{\Sigma_1'}{q}h_N(1)+\left(1-\frac{\Sigma_1'}{q}\right)h_N(0)-\frac{\mu_{11}}{q}h_N(1)-\left(1-\frac{\mu_{11}}{q}\right)h_N(0)\\ =& \frac{h_N(1)}{q}[\Sigma_1'-\mu_{11}] = \frac{g_N}{\Sigma_1'}[\Sigma_1'-\mu_{11}] \geq 0.\end{aligned}$$

Case 1, Part (ii) : $a = a_{21}$.

From Lemma 5, we have

$$\begin{aligned}\Gamma(0,a_{21}) =& \frac{\Sigma_1'}{q}h_N(1)+\left(1-\frac{\Sigma_1'}{q}\right)h_N(0)-\frac{\mu_{21}}{q}h_N(1)-\left(1-\frac{\mu_{21}}{q}\right)h_N(0)\\ =& \frac{h_N(1)}{q}[\Sigma_1'-\mu_{21}] = \frac{g_N}{\Sigma_1'}[\Sigma_1'-\mu_{21}] \geq 0.\end{aligned}$$

Case 1, Part (iii) : $a = a_{22}$.

From Lemma 5, we have

$$\Gamma(0, a_{22}) = \frac{\Sigma_1'}{q} h_N(1) + \left(1 - \frac{\Sigma_1'}{q}\right) h_N(0) = \frac{\Sigma_1'}{q} \frac{g_N q}{\Sigma_1'} = g_N \geq 0.$$

Thus we have proved $\Gamma(s, a) \geq 0$ for all policies a in Case 1.

Case 2: $s = 1, \ldots, N-1$.

In this case we have $d_N(s) = a_{12}$.

Case 2, Part (i) : $a = a_{22}$.

$$\begin{aligned}
&\Gamma(s, a_{22}) \\
=&\mu_{22} + \frac{\mu_{22} + (s-1)\theta}{q} h_N(s-1) + \left(1 - \frac{\mu_{11} + \mu_{22} + (s-1)\theta}{q}\right) h_N(s) + \frac{\mu_{11}}{q} h_N(s+1) \\
&- \Sigma_2' - \frac{\Sigma_2' + (s-1)\theta}{q} h_N(s-1) - \left(1 - \frac{\Sigma_2' + (s-1)\theta}{q}\right) h_N(s) \\
=&\mu_{22} - \Sigma_2' + \frac{\mu_{22}}{q}[h_N(s-1) - h_N(s)] - \frac{\mu_{11}}{q}[h_N(s) - h_N(s+1)] - \frac{\Sigma_2'}{q}[h_N(s-1) - h_N(s)] \\
=&\mu_{22} - \Sigma_2' + \frac{\Sigma_2' - \mu_{22}}{q}[h_N(s) - h_N(s-1)] + \frac{\mu_{11}}{q}[h_N(s+1) - h_N(s)]. \qquad (55)
\end{aligned}$$

We have, from (47) in proof of Lemma 5, that

$$\frac{\mu_{11}}{q}[h_N(s+1) - h_N(s)] = g_N - \mu_{22} + \frac{\mu_{22} + (s-1)\theta}{q}[h_N(s) - h_N(s-1)]. \qquad (56)$$

Therefore

$$\begin{aligned}
&\Gamma(s, a_{22}) \\
=&\mu_{22} - \Sigma_2' + \frac{\Sigma_2' - \mu_{22}}{q}[h_N(s) - h_N(s-1)] + g_N - \mu_{22} + \frac{\mu_{22} + (s-1)\theta}{q}[h_N(s) - h_N(s-1)] \\
=&g_N - \Sigma_2' + \frac{\Sigma_2' + (s-1)\theta}{q}[h_N(s) - h_N(s-1)].
\end{aligned}$$

From Lemma 5 we have

$$\begin{aligned}
&\Gamma(s, a_{22}) \\
=&g_N - \Sigma_2' + [\Sigma_2' + (s-1)\theta]\left[\sum_{i=0}^{s-2} \frac{g_N - \mu_{22}}{\mu_{11}^{i+1}} f(s-i, s) + \frac{g_N}{\Sigma_1' \mu_{11}^{s-1}} f(1, s)\right] \\
=&g_N \left(1 + \frac{\Sigma_2' + (s-1)\theta}{\Sigma_1' \mu_{11}^{s-1}} f(1, s) + [\Sigma_2' + (s-1)\theta] \sum_{i=0}^{s-2} \frac{1}{\mu_{11}^{i+1}} f(s-i, s)\right) \\
&- \left(\Sigma_2' + \mu_{22}[\Sigma_2' + (s-1)\theta] \sum_{i=0}^{s-2} \frac{1}{\mu_{11}^{i+1}} f(s-i, s)\right) \\
=&g_N \left(\frac{\Sigma_1' \mu_{11}^{s-1} + [\Sigma_2' + (s-1)\theta] f(1, s) + \Sigma_1' [\Sigma_2' + (s-1)\theta] \sum_{i=0}^{s-2} \mu_{11}^{s-i-2} f(s-i, s)}{\Sigma_1' \mu_{11}^{s-1}}\right) \\
&- \left(\frac{\Sigma_1' \Sigma_2' \mu_{11}^{s-1} + \mu_{22} \Sigma_1' [\Sigma_2' + (s-1)\theta] \sum_{i=0}^{s-2} \mu_{11}^{s-i-2} f(s-i, s)}{\Sigma_1' \mu_{11}^{s-1}}\right).
\end{aligned}$$

Let $s - i = i'$ and $j' = j + 1$. We have

$$\Gamma(s, a_{22})$$

$$
\begin{aligned}
=&g_N\left(\frac{\Sigma_1'\mu_{11}^{s-1}+[\Sigma_2'+(s-1)\theta]f(1,s)+\Sigma_1'[\Sigma_2'+(s-1)\theta]\sum_{i'=2}^{s}\mu_{11}^{i'-2}f(i',s)}{\Sigma_1'\mu_{11}^{s-1}}\right)\\
&-\left(\frac{\Sigma_1'\Sigma_2'\mu_{11}^{s-1}+\mu_{22}\Sigma_1'[\Sigma_2'+(s-1)\theta]\sum_{i'=2}^{s}\mu_{11}^{i'-2}f(i',s)}{\Sigma_1'\mu_{11}^{s-1}}\right)\\
=&g_N\left(\frac{\Sigma_1'\mu_{11}^{s-1}+[\Sigma_2'+(s-1)\theta]f(1,s)+\Sigma_1'[\Sigma_2'+(s-1)\theta]\alpha(s)}{\Sigma_1'\mu_{11}^{s-1}}\right)\\
&-\left(\frac{\Sigma_1'\Sigma_2'\mu_{11}^{s-1}+\mu_{22}\Sigma_1'[\Sigma_2'+(s-1)\theta]\alpha(s)}{\Sigma_1'\mu_{11}^{s-1}}\right)\\
\overset{(*)}{\geq}&\frac{\beta(s+1)}{\Delta(s+1)}\left(\frac{\Sigma_1'\mu_{11}^{s-1}+[\Sigma_2'+(s-1)\theta][f(1,s)+\Sigma_1'\alpha(s)]}{\Sigma_1'\mu_{11}^{s-1}}\right)\\
&-\Delta(s+1)\left(\frac{\Sigma_1'\Sigma_2'\mu_{11}^{s-1}+\mu_{22}\Sigma_1'[\Sigma_2'+(s-1)\theta]\alpha(s)}{\Delta(s+1)\Sigma_1'\mu_{11}^{s-1}}\right),
\end{aligned}
$$

where $(*)$ follows from $g_N \geq g_{s+1}$ (by the definition of N) and (17). From (18)-(19) we have

$$
\Gamma(s,a_{22}) \geq \frac{\beta(s+1)\Delta(s)-\Delta(s+1)\beta(s)}{\Delta(s+1)\Sigma_1'\mu_{11}^{s-1}} = \frac{\Sigma_1'\mu_{11}^{s-1}\tau'(s+1)}{\Delta(s+1)\Sigma_1'\mu_{11}^{s-1}} \geq 0
$$

from (24) and the definition of N.

Case 2, Part (ii) : $a = a_{21}$.

We have

$$
\begin{aligned}
\Gamma(s,a_{21}) =&\mu_{22}+\frac{\mu_{22}+(s-1)\theta}{q}h_N(s-1)+\left(1-\frac{\mu_{11}+\mu_{22}+(s-1)\theta}{q}\right)h_N(s)\\
&+\frac{\mu_{11}}{q}h_N(s+1)-\mu_{12}-\frac{\mu_{12}+(s-1)\theta}{q}h_N(s-1)-\frac{\mu_{21}}{q}h_N(s+1)\\
&-\left(1-\frac{\mu_{21}+\mu_{12}+(s-1)\theta}{q}\right)h_N(s)\\
=&\mu_{22}-\mu_{12}+\frac{\mu_{22}}{q}[h_N(s-1)-h_N(s)]+\frac{\mu_{11}}{q}[h_N(s+1)-h_N(s)]\\
&+\frac{\mu_{12}}{q}[h_N(s)-h_N(s-1)]+\frac{\mu_{21}}{q}[h_N(s)-h_N(s+1)]\\
=&\mu_{22}-\mu_{12}+\frac{\mu_{12}-\mu_{22}}{q}[h_N(s)-h_N(s-1))]+\frac{\mu_{11}-\mu_{21}}{q}[h_N(s+1)-h_N(s)]\\
=&\mu_{22}-\Sigma_2'+\frac{\Sigma_2'-\mu_{22}}{q}[h_N(s)-h_N(s-1))]+\frac{\mu_{11}}{q}[h_N(s+1)-h_N(s)]\\
&+\Sigma_2'-\mu_{12}+\frac{\Sigma_2'-\mu_{12}}{q}[h_N(s-1)-h_N(s)]+\frac{\mu_{21}}{q}[h_N(s)-h_N(s+1)]\\
=&\Gamma(s,a_{22})+\Gamma'(s,a_{21})
\end{aligned}
$$

from (55), where

$$
\Gamma'(s,a_{21}) = \Sigma_2'-\mu_{12}-\frac{\Sigma_2'-\mu_{12}}{q}[h_N(s)-h_N(s-1)]-\frac{\mu_{21}}{q}[h_N(s+1)-h_N(s)].
$$

Since we have shown that $\Gamma(s,a_{22})$ is non-negative, it is sufficient to show $\Gamma'(s,a_{21})$ is non-negative.

Using (56), we have

$\Gamma'(s,a_{21})$

$$
\begin{aligned}
=&\Sigma_2' - \mu_{12} - \frac{\Sigma_2' - \mu_{12}}{q}[h_N(s) - h_N(s-1)] - \frac{g_N\mu_{21} - \mu_{22}\mu_{21}}{\mu_{11}} \\
&- \frac{\mu_{21}\mu_{22} + \mu_{21}(s-1)\theta}{q\mu_{11}} \times [h_N(s) - h_N(s-1)] \\
\overset{(*)}{=}& \frac{\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21} - \mu_{21}g_N}{\mu_{11}} + \frac{-\mu_{11}(\Sigma_2' - \mu_{12}) - \mu_{21}\mu_{22} - \mu_{21}(s-1)\theta}{\Delta(N)\mu_{11}} \\
&\times \left[\Sigma_1'\mu_{11}^{N-s}[\Sigma_2' - \mu_{22}]\alpha(s) + [\Sigma_2' + (N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s) + \Sigma_2'\mu_{11}^{N-s}f(1,s)\right]
\end{aligned}
$$

where $(*)$ follows from Lemma 6. This yields

$$
\begin{aligned}
&\Gamma'(s,a_{21})\Delta(N)\mu_{11} \\
=&[\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21}]\Delta(N) - \mu_{21}\beta(N) + [-\mu_{11}(\Sigma_2' - \mu_{12}) - \mu_{21}\mu_{22} - \mu_{21}(s-1)\theta] \\
&\times \left[\Sigma_1'\mu_{11}^{N-s}[\Sigma_2' - \mu_{22}]\alpha(s) + [\Sigma_2' + (N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s) + \Sigma_2'\mu_{11}^{N-s}f(1,s)\right] \\
=&[\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21}][\Sigma_2' + (N-1)\theta][f(1,N) + \Sigma_1'\alpha(N)] + [\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21}]\Sigma_1'\mu_{11}^{N-1} \\
&- \mu_{21}[\Sigma_2' + (N-1)\theta]\Sigma_1'\mu_{22}\alpha(N) - \mu_{21}\Sigma_1'\Sigma_2'\mu_{11}^{N-1} \\
&+ [-\mu_{11}(\Sigma_2' - \mu_{12}) - \mu_{21}\mu_{22} - \mu_{21}(s-1)\theta]\Sigma_1'\mu_{11}^{N-s}[\Sigma_2' - \mu_{22}]\alpha(s) \\
&+ [-\mu_{11}(\Sigma_2' - \mu_{12}) - \mu_{21}\mu_{22} - \mu_{21}(s-1)\theta][\Sigma_2' + (N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s) \\
&+ [-\mu_{11}(\Sigma_2' - \mu_{12}) - \mu_{21}\mu_{22} - \mu_{21}(s-1)\theta]\Sigma_2'\mu_{11}^{N-s}f(1,s) \\
=&[\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21}]\Sigma_1'\mu_{11}^{N-1} - \mu_{21}\Sigma_1'\Sigma_2'\mu_{11}^{N-1} \\
&+ [-\mu_{11}(\Sigma_2' - \mu_{12}) - \mu_{21}\mu_{22} - \mu_{21}(s-1)\theta]\Sigma_1'\mu_{11}^{N-s}[\Sigma_2' - \mu_{22}]\alpha(s) \\
&+ [\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21}][\Sigma_2' + (N-1)\theta][f(1,N) + \Sigma_1'\alpha(N)] - \mu_{21}[\Sigma_2' + (N-1)\theta]\Sigma_1'\mu_{22}\alpha(N) \\
&+ [-\mu_{11}(\Sigma_2' - \mu_{12}) - \mu_{21}\mu_{22} - \mu_{21}(s-1)\theta][\Sigma_2' + (N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s) \\
&+ [-\mu_{11}(\Sigma_2' - \mu_{12}) - \mu_{21}\mu_{22} - \mu_{21}(s-1)\theta]\Sigma_2'\mu_{11}^{N-s}f(1,s) \\
=&[\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21} - \mu_{21}\Sigma_2']\Sigma_1'\mu_{11}^{N-1} - [\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{21}\mu_{22}]\Sigma_1'\mu_{11}^{N-s}[\Sigma_2' - \mu_{22}]\alpha(s) \\
&+ [\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21} - \mu_{21}\Sigma_2'](s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s) - (\mu_{11}(\Sigma_2' - \mu_{12}))(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s) \\
&+ [\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21}][\Sigma_2' + (N-1)\theta][f(1,N) + \Sigma_1'\alpha(N)] - \mu_{21}[\Sigma_2' + (N-1)\theta]\Sigma_1'\mu_{22}\alpha(N) \\
&+ [-\mu_{11}(\Sigma_2' - \mu_{12}) - \mu_{21}\mu_{22} - \mu_{21}(s-1)\theta][\Sigma_2' + (N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s) \\
&+ [-\mu_{11}(\Sigma_2' - \mu_{12}) - \mu_{21}\mu_{22} - \mu_{21}(s-1)\theta]\Sigma_2'\mu_{11}^{N-s}f(1,s) \\
=&[\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21} - \mu_{21}\Sigma_2']\left(\Sigma_1'\mu_{11}^{N-1} + (s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\right) \\
&- [\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{21}\mu_{22}]\Sigma_1'\mu_{11}^{N-s}[\Sigma_2' - \mu_{22}]\alpha(s) \\
&- (\mu_{11}(\Sigma_2' - \mu_{12}))(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s) \\
&+ [\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21}][\Sigma_2' + (N-1)\theta][f(1,N) + \Sigma_1'\alpha(N)] - \mu_{21}[\Sigma_2' + (N-1)\theta]\Sigma_1'\mu_{22}\alpha(N)
\end{aligned}
$$

$$
\begin{aligned}
&+[-\mu_{11}(\Sigma_2'-\mu_{12})-\mu_{21}\mu_{22}-\mu_{21}(s-1)\theta][\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&+[-\mu_{11}(\Sigma_2'-\mu_{12})-\mu_{21}\mu_{22}-\mu_{21}(s-1)\theta]\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
=&[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2']\Big(\Sigma_1'\mu_{11}^{N-1}+(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\Big)\\
&-[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{21}\mu_{22}]\Sigma_1'\mu_{11}^{N-s}[\Sigma_2'-\mu_{22}]\alpha(s)\\
&-(\mu_{11}(\Sigma_2'-\mu_{12}))(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\\
&+(-\mu_{11}(\Sigma_2'-\mu_{12})-\mu_{21}\mu_{22}-\Sigma_1'(s-1)\theta)[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&+(\Sigma_1'-\mu_{21})(s-1)\theta[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\Sigma_2'+(N-1)\theta][f(1,N)+\Sigma_1'\alpha(N)]-\mu_{21}[\Sigma_2'+(N-1)\theta]\Sigma_1'\mu_{22}\alpha(N)\\
&+[-\mu_{11}(\Sigma_2'-\mu_{12})-\mu_{21}\mu_{22}-\mu_{21}(s-1)\theta]\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
=&[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2']\Big(\Sigma_1'\mu_{11}^{N-1}+(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\Big)\\
&-[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{21}\mu_{22}]\Sigma_1'\mu_{11}^{N-s}[\Sigma_2'-\mu_{22}]\alpha(s)\\
&-(\mu_{11}(\Sigma_2'-\mu_{12}))(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\\
&+(-\mu_{11}(\Sigma_2'-\mu_{12})-\mu_{21}\mu_{22}-\Sigma_1'(s-1)\theta)[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&+(\Sigma_1'-\mu_{21})(s-1)\theta[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\Sigma_2'+(N-1)\theta][f(1,N)+\mu_{11}\alpha(N)]\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\Sigma_2'+(N-1)\theta][\lambda\mu_{11}+(\lambda+1)\mu_{21}]\alpha(N)-\mu_{21}[\Sigma_2'+(N-1)\theta]\Sigma_1'\mu_{22}\alpha(N)\\
&+[-\mu_{11}(\Sigma_2'-\mu_{12})-\mu_{21}\mu_{22}-\mu_{21}(s-1)\theta]\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
=&[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2']\Big(\Sigma_1'\mu_{11}^{N-1}+(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\Big)\\
&-[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{21}\mu_{22}]\Sigma_1'\mu_{11}^{N-s}[\Sigma_2'-\mu_{22}]\alpha(s)\\
&-(\mu_{11}(\Sigma_2'-\mu_{12}))(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\\
&+(-\mu_{11}(\Sigma_2'-\mu_{12})-\mu_{21}\mu_{22}-\Sigma_1'(s-1)\theta)[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&+(\Sigma_1'-\mu_{21})(s-1)\theta[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\Sigma_2'+(N-1)\theta][f(1,N)]
\end{aligned}
$$

$$
\begin{aligned}
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\Sigma_2'+(N-1)\theta]\Big[\sum_{k=2}^{N+1-s}\mu_{11}^{k-1}f(k,N)+\sum_{k=N+2-s}^{N}\mu_{11}^{k-1}f(k,N)\Big]\\
&+[\Sigma_2'+(N-1)\theta][[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\lambda\mu_{11}+(\lambda+1)\mu_{21}]-\mu_{21}\Sigma_1'\mu_{22}]\alpha(N)\\
&+[-\mu_{11}(\Sigma_2'-\mu_{12})-\mu_{21}\mu_{22}-\mu_{21}(s-1)\theta]\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
=&[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2']\Big(\Sigma_1'\mu_{11}^{N-1}+(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\Big)\\
&-[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{21}\mu_{22}]\Sigma_1'\mu_{11}^{N-s}[\Sigma_2'-\mu_{22}]\alpha(s)\\
&-(\mu_{11}(\Sigma_2'-\mu_{12}))(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\\
&+(-\mu_{11}(\Sigma_2'-\mu_{12})-\mu_{21}\mu_{22}-\Sigma_1'(s-1)\theta)[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&+(\Sigma_1'-\mu_{21})(s-1)\theta[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\Sigma_2'+(N-1)\theta][f(1,N)]\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\Sigma_2'+(N-1)\theta]\Big[\sum_{k=2}^{N+1-s}\mu_{11}^{k-1}f(k,N)\Big]\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\Sigma_2'+(N-1)\theta]\Big[\sum_{k=N+2-s}^{N}\mu_{11}^{k-1}f(k,N)\Big]\\
&+[\Sigma_2'+(N-1)\theta][[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\lambda\mu_{11}+(\lambda+1)\mu_{21}]-\mu_{21}\Sigma_1'\mu_{22}]\alpha(N)\\
&+[-\mu_{11}(\Sigma_2'-\mu_{12})-\mu_{21}\mu_{22}-\mu_{21}(s-1)\theta]\Sigma_2'\mu_{11}^{N-s}f(1,s)
\end{aligned}
$$

Let $k'+N-s=k$

$$
\begin{aligned}
=&[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2']\Big(\Sigma_1'\mu_{11}^{N-1}+(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\Big)\\
&-[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{21}\mu_{22}]\Sigma_1'\mu_{11}^{N-s}[\Sigma_2'-\mu_{22}]\alpha(s)\\
&-(\mu_{11}(\Sigma_2'-\mu_{12}))(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\\
&+(-\mu_{11}(\Sigma_2'-\mu_{12})-\mu_{21}\mu_{22}-\Sigma_1'(s-1)\theta)[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&+(\Sigma_1'-\mu_{21})(s-1)\theta[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\Sigma_2'+(N-1)\theta][f(1,N)]\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\Sigma_2'+(N-1)\theta]\Big[\sum_{k=2}^{N+1-s}\mu_{11}^{k-1}f(k,N)\Big]\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\Sigma_2'+(N-1)\theta]\Big[\sum_{k'=2}^{s}\mu_{11}^{k'+N-1-s}f(k'+N-s,N)\Big]\\
&+[\Sigma_2'+(N-1)\theta][[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\lambda\mu_{11}+(\lambda+1)\mu_{21}]-\mu_{21}\Sigma_1'\mu_{22}]\alpha(N)
\end{aligned}
$$

$$
\begin{aligned}
&+[-\mu_{11}(\Sigma_2'-\mu_{12})-\mu_{21}\mu_{22}-\mu_{21}(s-1)\theta]\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
=&[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2']\Big(\Sigma_1'\mu_{11}^{N-1}+(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\Big)\\
&-[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{21}\mu_{22}]\Sigma_1'\mu_{11}^{N-s}[\Sigma_2'-\mu_{22}]\alpha(s)\\
&-(\mu_{11}(\Sigma_2'-\mu_{12}))(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\\
&+(-\mu_{11}(\Sigma_2'-\mu_{12})-\mu_{21}\mu_{22}-\Sigma_1'(s-1)\theta)[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&+(\Sigma_1'-\mu_{21})(s-1)\theta[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\Sigma_2'+(N-1)\theta]\Big[\sum_{k=2}^{N+1-s}\mu_{11}^{k-1}f(k,N)+f(1,N)\Big]\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\Sigma_2'+(N-1)\theta]\Big[\sum_{k'=2}^{s}\mu_{11}^{k'+N-1-s}f(k'+N-s,N)\Big]\\
&+[\Sigma_2'+(N-1)\theta][[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\lambda\mu_{11}+(\lambda+1)\mu_{21}]-\mu_{21}\Sigma_1'\mu_{22}]\alpha(N)\\
&+[-\mu_{11}(\Sigma_2'-\mu_{12})-\mu_{21}\mu_{22}-\mu_{21}(s-1)\theta]\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
=&[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2']\Big(\Sigma_1'\mu_{11}^{N-1}+(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\Big)\\
&-[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{21}\mu_{22}]\Sigma_1'\mu_{11}^{N-s}[\Sigma_2'-\mu_{22}]\alpha(s)\\
&-(\mu_{11}(\Sigma_2'-\mu_{12}))(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\\
&+(-\mu_{11}(\Sigma_2'-\mu_{12})-\mu_{21}\mu_{22}-\Sigma_1'(s-1)\theta)[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&+(\Sigma_1'-\mu_{21})(s-1)\theta[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\Sigma_2'+(N-1)\theta]\Big[\sum_{k=1}^{N+1-s}\mu_{11}^{k-1}f(k,N)\Big]\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\Sigma_2'+(N-1)\theta]\Big[\sum_{k'=2}^{s}\mu_{11}^{k'+N-1-s}f(k'+N-s,N)\Big]\\
&+[\Sigma_2'+(N-1)\theta][[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\lambda\mu_{11}+(\lambda+1)\mu_{21}]-\mu_{21}\Sigma_1'\mu_{22}]\alpha(N)\\
&+[-\mu_{11}(\Sigma_2'-\mu_{12})-\mu_{21}\mu_{22}-\mu_{21}(s-1)\theta]\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
=&[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2']\Big(\Sigma_1'\mu_{11}^{N-1}+(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\Big)\\
&-[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{21}\mu_{22}]\Sigma_1'\mu_{11}^{N-s}[\Sigma_2'-\mu_{22}]\alpha(s)\\
&-(\mu_{11}(\Sigma_2'-\mu_{12}))(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)
\end{aligned}
$$

$$
\begin{aligned}
&+(-\mu_{11}(\Sigma_2'-\mu_{12})-\mu_{21}\mu_{22}-\Sigma_1'(s-1)\theta)[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&+(\Sigma_1'-\mu_{21})(s-1)\theta[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\Sigma_2'+(N-1)\theta]\Big[\sum_{k=1}^{N-s}\mu_{11}^{k-1}f(k,N)+\mu_{11}^{N-s}f(N+1-s,N)\Big]\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\Sigma_2'+(N-1)\theta]\Big[\sum_{k'=2}^{s}\mu_{11}^{k'+N-1-s}f(k'+N-s,N)\Big]\\
&+[\Sigma_2'+(N-1)\theta][[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\lambda\mu_{11}+(\lambda+1)\mu_{21}]-\mu_{21}\Sigma_1'\mu_{22}]\alpha(N)\\
&+[-\mu_{11}(\Sigma_2'-\mu_{12})-\mu_{21}\mu_{22}-\mu_{21}(s-1)\theta]\Sigma_2'\mu_{11}^{N-s}f(1,s)
\end{aligned}
$$

Let $k''-s=k$

$$
\begin{aligned}
=&[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2']\Big(\Sigma_1'\mu_{11}^{N-1}+(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\Big)\\
&-[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{21}\mu_{22}]\Sigma_1'\mu_{11}^{N-s}[\Sigma_2'-\mu_{22}]\alpha(s)\\
&-(\mu_{11}(\Sigma_2'-\mu_{12}))(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\\
&+[-\mu_{11}(\Sigma_2'-\mu_{12})-\mu_{21}\mu_{22}-\mu_{21}(s-1)\theta][\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\Sigma_2'+(N-1)\theta]\Big[\sum_{k''=s+1}^{N}\mu_{11}^{k''-s-1}f(k''-s,N)+\mu_{11}^{N-s}f(N+1-s,N)\Big]\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\Sigma_2'+(N-1)\theta]\Big[\sum_{k'=2}^{s}\mu_{11}^{k'+N-1-s}f(k'+N-s,N)\Big]\\
&+[\Sigma_2'+(N-1)\theta][[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\lambda\mu_{11}+(\lambda+1)\mu_{21}]-\mu_{21}\Sigma_1'\mu_{22}]\alpha(N)\\
&+[-\mu_{11}(\Sigma_2'-\mu_{12})-\mu_{21}\mu_{22}-\mu_{21}(s-1)\theta]\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
=&[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2']\Big(\Sigma_1'\mu_{11}^{N-1}+(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\Big)\\
&-[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{21}\mu_{22}]\Sigma_1'\mu_{11}^{N-s}[\Sigma_2'-\mu_{22}]\alpha(s)\\
&-(\mu_{11}(\Sigma_2'-\mu_{12}))(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\\
&+\frac{[-\mu_{11}(\Sigma_2'-\mu_{12})-\mu_{21}\mu_{22}-\mu_{21}(s-1)\theta]}{\mu_{22}+(s-1)\theta}[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s+1)\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\Sigma_2'+(N-1)\theta]\Big[\sum_{k''=s+1}^{N}\mu_{11}^{k''-s-1}f(k''-s,N)+\mu_{11}^{N-s}f(N+1-s,N)\Big]\\
&+\mu_{22}\frac{[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}+\mu_{21}(s-1)\theta]}{\mu_{22}+(s-1)\theta}[\Sigma_2'+(N-1)\theta]\sum_{k''=s+1}^{N}\mu_{11}^{k''-s-1}f(k''-s,N)
\end{aligned}
$$

$$
\begin{aligned}
&- \mu_{22}\frac{[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}+\mu_{21}(s-1)\theta]}{\mu_{22}+(s-1)\theta}[\Sigma_2'+(N-1)\theta]\sum_{k''=s+1}^{N}\mu_{11}^{k''-s-1}f(k''-s,N)\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\Sigma_2'+(N-1)\theta]\Big[\sum_{k'=2}^{s}\mu_{11}^{k'+N-1-s}f(k'+N-s,N)\Big]\\
&+[\Sigma_2'+(N-1)\theta][[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\lambda\mu_{11}+(\lambda+1)\mu_{21}]-\mu_{21}\Sigma_1'\mu_{22}]\alpha(N)\\
&+[-\mu_{11}(\Sigma_2'-\mu_{12})-\mu_{21}\mu_{22}-\mu_{21}(s-1)\theta]\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
=&[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2']\Big(\Sigma_1'\mu_{11}^{N-1}+(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\Big)\\
&-[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{21}\mu_{22}]\Sigma_1'\mu_{11}^{N-s}[\Sigma_2'-\mu_{22}]\alpha(s)\\
&-(\mu_{11}(\Sigma_2'-\mu_{12}))(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\\
&+\frac{[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}+\mu_{21}(s-1)\theta]}{\mu_{22}+(s-1)\theta}[\Sigma_2'+(N-1)\theta]\mu_{22}\times\\
&\Big[\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}[f(k-s,N)-f(k,N)f(1,s+1)]\Big]\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\Sigma_2'+(N-1)\theta]\Big[\mu_{11}^{N-s}f(N+1-s,N)\Big]\\
&+\Big[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{22}\frac{[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}+\mu_{21}(s-1)\theta]}{\mu_{22}+(s-1)\theta}\Big][\Sigma_2'+(N-1)\theta]\times\\
&\sum_{k''=s+1}^{N}\mu_{11}^{k''-s-1}f(k''-s,N)\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\Sigma_2'+(N-1)\theta]\Big[\sum_{k'=2}^{s}\mu_{11}^{k'+N-1-s}f(k'+N-s,N)\Big]\\
&+[\Sigma_2'+(N-1)\theta]\Big[[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\lambda\mu_{11}+(\lambda+1)\mu_{21}]-\mu_{21}\Sigma_1'\mu_{22}\Big]\alpha(N)\\
&+[-\mu_{11}(\Sigma_2'-\mu_{12})-\mu_{21}\mu_{22}-\mu_{21}(s-1)\theta]\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
=&[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2']\Big(\Sigma_1'\mu_{11}^{N-1}+(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\Big)\\
&+\frac{[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}+\mu_{21}(s-1)\theta]}{\mu_{22}+(s-1)\theta}[\Sigma_2'+(N-1)\theta]\mu_{22}\times\\
&\Big[\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}[f(k-s,N)-f(k,N)f(1,s+1)]\Big]\\
&+\Big[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{22}\frac{[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}+\mu_{21}(s-1)\theta]}{\mu_{22}+(s-1)\theta}\Big][\Sigma_2'+(N-1)\theta]\times\\
&\sum_{k''=s+1}^{N}\mu_{11}^{k''-s-1}f(k''-s,N)\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}]\mu_{11}^{N-s}\Big[\Sigma_2'\sum_{k'=2}^{s}\mu_{11}^{k'-1}f(k'+N-s,N)-\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(s)\Big]
\end{aligned}
$$

$$
\begin{aligned}
&- (\mu_{11}(\Sigma_2' - \mu_{12}))(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s) \\
&+ [\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21}][\Sigma_2' + (N-1)\theta]\Big[\mu_{11}^{N-s} f(N+1-s,N)\Big] \\
&+ [\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21}](N-1)\theta\Big[\sum_{k'=2}^{s}\mu_{11}^{k'+N-1-s} f(k'+N-s,N)\Big] \\
&+ [\Sigma_2' + (N-1)\theta]\Big[[\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21}][\lambda\mu_{11} + (\lambda+1)\mu_{21}] - \mu_{21}\Sigma_1'\mu_{22}\Big]\alpha(N) \\
&+ [-\mu_{11}(\Sigma_2' - \mu_{12}) - \mu_{21}\mu_{22} - \mu_{21}(s-1)\theta]\Sigma_2'\mu_{11}^{N-s} f(1,s) \\
=&[\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21} - \mu_{21}\Sigma_2']\Sigma_1'\mu_{11}^{N-1} \\
+&\frac{[\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21} + \mu_{21}(s-1)\theta]}{\mu_{22} + (s-1)\theta}[\Sigma_2' + (N-1)\theta]\mu_{22}\times \\
&\Big[\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}[f(k-s,N) - f(k,N)f(1,s+1)]\Big] \\
+&\Big[\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21} - \mu_{22}\frac{[\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21} + \mu_{21}(s-1)\theta]}{\mu_{22} + (s-1)\theta}\Big][\Sigma_2' + (N-1)\theta]\times \\
&\sum_{k''=s+1}^{N}\mu_{11}^{k''-s-1} f(k''-s,N) \\
&+ [\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21}]\mu_{11}^{N-s}\Big[\Sigma_2'\sum_{k'=2}^{s}\mu_{11}^{k'-1} f(k'+N-s,N) - \Sigma_1'[\Sigma_2' - \mu_{22}]\alpha(s)\Big] \\
&- (\mu_{21}\Sigma_2' - \mu_{22}\mu_{21})(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s) \\
&+ [\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21}][\Sigma_2' + (N-1)\theta]\Big[\mu_{11}^{N-s} f(N+1-s,N)\Big] \\
&+ (\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21} - \mu_{21}\Sigma_2' + \mu_{22}\mu_{21})(N-1)\theta\Big[\sum_{k'=2}^{s}\mu_{11}^{k'+N-1-s} f(k'+N-s,N)\Big] \\
&+ (\mu_{21}\Sigma_2' - \mu_{22}\mu_{21})(N-1)\theta\Big[\sum_{k'=2}^{s}\mu_{11}^{k'+N-1-s} f(k'+N-s,N)\Big] \\
&+ [\Sigma_2' + (N-1)\theta]\Big[[\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21}][\lambda\mu_{11} + (\lambda+1)\mu_{21}] - \mu_{21}\Sigma_1'\mu_{22}\Big]\alpha(N) \\
&+ [-\mu_{11}(\Sigma_2' - \mu_{12}) - \mu_{21}\mu_{22} - \mu_{21}(s-1)\theta]\Sigma_2'\mu_{11}^{N-s} f(1,s) \\
\\
=&[\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21} - \mu_{21}\Sigma_2']\Sigma_1'\mu_{11}^{N-1} \\
+&\frac{[\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21} + \mu_{21}(s-1)\theta]}{\mu_{22} + (s-1)\theta}[\Sigma_2' + (N-1)\theta]\mu_{22}\times \\
&\Big[\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}[f(k-s,N) - f(k,N)f(1,s+1)]\Big] \\
+&\Big[\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21} - \mu_{22}\frac{[\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21} + \mu_{21}(s-1)\theta]}{\mu_{22} + (s-1)\theta}\Big][\Sigma_2' + (N-1)\theta]\times
\end{aligned}
$$

$$
\begin{aligned}
&\sum_{k''=s+1}^{N} \mu_{11}^{k''-s-1} f(k''-s,N) \\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}]\mu_{11}^{N-s}\Big[\Sigma_2'\sum_{k'=2}^{s}\mu_{11}^{k'-1}f(k'+N-s,N)-\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(s)\Big] \\
&-(\mu_{21}\Sigma_2'-\mu_{22}\mu_{21})(s-1)\theta\mu_{11}\mu_{11}^{N-s}\alpha(s) \\
&-(\mu_{21}\Sigma_2'-\mu_{22}\mu_{21})(s-1)\theta[\lambda\mu_{11}+(\lambda+1)\mu_{21}]\mu_{11}^{N-s}\alpha(s) \\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\Sigma_2'+(N-1)\theta]\Big[\mu_{11}^{N-s}f(N+1-s,N)\Big] \\
&+(\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2'+\mu_{22}\mu_{21})(N-1)\theta\Big[\sum_{k'=2}^{s}\mu_{11}^{k'+N-1-s}f(k'+N-s,N)\Big] \\
&+(\mu_{21}\Sigma_2'-\mu_{22}\mu_{21})(N-1)\theta\Big[\sum_{k'=2}^{s}\mu_{11}^{k'+N-1-s}f(k'+N-s,N)\Big] \\
&+[\Sigma_2'+(N-1)\theta]\Big[[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\lambda\mu_{11}+(\lambda+1)\mu_{21}]-\mu_{21}\Sigma_1'\mu_{22}\Big]\alpha(N) \\
&+[-\mu_{11}(\Sigma_2'-\mu_{12})-\mu_{21}\mu_{22}-\mu_{21}(s-1)\theta]\Sigma_2'\mu_{11}^{N-s}f(1,s) \\
=&[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2']\Sigma_1'\mu_{11}^{N-1} \\
&+\frac{[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}+\mu_{21}(s-1)\theta]}{\mu_{22}+(s-1)\theta}[\Sigma_2'+(N-1)\theta]\mu_{22}\times \\
&\Big[\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}[f(k-s,N)-f(k,N)f(1,s+1)]\Big] \\
&+\Big[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{22}\frac{[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}+\mu_{21}(s-1)\theta]}{\mu_{22}+(s-1)\theta}\Big][\Sigma_2'+(N-1)\theta]\times \\
&\sum_{k''=s+1}^{N}\mu_{11}^{k''-s-1}f(k''-s,N) \\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}]\mu_{11}^{N-s}\Big[\Sigma_2'\sum_{k'=2}^{s}\mu_{11}^{k'-1}f(k'+N-s,N)-\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(s)\Big] \\
&+(\mu_{21}\Sigma_2'-\mu_{22}\mu_{21})\mu_{11}^{N-s+1}\Big[(N-1)\theta\sum_{k'=2}^{s}\mu_{11}^{k'-2}f(k'+N-s,N)-(s-1)\theta\alpha(s)\Big] \\
&-(\mu_{21}\Sigma_2'-\mu_{22}\mu_{21})(s-1)\theta[\lambda\mu_{11}+(\lambda+1)\mu_{21}]\mu_{11}^{N-s}\alpha(s) \\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\Sigma_2'+(N-1)\theta]\Big[\mu_{11}^{N-s}f(N+1-s,N)\Big] \\
&+(\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2'+\mu_{22}\mu_{21})(N-1)\theta\Big[\sum_{k'=2}^{s}\mu_{11}^{k'+N-1-s}f(k'+N-s,N)\Big] \\
&+[\Sigma_2'+(N-1)\theta]\Big[[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\lambda\mu_{11}+(\lambda+1)\mu_{21}]-\mu_{21}\Sigma_1'\mu_{22}\Big]\alpha(N) \\
&+[-\mu_{11}(\Sigma_2'-\mu_{12})-\mu_{21}\mu_{22}-\mu_{21}(s-1)\theta]\Sigma_2'\mu_{11}^{N-s}f(1,s)
\end{aligned}
$$

$$
\begin{aligned}
&=[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2']\Sigma_1'\mu_{11}^{N-1}\\
&+\frac{[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}+\mu_{21}(s-1)\theta]}{\mu_{22}+(s-1)\theta}[\Sigma_2'+(N-1)\theta]\mu_{22}\times\\
&\quad\Big[\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}[f(k-s,N)-f(k,N)f(1,s+1)]\Big]\\
&+\Big[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{22}\frac{[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}+\mu_{21}(s-1)\theta]}{\mu_{22}+(s-1)\theta}\Big][\Sigma_2'+(N-1)\theta]\times\\
&\quad\sum_{k''=s+1}^{N}\mu_{11}^{k''-s-1}f(k''-s,N)\\
&\quad+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}]\mu_{11}^{N-s}\Big[\Sigma_2'\sum_{k'=2}^{s}\mu_{11}^{k'-1}f(k'+N-s,N)-\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(s)\Big]\\
&\quad+(\mu_{21}\Sigma_2'-\mu_{22}\mu_{21})\mu_{11}^{N-s+1}\Big[(N-1)\theta\sum_{k'=2}^{s}\mu_{11}^{k'-2}f(k'+N-s,N)-(s-1)\theta\alpha(s)\Big]\\
&\quad-(\mu_{21}\Sigma_2'-\mu_{22}\mu_{21})(s-1)\theta[\lambda\mu_{11}+(\lambda+1)\mu_{21}]\mu_{11}^{N-s}\alpha(s)\\
&\quad+(N-1)\theta[[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\lambda\mu_{11}+(\lambda+1)\mu_{21}]-\mu_{21}\Sigma_1'\mu_{22}]\alpha(N)\\
&\quad+\mu_{22}\mu_{21}(N-1)\theta\Big[\sum_{k'=2}^{s}\mu_{11}^{k'+N-1-s}f(k'+N-s,N)\Big]\\
&\quad+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\Sigma_2'+(N-1)\theta]\Big[\mu_{11}^{N-s}f(N+1-s,N)\Big]\\
&\quad+(\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2')(N-1)\theta\Big[\sum_{k'=2}^{s}\mu_{11}^{k'+N-1-s}f(k'+N-s,N)\Big]\\
&\quad+\Sigma_2'\Big[[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\lambda\mu_{11}+(\lambda+1)\mu_{21}]-\mu_{21}\Sigma_1'\mu_{22}\Big]\alpha(N)\\
&\quad+[-\mu_{11}(\Sigma_2'-\mu_{12})-\mu_{21}\mu_{22}-\mu_{21}(s-1)\theta]\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
&=[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2']\Sigma_1'\mu_{11}^{N-1}\\
&+\frac{[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}+\mu_{21}(s-1)\theta]}{\mu_{22}+(s-1)\theta}[\Sigma_2'+(N-1)\theta]\mu_{22}\times\\
&\quad\Big[\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}[f(k-s,N)-f(k,N)f(1,s+1)]\Big]\\
&+\Big[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{22}\frac{[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}+\mu_{21}(s-1)\theta]}{\mu_{22}+(s-1)\theta}\Big][\Sigma_2'+(N-1)\theta]\times\\
&\quad\sum_{k''=s+1}^{N}\mu_{11}^{k''-s-1}f(k''-s,N)\\
&\quad+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}]\mu_{11}^{N-s}\Big[\Sigma_2'\sum_{k'=2}^{s}\mu_{11}^{k'-1}f(k'+N-s,N)-\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(s)\Big]
\end{aligned}
$$

$$
\begin{aligned}
&+(\mu_{21}\Sigma_2'-\mu_{22}\mu_{21})\mu_{11}^{N-s+1}\Big[(N-1)\theta\sum_{k'=2}^{s}\mu_{11}^{k'-2}f(k'+N-s,N)-(s-1)\theta\alpha(s)\Big]\\
&-(\mu_{21}\Sigma_2'-\mu_{22}\mu_{21})(s-1)\theta[\lambda\mu_{11}+(\lambda+1)\mu_{21}]\mu_{11}^{N-s}\alpha(s)\\
&+(\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2')(s-1)\theta[\lambda\mu_{11}+(\lambda+1)\mu_{21}]\mu_{11}^{N-s}\alpha(s)\\
&-[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2'](s-1)\theta[\lambda\mu_{11}+(\lambda+1)\mu_{21}]\mu_{11}^{N-s}\alpha(s)\\
&+(N-1)\theta[[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\lambda\mu_{11}+(\lambda+1)\mu_{21}]-\mu_{21}\Sigma_1'\mu_{22}]\alpha(N)\\
&+\mu_{22}\mu_{21}(N-1)\theta\Big[\sum_{k'=2}^{s}\mu_{11}^{k'+N-1-s}f(k'+N-s,N)\Big]\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\Sigma_2'+(N-1)\theta]\Big[\mu_{11}^{N-s}f(N+1-s,N)\Big]\\
&+(\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2')(N-1)\theta\Big[\sum_{k'=2}^{s}\mu_{11}^{k'+N-1-s}f(k'+N-s,N)\Big]\\
&+\Sigma_2'\Big[[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\lambda\mu_{11}+(\lambda+1)\mu_{21}]-\mu_{21}\Sigma_1'\mu_{22}\Big]\alpha(N)\\
&+[-\mu_{11}(\Sigma_2'-\mu_{12})-\mu_{21}\mu_{22}-\mu_{21}(s-1)\theta]\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
=&[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2']\Sigma_1'\mu_{11}^{N-1}\\
&+\frac{[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}+\mu_{21}(s-1)\theta]}{\mu_{22}+(s-1)\theta}[\Sigma_2'+(N-1)\theta]\mu_{22}\times\\
&\Big[\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}[f(k-s,N)-f(k,N)f(1,s+1)]\Big]\\
&+\Big[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{22}\frac{[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}+\mu_{21}(s-1)\theta]}{\mu_{22}+(s-1)\theta}\Big][\Sigma_2'+(N-1)\theta]\times\\
&\sum_{k''=s+1}^{N}\mu_{11}^{k''-s-1}f(k''-s,N)\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}]\mu_{11}^{N-s}\Big[\Sigma_2'\sum_{k'=2}^{s}\mu_{11}^{k'-1}f(k'+N-s,N)-\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(s)\Big]\\
&+(\mu_{21}\Sigma_2'-\mu_{22}\mu_{21})\mu_{11}^{N-s+1}\Big[(N-1)\theta\sum_{k'=2}^{s}\mu_{11}^{k'-2}f(k'+N-s,N)-(s-1)\theta\alpha(s)\Big]\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2'](s-1)\theta[\lambda\mu_{11}+(\lambda+1)\mu_{21}]\mu_{11}^{N-s}\alpha(s)\\
&-(\mu_{11}(\Sigma_2'-\mu_{12}))(s-1)\theta[\lambda\mu_{11}+(\lambda+1)\mu_{21}]\mu_{11}^{N-s}\alpha(s)\\
&+(N-1)\theta[[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\lambda\mu_{11}+(\lambda+1)\mu_{21}]-\mu_{21}\Sigma_1'\mu_{22}]\alpha(N)\\
&+\mu_{22}\mu_{21}(N-1)\theta\Big[\sum_{k'=2}^{s}\mu_{11}^{k'+N-1-s}f(k'+N-s,N)\Big]\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\Sigma_2'+(N-1)\theta]\Big[\mu_{11}^{N-s}f(N+1-s,N)\Big]
\end{aligned}
$$

$$
\begin{aligned}
&+(\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2')(N-1)\theta\Big[\sum_{k'=2}^{s}\mu_{11}^{k'+N-1-s}f(k'+N-s,N)\Big]\\
&+\Sigma_2'\Big[[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\lambda\mu_{11}+(\lambda+1)\mu_{21}]-\mu_{21}\Sigma_1'\mu_{22}\Big]\alpha(N)\\
&+[-\mu_{11}(\Sigma_2'-\mu_{12})-\mu_{21}\mu_{22}-\mu_{21}(s-1)\theta]\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
=&[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2']\Sigma_1'\mu_{11}^{N-1}\\
+&\frac{[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}+\mu_{21}(s-1)\theta]}{\mu_{22}+(s-1)\theta}[\Sigma_2'+(N-1)\theta]\mu_{22}\times\\
&\Big[\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}[f(k-s,N)-f(k,N)f(1,s+1)]\Big]\\
+&\Big[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{22}\frac{[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}+\mu_{21}(s-1)\theta]}{\mu_{22}+(s-1)\theta}\Big][\Sigma_2'+(N-1)\theta]\times\\
&\sum_{k''=s+1}^{N}\mu_{11}^{k''-s-1}f(k''-s,N)\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}]\mu_{11}^{N-s}\Big[\Sigma_2'\sum_{k'=2}^{s}\mu_{11}^{k'-1}f(k'+N-s,N)-\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(s)\Big]\\
&+(\mu_{21}\Sigma_2'-\mu_{22}\mu_{21})\mu_{11}^{N-s+1}\Big[(N-1)\theta\sum_{k'=2}^{s}\mu_{11}^{k'-2}f(k'+N-s,N)-(s-1)\theta\alpha(s)\Big]\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2'](s-1)\theta[\lambda\mu_{11}+(\lambda+1)\mu_{21}]\mu_{11}^{N-s}\alpha(s)\\
&-(\mu_{11}(\Sigma_2'-\mu_{12}))(s-1)\theta[\lambda\mu_{11}+(\lambda+1)\mu_{21}]\mu_{11}^{N-s}\alpha(s)\\
&+(N-1)\theta\Big[[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\lambda\mu_{11}+(\lambda+1)\mu_{21}]-\mu_{21}\Sigma_1'\mu_{22}\Big]\alpha(N)\\
&+\mu_{22}\mu_{21}(N-1)\theta\Big[\sum_{k'=2}^{s}\mu_{11}^{k'+N-1-s}f(k'+N-s,N)\Big]\\
&+(\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2')(N-1)\theta\Big[\sum_{k'=2}^{s}\mu_{11}^{k'+N-1-s}f(k'+N-s,N)\Big]\\
&+\Sigma_2'\Big[[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\lambda\mu_{11}+(\lambda+1)\mu_{21}]-\mu_{21}\Sigma_1'\mu_{22}\Big]\alpha(N)\\
&+[-\mu_{11}(\Sigma_2'-\mu_{12})-\mu_{21}\mu_{22}-\mu_{21}(s-1)\theta]\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\Sigma_2'+(N-1)\theta]\Big[\mu_{11}^{N-s}f(N+1-s,N)\Big]\\
=&[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2']\Sigma_1'\mu_{11}^{N-1}\\
+&\frac{[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}+\mu_{21}(s-1)\theta]}{\mu_{22}+(s-1)\theta}[\Sigma_2'+(N-1)\theta]\mu_{22}\times\\
&\Big[\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}[f(k-s,N)-f(k,N)f(1,s+1)]\Big]
\end{aligned}
$$

$$
\begin{aligned}
&+\left[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{22}\frac{[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}+\mu_{21}(s-1)\theta]}{\mu_{22}+(s-1)\theta}\right][\Sigma_2'+(N-1)\theta]\times\\
&\quad\sum_{k''=s+1}^{N}\mu_{11}^{k''-s-1}f(k''-s,N)\\
&\quad+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}]\mu_{11}^{N-s}\left[\Sigma_2'\sum_{k'=2}^{s}\mu_{11}^{k'-1}f(k'+N-s,N)-\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(s)\right]\\
&\quad+(\mu_{21}\Sigma_2'-\mu_{22}\mu_{21})\mu_{11}^{N-s+1}\left[(N-1)\theta\sum_{k'=2}^{s}\mu_{11}^{k'-2}f(k'+N-s,N)-(s-1)\theta\alpha(s)\right]\\
&\quad+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2'](s-1)\theta[\lambda\mu_{11}+(\lambda+1)\mu_{21}]\mu_{11}^{N-s}\alpha(s)\\
&\quad-(\mu_{11}(\Sigma_2'-\mu_{12}))(s-1)\theta[\lambda\mu_{11}+(\lambda+1)\mu_{21}]\mu_{11}^{N-s}\alpha(s)\\
&\quad+(N-1)\theta\left[[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\lambda\mu_{11}+(\lambda+1)\mu_{21}]-\mu_{21}\Sigma_1'\mu_{22}\right]\alpha(N)\\
&\quad+\mu_{22}\mu_{21}(N-1)\theta\left[\sum_{k'=2}^{s}\mu_{11}^{k'+N-1-s}f(k'+N-s,N)\right]\\
&\quad+(\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2')(N-1)\theta\left[\sum_{k'=2}^{s}\mu_{11}^{k'+N-1-s}f(k'+N-s,N)\right]\\
&\quad+\Sigma_2'\left[[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\lambda\mu_{11}+(\lambda+1)\mu_{21}]-\mu_{21}\Sigma_1'\mu_{22}\right]\alpha(N)\\
&\quad-[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{21}\mu_{22}]\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
&\quad-\mu_{21}(s-1)\theta\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
&\quad+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}]\Sigma_2'\left[\mu_{11}^{N-s}f(N+1-s,N)\right]\\
&\quad+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}](N-1)\theta\left[\mu_{11}^{N-s}f(N+1-s,N)\right]\\
&=[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2']\Sigma_1'\mu_{11}^{N-1}\\
&+\frac{[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}+\mu_{21}(s-1)\theta]}{\mu_{22}+(s-1)\theta}[\Sigma_2'+(N-1)\theta]\mu_{22}\times\\
&\quad\left[\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}[f(k-s,N)-f(k,N)f(1,s+1)]\right]\\
&+\left[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{22}\frac{[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}+\mu_{21}(s-1)\theta]}{\mu_{22}+(s-1)\theta}\right][\Sigma_2'+(N-1)\theta]\times\\
&\quad\sum_{k''=s+1}^{N}\mu_{11}^{k''-s-1}f(k''-s,N)\\
&\quad+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}]\mu_{11}^{N-s}\left[\Sigma_2'\sum_{k'=2}^{s}\mu_{11}^{k'-1}f(k'+N-s,N)-\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(s)\right]\\
&\quad+(\mu_{21}\Sigma_2'-\mu_{22}\mu_{21})\mu_{11}^{N-s+1}\left[(N-1)\theta\sum_{k'=2}^{s}\mu_{11}^{k'-2}f(k'+N-s,N)-(s-1)\theta\alpha(s)\right]
\end{aligned}
$$

$$
\begin{aligned}
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2'](s-1)\theta[\lambda\mu_{11}+(\lambda+1)\mu_{21}]\mu_{11}^{N-s}\alpha(s)\\
&-(\mu_{11}(\Sigma_2'-\mu_{12}))(s-1)\theta[\lambda\mu_{11}+(\lambda+1)\mu_{21}]\mu_{11}^{N-s}\alpha(s)\\
&+(N-1)\theta\Big[[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\lambda\mu_{11}+(\lambda+1)\mu_{21}]-\mu_{21}\Sigma_1'\mu_{22}\Big]\alpha(N)\\
&+\mu_{22}\mu_{21}(N-1)\theta\Big[\sum_{k'=2}^{s}\mu_{11}^{k'+N-1-s}f(k'+N-s,N)\Big]\\
&+(\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2')(N-1)\theta\Big[\sum_{k'=2}^{s}\mu_{11}^{k'+N-1-s}f(k'+N-s,N)\Big]\\
&+\Sigma_2'\Big[[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}][\lambda\mu_{11}+(\lambda+1)\mu_{21}]-\mu_{21}\Sigma_1'\mu_{22}\Big]\alpha(N)\\
&-\mu_{21}(s-1)\theta\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}]\Sigma_2'\mu_{11}^{N-s}\Big[f(N+1-s,N)-f(1,s)\Big]\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2'](N-1)\theta\Big[\mu_{11}^{N-s}f(N+1-s,N)\Big]\\
&+\mu_{21}\Sigma_2'(N-1)\theta\Big[\mu_{11}^{N-s}f(N+1-s,N)\Big]\\
=&[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2']\Sigma_1'\mu_{11}^{N-1} &(57)\\
&+\frac{[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}+\mu_{21}(s-1)\theta]}{\mu_{22}+(s-1)\theta}[\Sigma_2'+(N-1)\theta]\mu_{22}\\
&\times\Big[\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}[f(k-s,N)-f(k,N)f(1,s+1)]\Big] &(58)\\
&+\Big[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{22}\frac{[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}+\mu_{21}(s-1)\theta]}{\mu_{22}+(s-1)\theta}\Big][\Sigma_2'+(N-1)\theta]\\
&\times\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k-s,N) &(59)\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}]\mu_{11}^{N-s}\Big[\Sigma_2'\sum_{k=2}^{s}\mu_{11}^{k-1}f(k+N-s,N)-\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(s)\Big] &(60)\\
&+\Sigma_2'\big[[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}](\Sigma_1'-\mu_{11})-\mu_{21}\Sigma_1'\mu_{22}\big]\alpha(N) &(61)\\
&+\mu_{21}(\Sigma_2'-\mu_{22})\mu_{11}^{N-s+1}\Big[(N-1)\theta\sum_{k=2}^{s}\mu_{11}^{k-2}f(k+N-s,N)-(s-1)\theta\alpha(s)\Big] &(62)\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{21}\Sigma_2'](s-1)\theta(\Sigma_1'-\mu_{11})\mu_{11}^{N-s}\alpha(s) &(63)\\
&-\mu_{11}(\Sigma_2'-\mu_{12})(s-1)\theta(\Sigma_1'-\mu_{11})\mu_{11}^{N-s}\alpha(s) &(64)\\
&+(N-1)\theta\big[[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}](\Sigma_1'-\mu_{11})-\mu_{21}\Sigma_1'\mu_{22}\big]\alpha(N) &(65)\\
&+\mu_{22}\mu_{21}(N-1)\theta\Big[\sum_{k=2}^{s}\mu_{11}^{N+k-s-1}f(k+N-s,N)\Big] &(66)
\end{aligned}
$$

$$+ [\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21} - \mu_{21}\Sigma_2'](N-1)\theta\Big[\sum_{k=2}^{s} \mu_{11}^{N+k-s-1} f(k+N-s, N)\Big] \tag{67}$$

$$+ \mu_{21}\Sigma_2'\mu_{11}^{N-s}\theta\Big[(N-1)f(N+1-s, N) - (s-1)f(1,s)\Big] \tag{68}$$

$$+ [\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21}]\Sigma_2'\mu_{11}^{N-s}\Big[f(N+1-s, N) - f(1,s)\Big] \tag{69}$$

$$+ [\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21} - \mu_{21}\Sigma_2'](N-1)\theta\mu_{11}^{N-s} f(N+1-s, N). \tag{70}$$

We now show that the sum of the terms (57)-(70) is positive. Note that as $N \geq 2$ for this case to be possible, we have $\tau'(2) \geq 0$. Note that

$$\begin{aligned}
\tau'(2) &= [\Sigma_2' + \theta]\Sigma_2'\mu_{22} - [\Sigma_2' + \theta][\Sigma_1'(\Sigma_2' - \mu_{22}) + \Sigma_2'\mu_{22}] + \mu_{11}(\Sigma_2')^2 \\
&= -[\Sigma_2' + \theta][\Sigma_1'(\Sigma_2' - \mu_{22})] + \mu_{11}(\Sigma_2')^2 \\
&= -\Sigma_2'\Sigma_1'(\Sigma_2' - \mu_{22}) + (\Sigma_2')^2\mu_{11} - \theta\Sigma_1'(\Sigma_2' - \mu_{22}) \qquad (71) \\
&= -\Sigma_2'\Sigma_1'(\lambda\Sigma_2 + \mu_{12}) + (\Sigma_2')^2\mu_{11} - \theta\Sigma_1'(\Sigma_2' - \mu_{22}) \\
&= \Sigma_2'(1+\lambda)[-\lambda\Sigma_1\Sigma_2 - \mu_{12}\Sigma_1 + \Sigma_2\mu_{11}] - \theta\Sigma_1'(\Sigma_2' - \mu_{22}) \\
&= \Sigma_2'(1+\lambda)[-\lambda\Sigma_1\Sigma_2 + \mu_{11}\mu_{22} - \mu_{21}\mu_{12}] - \theta\Sigma_1'(\Sigma_2' - \mu_{22}). \qquad (72)
\end{aligned}$$

That is, $\tau'(2) \geq 0$ implies that $-\lambda\Sigma_1\Sigma_2 + \mu_{11}\mu_{22} - \mu_{21}\mu_{12} \geq 0$, and hence

$$\lambda \leq \frac{\mu_{11}\mu_{22} - \mu_{21}\mu_{12}}{\Sigma_1\Sigma_2}. \tag{73}$$

For (57) we have

$$\begin{aligned}
&\mu_{11}(\Sigma_2' - \mu_{12}) + \mu_{22}\mu_{21} - \mu_{21}\Sigma_2' \\
=&\mu_{11}[(\lambda+1)\mu_{22} + \lambda\mu_{12}] + \mu_{22}\mu_{21} - \mu_{21}[(\lambda+1)\mu_{22} + (\lambda+1)\mu_{12}] \\
=&(\lambda+1)\mu_{11}\mu_{22} + \lambda\mu_{11}\mu_{12} + \mu_{22}\mu_{21} - (\lambda+1)\mu_{21}\mu_{22} - (\lambda+1)\mu_{21}\mu_{12} \\
=&(\lambda+1)(\mu_{11}\mu_{22} - \mu_{21}\mu_{12}) + \lambda(\mu_{11}\mu_{12} - \mu_{21}\mu_{22}) \\
=&(\mu_{11}\mu_{22} - \mu_{21}\mu_{12}) + \lambda(\mu_{11}\mu_{22} - \mu_{21}\mu_{12} + \mu_{11}\mu_{12} - \mu_{21}\mu_{22}) \\
=&(\mu_{11}\mu_{22} - \mu_{21}\mu_{12}) + \lambda(\mu_{11}\mu_{22} + \mu_{11}\mu_{12}) - \lambda\mu_{21}\Sigma_2 \\
\overset{(*)}{\geq}&(\mu_{11}\mu_{22} - \mu_{21}\mu_{12}) + \lambda(\mu_{11}\mu_{22} + \mu_{11}\mu_{12}) - \frac{\mu_{21}}{\Sigma_1}(\mu_{11}\mu_{22} - \mu_{21}\mu_{12}) \\
=&(\mu_{11}\mu_{22} - \mu_{21}\mu_{12})\Big[1 - \frac{\mu_{21}}{\Sigma_1}\Big] + \lambda(\mu_{11}\mu_{22} + \mu_{11}\mu_{12}) \\
=&(\mu_{11}\mu_{22} - \mu_{21}\mu_{12})\frac{\mu_{11}}{\Sigma_1} + \lambda(\mu_{11}\mu_{22} + \mu_{11}\mu_{12}) \\
\overset{(**)}{\geq}&0, \qquad (74)
\end{aligned}$$

where in $(*)$ we have used (73) and in $(**)$ we have used $\mu_{11}\mu_{22} - \mu_{21}\mu_{12} \geq 0$. It follows that (57) is non-negative.

The non-negativity of (58) follows because for $s+1 \leq k \leq N$, we have

$$\begin{aligned}
&f(k-s, N) - f(1, s+1)f(k, N) \\
&= f(k-s, k)f(k, N) - f(1, s+1)f(k, N) \\
&\geq f(1, s+1)f(k, N) - f(1, s+1)f(k, N) = 0.
\end{aligned}$$

The non-negativity of (59) follows from

$$\begin{aligned}
&\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}-\mu_{22}\frac{[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}+\mu_{21}(s-1)\theta]}{\mu_{22}+(s-1)\theta}\\
\geq&\frac{\mu_{22}\mu_{21}(s-1)\theta-\mu_{22}\mu_{21}(s-1)\theta+\mu_{22}[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}]-\mu_{22}[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}]}{\mu_{22}+(s-1)\theta}\\
=&0.
\end{aligned}$$

Next we show (60)+(61) is positive. From (60) we have

$$\begin{aligned}
&[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}]\mu_{11}^{N-s}\left[\Sigma_2'\sum_{k=2}^{s}\mu_{11}^{k-1}f(k+N-s,N)-\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(s)\right]\\
\overset{(*)}{\geq}&[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}]\mu_{11}^{N-s}\left[\Sigma_2'\mu_{11}\sum_{k=2}^{s}\mu_{11}^{k-2}f(k,s)-\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(s)\right]\\
=&[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}]\mu_{11}^{N-s}\left[\Sigma_2'\mu_{11}\alpha(s)-\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(s)\right]\\
=&[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}]\mu_{11}^{N-s}\alpha(s)[\Sigma_2'\mu_{11}-\Sigma_1'\Sigma_2'+\Sigma_1'\mu_{22}],
\end{aligned}$$

where $(*)$ follows since

$$\sum_{k=2}^{s}\mu_{11}^{k-2}f(k+N-s,N)\geq\sum_{k=2}^{s}\mu_{11}^{k-2}f(k,s)=\alpha(s)\text{ for }s\leq N-1. \tag{75}$$

And for (61) we have

$$\begin{aligned}
&[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}](\Sigma_1'-\mu_{11})-\mu_{21}\Sigma_1'\mu_{22}\\
=&[\mu_{11}(1+\lambda)\mu_{22}+\lambda\mu_{11}\mu_{12}+\mu_{22}\mu_{21}][\lambda\mu_{11}+(\lambda+1)\mu_{21}]-(1+\lambda)\mu_{21}^2\mu_{22}-(1+\lambda)\mu_{21}\mu_{22}\mu_{11}\\
\geq&(1+\lambda)\mu_{21}^2\mu_{22}+(1+\lambda)^2\mu_{22}\mu_{11}\mu_{21}-(1+\lambda)\mu_{21}^2\mu_{22}-(1+\lambda)\mu_{21}\mu_{22}\mu_{11}\\
=&0.
\end{aligned} \tag{76}$$

It now follows from (37) that (61) satisfies

$$\begin{aligned}
&\Sigma_2'\left[[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}](\Sigma_1'-\mu_{11})-\mu_{21}\Sigma_1'\mu_{22}\right]\alpha(N)\\
\geq&\Sigma_2'\left[[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}](\Sigma_1'-\mu_{11})-\mu_{21}\Sigma_1'\mu_{22}\right]\mu_{11}^{N-s}\alpha(s).
\end{aligned}$$

That is

$$\begin{aligned}
(60)+(61)\geq&\mu_{11}^{N-s}\alpha(s)\Big[\Sigma_2'\left[[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}](\Sigma_1'-\mu_{11})-\mu_{21}\Sigma_1'\mu_{22}\right]\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{21}\mu_{22}][\Sigma_2'\mu_{11}-\Sigma_1'\Sigma_2'+\Sigma_1'\mu_{22}]\Big]\\
=&\mu_{11}^{N-s}\alpha(s)\Big[\Sigma_2'\left[[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}](\Sigma_1'-\mu_{11})-\mu_{21}\Sigma_1'\mu_{22}\right]\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{21}\mu_{22}][\Sigma_2'(\mu_{11}-\Sigma_1')+\Sigma_1'\mu_{22}]\Big]\\
=&\mu_{11}^{N-s}\alpha(s)\Big[\Sigma_2'[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}](\Sigma_1'-\mu_{11})-\Sigma_2'\Sigma_1'\mu_{22}\mu_{21}\\
&-\Sigma_2'[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}](\Sigma_1'-\mu_{11})+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{21}\mu_{22}]\Sigma_1'\mu_{22}\Big]
\end{aligned}$$

$$
\begin{aligned}
=&\mu_{11}^{N-s}\alpha(s)\Big[-\Sigma_2'\Sigma_1'\mu_{22}\mu_{21}+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{21}\mu_{22}]\Sigma_1'\mu_{22}\Big]\\
=&\mu_{11}^{N-s}\alpha(s)\Sigma_1'\mu_{22}[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{21}\mu_{22}-\Sigma_2'\mu_{21}]\geq 0,
\end{aligned}
$$

where the last inequality follows from (74). Hence (60) + (61) ≥ 0.

The non-negativity of (62) follows since $1 \leq s \leq N-1$ implies that

$$
\begin{aligned}
&(N-1)\theta\sum_{k=2}^{s}\mu_{11}^{k-2}f(k+N-s,N)-(s-1)\theta\alpha(s)\\
\geq&(s-1)\theta\Big[\sum_{k=2}^{s}\mu_{11}^{k-2}f(k+N-s,N)-\alpha(s)\Big]\geq 0,
\end{aligned}
$$

where the last inequality follows from (75).

The non-negativity of (63) follows directly from (74).

Next we show (64)+(65)+(66) is positive. We have

$$
\begin{aligned}
(64)+(65)+(66)=&-[\mu_{11}(\Sigma_2'-\mu_{12}))(s-1)\theta(\Sigma_1'-\mu_{11})\mu_{11}^{N-s}\alpha(s)\\
&+(N-1)\theta\big[[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}](\Sigma_1'-\mu_{11})-\mu_{21}\Sigma_1'\mu_{22}\big]\alpha(N)\\
&+\mu_{22}\mu_{21}(N-1)\theta\Big[\sum_{k=2}^{s}\mu_{11}^{N+k-s-1}f(k+N-s,N)\Big]\\
\overset{(*)}{\geq}&-[\mu_{11}(\Sigma_2'-\mu_{12})](s-1)\theta(\Sigma_1'-\mu_{11})\mu_{11}^{N-s}\alpha(s)\\
&+(s-1)\theta\big[[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}](\Sigma_1'-\mu_{11})-\mu_{21}\Sigma_1'\mu_{22}\big]\mu_{11}^{N-s}\alpha(s)\\
&+\mu_{22}\mu_{21}(N-1)\theta\mu_{11}^{N-s+1}\Big[\sum_{k=2}^{s}\mu_{11}^{k-2}f(k+N-s,N)\Big]\\
\overset{(**)}{\geq}&-[\mu_{11}(\Sigma_2'-\mu_{12})](s-1)\theta(\Sigma_1'-\mu_{11})\mu_{11}^{N-s}\alpha(s)\\
&+(s-1)\theta\big[[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}](\Sigma_1'-\mu_{11})-\mu_{21}\Sigma_1'\mu_{22}\big]\mu_{11}^{N-s}\alpha(s)\\
&+\mu_{22}\mu_{21}(s-1)\theta\mu_{11}^{N-s+1}\alpha(s)\\
=&(s-1)\theta\mu_{11}^{N-s}\alpha(s)\Big[-[\mu_{11}(\Sigma_2'-\mu_{12})](\Sigma_1'-\mu_{11})\\
&+[\mu_{11}(\Sigma_2'-\mu_{12})+\mu_{22}\mu_{21}](\Sigma_1'-\mu_{11})-\mu_{21}\Sigma_1'\mu_{22}+\mu_{22}\mu_{21}\mu_{11}\Big]\\
=&(s-1)\theta\mu_{11}^{N-s}\alpha(s)\Big[\mu_{22}\mu_{21}(\Sigma_1'-\mu_{11})-\mu_{21}\Sigma_1'\mu_{22}+\mu_{22}\mu_{21}\mu_{11}\Big]\\
=&0,
\end{aligned}
$$

where $(*)$ follows from (37) and (76) and $(**)$ follows from (75) and $s \leq N-1$. Hence (64) + (65) + (66) ≥ 0.

The non-negativity of (67) follows from (74).

The non-negativity of (68) follows since $s \leq N-1$ and hence

$$
(N-1)f(N+1-s,N)-(s-1)f(1,s)\geq(s-1)\Big[f(N+1-s,N)-f(1,s)\Big]\geq 0,
$$

where the last inequality follows from $f(N+1-s,N)-f(1,s)\geq 0$ for $s \leq N-1$.

Similarly, the non-negativity of (69) also follows from $f(N+1-s,N)-f(1,s)\geq 0$ for $s \leq N-1$.

Finally, (70) is non-negative follows from (74). Hence we have shown $\Gamma'(s, a_{21}) \geq 0$ and therefore $\Gamma(s, a_{21}) \geq 0$.

Case 2, Part (iii) : $a = a_{11}$.

We have

$$\begin{aligned}
\Gamma(s, a_{11}) =& \mu_{22} + \frac{\mu_{11}}{q} h_N(s+1) + \frac{\mu_{22} + (s-1)\theta}{q} h_N(s-1) + \left(1 - \frac{\mu_{11} + \mu_{22} + (s-1)\theta}{q}\right) h_N(s) \\
& - \frac{\Sigma_1'}{q} h_N(s+1) - \frac{s\theta}{q} h_N(s-1) - \left(1 - \frac{\Sigma_1' + s\theta}{q}\right) h_N(s) \\
=& \mu_{22} + \frac{\mu_{11} - \Sigma_1'}{q} h_N(s+1) + \frac{\mu_{22} - \theta}{q} h_N(s-1) + \frac{\Sigma_1' + s\theta - \mu_{11} - \mu_{22} - (s-1)\theta}{q} h_N(s) \\
=& \mu_{22} - \frac{\Sigma_1' - \mu_{11}}{q} [h_N(s+1) - h_N(s)] - \frac{\mu_{22}}{q} [h_N(s) - h_N(s-1)] + \frac{\theta}{q} [h_N(s) - h_N(s-1)] \\
=& \Gamma''(s, a_{11}) + \frac{\theta}{q} [h_N(s) - h_N(s-1)],
\end{aligned}$$

where

$$\Gamma''(s, a_{11}) = \mu_{22} - \frac{\Sigma_1' - \mu_{11}}{q} [h_N(s+1) - h_N(s)] - \frac{\mu_{22}}{q} [h_N(s) - h_N(s-1)].$$

Since $h_N(s) - h_N(s-1) \geq 0$ from Lemma 6, it is sufficient to show that $\Gamma''(s, a_{11})$ is non negative. Using (56), we have

$$\begin{aligned}
\frac{\Sigma_1' - \mu_{11}}{q} [h_N(s+1) - h_N(s)] =& \frac{(\Sigma_1' - \mu_{11})(g_N - \mu_{22})}{\mu_{11}} \\
& + \frac{(\Sigma_1' - \mu_{11})[\mu_{22} + (s-1)\theta]}{\mu_{11} q} [h_N(s) - h_N(s-1)].
\end{aligned}$$

Hence,

$$\begin{aligned}
& \Gamma''(s, a_{11}) \\
=& \mu_{22} - \frac{(\Sigma_1' - \mu_{11})(g_N - \mu_{22})}{\mu_{11}} - \frac{(\Sigma_1' - \mu_{11})[\mu_{22} + (s-1)\theta]}{\mu_{11} q} [h_N(s) - h_N(s-1)] \\
& - \frac{\mu_{22}}{q} [h_N(s) - h_N(s-1)] \\
=& \frac{\mu_{22}\mu_{11} + (\mu_{11} - \Sigma_1')g_N + \Sigma_1'\mu_{22} - \mu_{22}\mu_{11}}{\mu_{11}} \\
& + \frac{(\mu_{11} - \Sigma_1')\mu_{22} - \mu_{22}\mu_{11} + (\mu_{11} - \Sigma_1')(s-1)\theta}{q\mu_{11}} [h_N(s) - h_N(s-1)] \\
=& \frac{\Sigma_1'\mu_{22} + (\mu_{11} - \Sigma_1')g_N}{\mu_{11}} + \frac{-\Sigma_1'\mu_{22} + (\mu_{11} - \Sigma_1')(s-1)\theta}{q\mu_{11}} [h_N(s) - h_N(s-1)] \\
=& \frac{\Sigma_1'\mu_{22} + (\mu_{11} - \Sigma_1')g_N}{\mu_{11}} + \frac{-\Sigma_1'\mu_{22} + (\mu_{11} - \Sigma_1')(s-1)\theta}{\mu_{11}\Delta(N)} \Bigg[\Sigma_1'\mu_{11}^{N-s} [\Sigma_2' - \mu_{22}]\alpha(s) \\
& + [\Sigma_2' + (N-1)\theta]\mu_{22} \sum_{k=s+1}^{N} \mu_{11}^{k-s-1} f(k, N) f(1, s) + \Sigma_2'\mu_{11}^{N-s} f(1, s) \Bigg],
\end{aligned}$$

where in the last equality we have used Lemma 6. Therefore, (17) yields

$$\begin{aligned}
& \Gamma''(s, a_{11})\Delta(N)\mu_{11} \\
=& \Sigma_1'\mu_{22}\Delta(N) + (\mu_{11} - \Sigma_1')\beta(N) + [-\Sigma_1'\mu_{22} + (\mu_{11} - \Sigma_1')(s-1)\theta] \Big[\Sigma_1'\mu_{11}^{N-s} [\Sigma_2' - \mu_{22}]\alpha(s)
\end{aligned}$$

$$+[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)+\Sigma_2'\mu_{11}^{N-s}f(1,s)\Big].$$

Now (18)-(19) yield

$$\begin{aligned}
&\Gamma''(s,a_{11})\Delta(N)\mu_{11}\\
=&\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta][f(1,N)+\Sigma_1'\alpha(N)]+\Sigma_1'\mu_{22}\Sigma_1'\mu_{11}^{N-1}+(\mu_{11}-\Sigma_1')[\Sigma_2'+(N-1)\theta]\Sigma_1'\mu_{22}\alpha(N)\\
&+(\mu_{11}-\Sigma_1')\Sigma_1'\Sigma_2'\mu_{11}^{N-1}+(-\Sigma_1'\mu_{22}+(\mu_{11}-\Sigma_1')(s-1)\theta)\Sigma_1'\mu_{11}^{N-s}[\Sigma_2'-\mu_{22}]\alpha(s)\\
&+(-\Sigma_1'\mu_{22}+(\mu_{11}-\Sigma_1')(s-1)\theta)[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&+(-\Sigma_1'\mu_{22}+(\mu_{11}-\Sigma_1')(s-1)\theta)\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
=&\Sigma_1'\mu_{22}\Sigma_1'\mu_{11}^{N-1}+(\mu_{11}-\Sigma_1')\Sigma_1'\Sigma_2'\mu_{11}^{N-1}+(-\Sigma_1'\mu_{22}+(\mu_{11}-\Sigma_1')(s-1)\theta)\Sigma_1'\mu_{11}^{N-s}[\Sigma_2'-\mu_{22}]\alpha(s)\\
&+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta][f(1,N)+\Sigma_1'\alpha(N)]+(\mu_{11}-\Sigma_1')[\Sigma_2'+(N-1)\theta]\Sigma_1'\mu_{22}\alpha(N)\\
&+(-\Sigma_1'\mu_{22}+(\mu_{11}-\Sigma_1')(s-1)\theta)[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&+(-\Sigma_1'\mu_{22}+(\mu_{11}-\Sigma_1')(s-1)\theta)\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
=&(\Sigma_1'\mu_{22}+\mu_{11}\Sigma_2'-\Sigma_1'\Sigma_2')\Sigma_1'\mu_{11}^{N-1}+(\mu_{11}-\Sigma_1')[\Sigma_2'-\mu_{22}](s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\\
&-\Sigma_1'\mu_{22}\Sigma_1'\mu_{11}^{N-s}[\Sigma_2'-\mu_{22}]\alpha(s)\\
&+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta][f(1,N)+\Sigma_1'\alpha(N)]+(\mu_{11}-\Sigma_1')[\Sigma_2'+(N-1)\theta]\Sigma_1'\mu_{22}\alpha(N)\\
&+(-\Sigma_1'\mu_{22}+(\mu_{11}-\Sigma_1')(s-1)\theta)[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&+(-\Sigma_1'\mu_{22}+(\mu_{11}-\Sigma_1')(s-1)\theta)\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
=&(\Sigma_1'\mu_{22}+\mu_{11}\Sigma_2'-\Sigma_1'\Sigma_2')\Sigma_1'\mu_{11}^{N-1}+(\Sigma_1'\mu_{22}+\mu_{11}\Sigma_2'-\Sigma_1'\Sigma_2')(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\\
&-\mu_{22}(s-1)\theta\Sigma_1'\mu_{11}^{N-s+1}\alpha(s)-\Sigma_1'\mu_{22}\Sigma_1'\mu_{11}^{N-s}[\Sigma_2'-\mu_{22}]\alpha(s)\\
&+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta][f(1,N)+\Sigma_1'\alpha(N)]+(\mu_{11}-\Sigma_1')[\Sigma_2'+(N-1)\theta]\Sigma_1'\mu_{22}\alpha(N)\\
&+(-\Sigma_1'\mu_{22}+(\mu_{11}-\Sigma_1')(s-1)\theta)[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&+(-\Sigma_1'\mu_{22}+(\mu_{11}-\Sigma_1')(s-1)\theta)\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
=&(\Sigma_1'\mu_{22}+\mu_{11}\Sigma_2'-\Sigma_1'\Sigma_2')\Big(\Sigma_1'\mu_{11}^{N-1}+(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\Big)\\
&-\mu_{22}(s-1)\theta\Sigma_1'\mu_{11}^{N-s+1}\alpha(s)-\Sigma_1'\mu_{22}\Sigma_1'\mu_{11}^{N-s}[\Sigma_2'-\mu_{22}]\alpha(s)\\
&+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta][f(1,N)+\Sigma_1'\alpha(N)]+(\mu_{11}-\Sigma_1')[\Sigma_2'+(N-1)\theta]\Sigma_1'\mu_{22}\alpha(N)\\
&+(-\Sigma_1'\mu_{22}+(\mu_{11}-\Sigma_1')(s-1)\theta)[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&+(-\Sigma_1'\mu_{22}+(\mu_{11}-\Sigma_1')(s-1)\theta)\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
=&(\Sigma_1'\mu_{22}+\mu_{11}\Sigma_2'-\Sigma_1'\Sigma_2')\Big(\Sigma_1'\mu_{11}^{N-1}+(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\Big)
\end{aligned}$$

$$
\begin{aligned}
&- \mu_{22}(s-1)\theta\Sigma_1'\mu_{11}^{N-s+1}\alpha(s) - \Sigma_1'\mu_{22}\Sigma_1'\mu_{11}^{N-s}[\Sigma_2' - \mu_{22}]\alpha(s) \\
&+ (-\Sigma_1'\mu_{22} + (\mu_{11} - \Sigma_1')(s-1)\theta)[\Sigma_2' + (N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s) \\
&+ \Sigma_1'\mu_{22}[\Sigma_2' + (N-1)\theta][f(1,N) + \Sigma_1'\alpha(N)] + (\mu_{11} - \Sigma_1')[\Sigma_2' + (N-1)\theta]\Sigma_1'\mu_{22}\alpha(N) \\
&+ (-\Sigma_1'\mu_{22} + (\mu_{11} - \Sigma_1')(s-1)\theta)\Sigma_2'\mu_{11}^{N-s}f(1,s) \\
=&(\Sigma_1'\mu_{22} + \mu_{11}\Sigma_2' - \Sigma_1'\Sigma_2')\left(\Sigma_1'\mu_{11}^{N-1} + (s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\right) \\
&- \mu_{22}(s-1)\theta\Sigma_1'\mu_{11}^{N-s+1}\alpha(s) - \Sigma_1'\mu_{22}\Sigma_1'\mu_{11}^{N-s}[\Sigma_2' - \mu_{22}]\alpha(s) \\
&+ (-\Sigma_1'\mu_{22} - \Sigma_1'(s-1)\theta)[\Sigma_2' + (N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s) \\
&+ \mu_{11}(s-1)\theta[\Sigma_2' + (N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s) \\
&+ \Sigma_1'\mu_{22}[\Sigma_2' + (N-1)\theta][f(1,N) + \Sigma_1'\alpha(N)] + (\mu_{11} - \Sigma_1')[\Sigma_2' + (N-1)\theta]\Sigma_1'\mu_{22}\alpha(N) \\
&+ (-\Sigma_1'\mu_{22} + (\mu_{11} - \Sigma_1')(s-1)\theta)\Sigma_2'\mu_{11}^{N-s}f(1,s) \\
=&(\Sigma_1'\mu_{22} + \mu_{11}\Sigma_2' - \Sigma_1'\Sigma_2')\left(\Sigma_1'\mu_{11}^{N-1} + (s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\right) \\
&- \mu_{22}(s-1)\theta\Sigma_1'\mu_{11}^{N-s+1}\alpha(s) - \Sigma_1'\mu_{22}\Sigma_1'\mu_{11}^{N-s}[\Sigma_2' - \mu_{22}]\alpha(s) \\
&- \Sigma_1'(\mu_{22} + (s-1)\theta)[\Sigma_2' + (N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s) \\
&+ \mu_{11}(s-1)\theta[\Sigma_2' + (N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s) \\
&+ \Sigma_1'\mu_{22}[\Sigma_2' + (N-1)\theta][f(1,N) + \Sigma_1'\alpha(N)] + (\mu_{11} - \Sigma_1')[\Sigma_2' + (N-1)\theta]\Sigma_1'\mu_{22}\alpha(N) \\
&+ (-\Sigma_1'\mu_{22} + (\mu_{11} - \Sigma_1')(s-1)\theta)\Sigma_2'\mu_{11}^{N-s}f(1,s) \\
=&(\Sigma_1'\mu_{22} + \mu_{11}\Sigma_2' - \Sigma_1'\Sigma_2')\left(\Sigma_1'\mu_{11}^{N-1} + (s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\right) \\
&- \mu_{22}(s-1)\theta\Sigma_1'\mu_{11}^{N-s+1}\alpha(s) - \Sigma_1'\mu_{22}\Sigma_1'\mu_{11}^{N-s}[\Sigma_2' - \mu_{22}]\alpha(s) \\
&- \Sigma_1'[\Sigma_2' + (N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s+1) \\
&+ \mu_{11}(s-1)\theta[\Sigma_2' + (N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s) \\
&+ \Sigma_1'\mu_{22}[\Sigma_2' + (N-1)\theta][f(1,N) + \Sigma_1'\alpha(N)] + (\mu_{11} - \Sigma_1')[\Sigma_2' + (N-1)\theta]\Sigma_1'\mu_{22}\alpha(N) \\
&+ (-\Sigma_1'\mu_{22} + (\mu_{11} - \Sigma_1')(s-1)\theta)\Sigma_2'\mu_{11}^{N-s}f(1,s) \\
=&(\Sigma_1'\mu_{22} + \mu_{11}\Sigma_2' - \Sigma_1'\Sigma_2')\left(\Sigma_1'\mu_{11}^{N-1} + (s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\right) \\
&- \mu_{22}(s-1)\theta\Sigma_1'\mu_{11}^{N-s+1}\alpha(s) - \Sigma_1'\mu_{22}\Sigma_1'\mu_{11}^{N-s}[\Sigma_2' - \mu_{22}]\alpha(s)
\end{aligned}
$$

$$
\begin{aligned}
&- \Sigma_1'[\Sigma_2' + (N-1)\theta]\mu_{22} \sum_{k=s+1}^{N} \mu_{11}^{k-s-1} f(k,N) f(1,s+1) \\
&+ \mu_{11}(s-1)\theta[\Sigma_2' + (N-1)\theta]\mu_{22} \sum_{k=s+1}^{N} \mu_{11}^{k-s-1} f(k,N) f(1,s) \\
&+ \Sigma_1'\mu_{22}[\Sigma_2' + (N-1)\theta][f(1,N)] + \Sigma_1'\mu_{22}[\Sigma_2' + (N-1)\theta][\Sigma_1'\alpha(N) + (\mu_{11} - \Sigma_1')\alpha(N)] \\
&+ (-\Sigma_1'\mu_{22} + (\mu_{11} - \Sigma_1')(s-1)\theta)\Sigma_2'\mu_{11}^{N-s} f(1,s) \\
=&(\Sigma_1'\mu_{22} + \mu_{11}\Sigma_2' - \Sigma_1'\Sigma_2')\Big(\Sigma_1'\mu_{11}^{N-1} + (s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\Big) \\
&- \mu_{22}(s-1)\theta\Sigma_1'\mu_{11}^{N-s+1}\alpha(s) - \Sigma_1'\mu_{22}\Sigma_1'\mu_{11}^{N-s}[\Sigma_2' - \mu_{22}]\alpha(s) \\
&- \Sigma_1'[\Sigma_2' + (N-1)\theta]\mu_{22} \sum_{k=s+1}^{N} \mu_{11}^{k-s-1} f(k,N) f(1,s+1) \\
&+ \mu_{11}(s-1)\theta[\Sigma_2' + (N-1)\theta]\mu_{22} \sum_{k=s+1}^{N} \mu_{11}^{k-s-1} f(k,N) f(1,s) \\
&+ \Sigma_1'\mu_{22}[\Sigma_2' + (N-1)\theta][f(1,N)] + \Sigma_1'\mu_{22}[\Sigma_2' + (N-1)\theta]\mu_{11}\alpha(N) \\
&+ (-\Sigma_1'\mu_{22} + (\mu_{11} - \Sigma_1')(s-1)\theta)\Sigma_2'\mu_{11}^{N-s} f(1,s) \\
=&(\Sigma_1'\mu_{22} + \mu_{11}\Sigma_2' - \Sigma_1'\Sigma_2')\Big(\Sigma_1'\mu_{11}^{N-1} + (s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\Big) \\
&- \mu_{22}(s-1)\theta\Sigma_1'\mu_{11}^{N-s+1}\alpha(s) - \Sigma_1'\mu_{22}\Sigma_1'\mu_{11}^{N-s}[\Sigma_2' - \mu_{22}]\alpha(s) \\
&- \Sigma_1'[\Sigma_2' + (N-1)\theta]\mu_{22} \sum_{k=s+1}^{N} \mu_{11}^{k-s-1} f(k,N) f(1,s+1) \\
&+ \mu_{11}(s-1)\theta[\Sigma_2' + (N-1)\theta]\mu_{22} \sum_{k=s+1}^{N} \mu_{11}^{k-s-1} f(k,N) f(1,s) \\
&+ \Sigma_1'\mu_{22}[\Sigma_2' + (N-1)\theta][f(1,N)] \\
&+ \Sigma_1'\mu_{22}[\Sigma_2' + (N-1)\theta]\mu_{11}\Big[\sum_{k=2}^{N+1-s} \mu_{11}^{k-2} f(k,N) + \sum_{k=N+2-s}^{N} \mu_{11}^{k-2} f(k,N)\Big] \\
&+ (-\Sigma_1'\mu_{22} + (\mu_{11} - \Sigma_1')(s-1)\theta)\Sigma_2'\mu_{11}^{N-s} f(1,s) \\
=&(\Sigma_1'\mu_{22} + \mu_{11}\Sigma_2' - \Sigma_1'\Sigma_2')\Big(\Sigma_1'\mu_{11}^{N-1} + (s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\Big) \\
&- \mu_{22}(s-1)\theta\Sigma_1'\mu_{11}^{N-s+1}\alpha(s) - \Sigma_1'\mu_{22}\Sigma_1'\mu_{11}^{N-s}[\Sigma_2' - \mu_{22}]\alpha(s) \\
&- \Sigma_1'[\Sigma_2' + (N-1)\theta]\mu_{22} \sum_{k=s+1}^{N} \mu_{11}^{k-s-1} f(k,N) f(1,s+1) \\
&+ \mu_{11}(s-1)\theta[\Sigma_2' + (N-1)\theta]\mu_{22} \sum_{k=s+1}^{N} \mu_{11}^{k-s-1} f(k,N) f(1,s)
\end{aligned}
$$

$$+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta][f(1,N)]+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta]\Big[\sum_{k=2}^{N+1-s}\mu_{11}^{k-1}f(k,N)\Big]$$

$$+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta]\sum_{k=N+2-s}^{N}\mu_{11}^{k-1}f(k,N)$$

$$+(-\Sigma_1'\mu_{22}+(\mu_{11}-\Sigma_1')(s-1)\theta)\Sigma_2'\mu_{11}^{N-s}f(1,s)$$

Let $k'+N-s=k$

$$=(\Sigma_1'\mu_{22}+\mu_{11}\Sigma_2'-\Sigma_1'\Sigma_2')\Big(\Sigma_1'\mu_{11}^{N-1}+(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\Big)$$

$$-\mu_{22}(s-1)\theta\Sigma_1'\mu_{11}^{N-s+1}\alpha(s)-\Sigma_1'\mu_{22}\Sigma_1'\mu_{11}^{N-s}[\Sigma_2'-\mu_{22}]\alpha(s)$$

$$-\Sigma_1'[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s+1)$$

$$+\mu_{11}(s-1)\theta[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)$$

$$+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta][f(1,N)]+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta]\Big[\sum_{k=2}^{N+1-s}\mu_{11}^{k-1}f(k,N)\Big]$$

$$+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta]\sum_{k'=2}^{s}\mu_{11}^{k'-1+N-s}f(k'+N-s,N)$$

$$+(-\Sigma_1'\mu_{22}+(\mu_{11}-\Sigma_1')(s-1)\theta)\Sigma_2'\mu_{11}^{N-s}f(1,s)$$

$$=(\Sigma_1'\mu_{22}+\mu_{11}\Sigma_2'-\Sigma_1'\Sigma_2')\Big(\Sigma_1'\mu_{11}^{N-1}+(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\Big)$$

$$-\mu_{22}(s-1)\theta\Sigma_1'\mu_{11}^{N-s+1}\alpha(s)-\Sigma_1'\mu_{22}\Sigma_1'\mu_{11}^{N-s}[\Sigma_2'-\mu_{22}]\alpha(s)$$

$$-\Sigma_1'[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s+1)$$

$$+\mu_{11}(s-1)\theta[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)$$

$$+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta]\Big[\sum_{k=2}^{N+1-s}\mu_{11}^{k-1}f(k,N)+f(1,N)\Big]$$

$$+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta]\sum_{k'=2}^{s}\mu_{11}^{k'-1+N-s}f(k'+N-s,N)$$

$$+(-\Sigma_1'\mu_{22}+(\mu_{11}-\Sigma_1')(s-1)\theta)\Sigma_2'\mu_{11}^{N-s}f(1,s)$$

$$=(\Sigma_1'\mu_{22}+\mu_{11}\Sigma_2'-\Sigma_1'\Sigma_2')\Big(\Sigma_1'\mu_{11}^{N-1}+(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\Big)$$

$$-\mu_{22}(s-1)\theta\Sigma_1'\mu_{11}^{N-s+1}\alpha(s)-\Sigma_1'\mu_{22}\Sigma_1'\mu_{11}^{N-s}[\Sigma_2'-\mu_{22}]\alpha(s)$$

$$-\Sigma_1'[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s+1)$$

$$
\begin{aligned}
&+\mu_{11}(s-1)\theta[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta]\Big[\sum_{k=1}^{N+1-s}\mu_{11}^{k-1}f(k,N)\Big]\\
&+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta]\sum_{k'=2}^{s}\mu_{11}^{k'-1+N-s}f(k'+N-s,N)\\
&+(-\Sigma_1'\mu_{22}+(\mu_{11}-\Sigma_1')(s-1)\theta)\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
=&(\Sigma_1'\mu_{22}+\mu_{11}\Sigma_2'-\Sigma_1'\Sigma_2')\Big(\Sigma_1'\mu_{11}^{N-1}+(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\Big)\\
&-\mu_{22}(s-1)\theta\Sigma_1'\mu_{11}^{N-s+1}\alpha(s)-\Sigma_1'\mu_{22}\Sigma_1'\mu_{11}^{N-s}[\Sigma_2'-\mu_{22}]\alpha(s)\\
&-\Sigma_1'[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s+1)\\
&+\mu_{11}(s-1)\theta[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta]\Big[\sum_{k=1}^{N-s}\mu_{11}^{k-1}f(k,N)+\mu_{11}^{N-s}f(N+1-s,N)\Big]\\
&+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta]\sum_{k'=2}^{s}\mu_{11}^{k'-1+N-s}f(k'+N-s,N)\\
&+(-\Sigma_1'\mu_{22}+(\mu_{11}-\Sigma_1')(s-1)\theta)\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
\text{Let } k''-s=k&\\
=&(\Sigma_1'\mu_{22}+\mu_{11}\Sigma_2'-\Sigma_1'\Sigma_2')\Big(\Sigma_1'\mu_{11}^{N-1}+(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\Big)\\
&-\mu_{22}(s-1)\theta\Sigma_1'\mu_{11}^{N-s+1}\alpha(s)-\Sigma_1'\mu_{22}\Sigma_1'\mu_{11}^{N-s}[\Sigma_2'-\mu_{22}]\alpha(s)\\
&-\Sigma_1'[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s+1)\\
&+\mu_{11}(s-1)\theta[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta]\Big[\sum_{k''=s+1}^{N}\mu_{11}^{k''-s-1}f(k''-s,N)+\mu_{11}^{N-s}f(N+1-s,N)\Big]\\
&+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta]\sum_{k'=2}^{s}\mu_{11}^{k'-1+N-s}f(k'+N-s,N)\\
&+(-\Sigma_1'\mu_{22}+(\mu_{11}-\Sigma_1')(s-1)\theta)\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
=&(\Sigma_1'\mu_{22}+\mu_{11}\Sigma_2'-\Sigma_1'\Sigma_2')\Big(\Sigma_1'\mu_{11}^{N-1}+(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\Big)
\end{aligned}
$$

$$
\begin{aligned}
&+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta]\Big[\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}[f(k-s,N)-f(k,n)f(1,s+1)]\Big]\\
&-\mu_{22}(s-1)\theta\Sigma_1'\mu_{11}^{N-s+1}\alpha(s)-\Sigma_1'\mu_{22}\Sigma_1'\mu_{11}^{N-s}[\Sigma_2'-\mu_{22}]\alpha(s)\\
&+\mu_{11}(s-1)\theta[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta]\Big[\mu_{11}^{N-s}f(N+1-s,N)\Big]\\
&+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta]\sum_{k'=2}^{s}\mu_{11}^{k'-1+N-s}f(k'+N-s,N)\\
&+(-\Sigma_1'\mu_{22}+(\mu_{11}-\Sigma_1')(s-1)\theta)\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
=&(\Sigma_1'\mu_{22}+\mu_{11}\Sigma_2'-\Sigma_1'\Sigma_2')\Big(\Sigma_1'\mu_{11}^{N-1}+(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\Big)\\
&+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta]\Big[\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}[f(k-s,N)-f(k,n)f(1,s+1)]\Big]\\
&+\mu_{11}(s-1)\theta[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&-\mu_{22}(s-1)\theta\Sigma_1'\mu_{11}^{N-s+1}\alpha(s)-\Sigma_1'\mu_{22}\Sigma_1'\mu_{11}^{N-s}[\Sigma_2'-\mu_{22}]\alpha(s)\\
&+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta]\Big[\mu_{11}^{N-s}f(N+1-s,N)\Big]\\
&+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta]\sum_{k'=2}^{s}\mu_{11}^{k'-1+N-s}f(k'+N-s,N)\\
&+(-\Sigma_1'\mu_{22}+(\mu_{11}-\Sigma_1')(s-1)\theta)\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
=&(\Sigma_1'\mu_{22}+\mu_{11}\Sigma_2'-\Sigma_1'\Sigma_2')\Big(\Sigma_1'\mu_{11}^{N-1}+(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\Big)\\
&+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta]\Big[\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}[f(k-s,N)-f(k,n)f(1,s+1)]\Big]\\
&+\mu_{11}(s-1)\theta[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&-\mu_{22}(s-1)\theta\Sigma_1'\mu_{11}^{N-s+1}\alpha(s)-\Sigma_1'\mu_{22}\Sigma_1'\mu_{11}^{N-s}[\Sigma_2'-\mu_{22}]\alpha(s)\\
&+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta]\Big[\mu_{11}^{N-s}f(N+1-s,N)\Big]\\
&+\Sigma_1'\mu_{22}\Sigma_2'\sum_{k'=2}^{s}\mu_{11}^{k'-1+N-s}f(k'+N-s,N)\\
&+\Sigma_1'\mu_{22}[(N-1)\theta]\sum_{k'=2}^{s}\mu_{11}^{k'-1+N-s}f(k'+N-s,N)
\end{aligned}
$$

$$
\begin{aligned}
&+(-\Sigma_1'\mu_{22}+(\mu_{11}-\Sigma_1')(s-1)\theta)\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
=&(\Sigma_1'\mu_{22}+\mu_{11}\Sigma_2'-\Sigma_1'\Sigma_2')\Big(\Sigma_1'\mu_{11}^{N-1}+(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\Big)\\
&+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta]\Big[\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}[f(k-s,N)-f(k,n)f(1,s+1)]\Big]\\
&+\mu_{11}(s-1)\theta[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&+\Sigma_1'\mu_{22}\mu_{11}^{N-s}\Big[\Sigma_2'\sum_{k'=2}^{s}\mu_{11}^{k'-1}f(k'+N-s,N)-\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(s)\Big]\\
&-\mu_{22}(s-1)\theta\Sigma_1'\mu_{11}^{N-s+1}\alpha(s)\\
&+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta]\Big[\mu_{11}^{N-s}f(N+1-s,N)\Big]\\
&+\Sigma_1'\mu_{22}[(N-1)\theta]\sum_{k'=2}^{s}\mu_{11}^{k'-1+N-s}f(k'+N-s,N)\\
&+(-\Sigma_1'\mu_{22}+(\mu_{11}-\Sigma_1')(s-1)\theta)\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
=&(\Sigma_1'\mu_{22}+\mu_{11}\Sigma_2'-\Sigma_1'\Sigma_2')\Big(\Sigma_1'\mu_{11}^{N-1}+(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\Big)\\
&+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta]\Big[\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}[f(k-s,N)-f(k,n)f(1,s+1)]\Big]\\
&+\mu_{11}(s-1)\theta[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)\\
&+\Sigma_1'\mu_{22}\mu_{11}^{N-s}\Big[\Sigma_2'\sum_{k'=2}^{s}\mu_{11}^{k'-1}f(k'+N-s,N)-\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(s)\Big]\\
&+\Sigma_1'\mu_{22}\mu_{11}^{N-s+1}\Big[(N-1)\theta\sum_{k'=2}^{s}\mu_{11}^{k'-2}f(k'+N-s,N)-(s-1)\theta\alpha(s)\Big]\\
&+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta]\Big[\mu_{11}^{N-s}f(N+1-s,N)\Big]\\
&+(-\Sigma_1'\mu_{22}+(\mu_{11}-\Sigma_1')(s-1)\theta)\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
=&(\Sigma_1'\mu_{22}+\mu_{11}\Sigma_2'-\Sigma_1'\Sigma_2')\Big(\Sigma_1'\mu_{11}^{N-1}+(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\Big)\\
&+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta]\Big[\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}[f(k-s,N)-f(k,n)f(1,s+1)]\Big]\\
&+\mu_{11}(s-1)\theta[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s)
\end{aligned}
$$

$$
\begin{aligned}
&+\Sigma_1'\mu_{22}\mu_{11}^{N-s}\left[\Sigma_2'\sum_{k'=2}^{s}\mu_{11}^{k'-1}f(k'+N-s,N)-\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(s)\right]\\
&+\Sigma_1'\mu_{22}\mu_{11}^{N-s+1}\left[(N-1)\theta\sum_{k'=2}^{s}\mu_{11}^{k'-2}f(k'+N-s,N)-(s-1)\theta\alpha(s)\right]\\
&+\Sigma_1'\mu_{22}\Sigma_2'\left[\mu_{11}^{N-s}f(N+1-s,N)\right]+\Sigma_1'\mu_{22}[(N-1)\theta]\left[\mu_{11}^{N-s}f(N+1-s,N)\right]\\
&-\Sigma_1'\mu_{22}\Sigma_2'\mu_{11}^{N-s}f(1,s)+(\mu_{11}-\Sigma_1')(s-1)\theta\Sigma_2'\mu_{11}^{N-s}f(1,s)\\
=&[\Sigma_1'\mu_{22}+\mu_{11}\Sigma_2'-\Sigma_1'\Sigma_2']\left(\Sigma_1'\mu_{11}^{N-1}+(s-1)\theta\Sigma_1'\mu_{11}^{N-s}\alpha(s)\right) &(77)\\
&+\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta]\left[\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}[f(k-s,N)-f(k,N)f(1,s+1)]\right] &(78)\\
&+\mu_{11}(s-1)\theta[\Sigma_2'+(N-1)\theta]\mu_{22}\sum_{k=s+1}^{N}\mu_{11}^{k-s-1}f(k,N)f(1,s) &(79)\\
&+\Sigma_1'\mu_{22}\mu_{11}^{N-s}\left[\Sigma_2'\sum_{k=2}^{s}\mu_{11}^{k-1}f(k+N-s,N)-\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(s)\right] &(80)\\
&+\Sigma_1'\mu_{22}\mu_{11}^{N-s+1}\left[(N-1)\theta\sum_{k=2}^{s}\mu_{11}^{k-2}f(k+N-s,N)-(s-1)\theta\alpha(s)\right] &(81)\\
&+\Sigma_1'\mu_{22}\Sigma_2'\mu_{11}^{N-s}\left[f(N+1-s,N)-f(1,s)\right] &(82)\\
&+\mu_{11}^{N-s}\left[\Sigma_1'\mu_{22}(N-1)\theta f(N+1-s,N)+(\mu_{11}-\Sigma_1')(s-1)\theta\Sigma_2'f(1,s)\right]. &(83)
\end{aligned}
$$

We show that all the terms (77)-(83) are non-negative.

The non-negativity of (77) follows from

$$
\begin{aligned}
&\Sigma_1'\mu_{22}+\mu_{11}\Sigma_2'-\Sigma_1'\Sigma_2'\\
=&(1+\lambda)(\mu_{11}+\mu_{21})\mu_{22}+(1+\lambda)(\mu_{12}+\mu_{22})\mu_{11}-(1+\lambda)^2\Sigma_1\Sigma_2\\
=&(1+\lambda)\left[\mu_{11}\mu_{22}+\mu_{21}\mu_{22}+\mu_{12}\mu_{11}+\mu_{22}\mu_{11}-(\mu_{11}+\mu_{21})(\mu_{22}+\mu_{12})-\lambda\Sigma_1\Sigma_2\right]\\
=&(1+\lambda)\left[\mu_{11}\mu_{22}+\mu_{21}\mu_{22}+\mu_{12}\mu_{11}+\mu_{22}\mu_{11}-\mu_{11}\mu_{22}-\mu_{11}\mu_{12}-\mu_{21}\mu_{22}-\mu_{21}\mu_{12}-\lambda\Sigma_1\Sigma_2\right]\\
=&(1+\lambda)\left[\mu_{11}\mu_{22}-\mu_{21}\mu_{12}-\lambda\Sigma_1\Sigma_2\right]\geq 0. &(84)
\end{aligned}
$$

from (73) since $\tau'(2)\geq 0$ for this case to be possible.

The non-negativity of (78) follows since $s+1\leq k\leq N$ and

$$f(k-s,N)-f(1,s+1)f(k,N)=f(k-s,k)f(k,N)-f(1,s+1)f(k,N)\geq 0$$

as $f(k-s,k)\geq f(1,s+1)$ for $k\geq s+1$.

Clearly, (79) is non-negative.

For (80) we have

$$\Sigma_2'\sum_{k=2}^{s}\mu_{11}^{k-1}f(k+N-s,N)-\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(s)\geq\Sigma_2'\mu_{11}\sum_{k=2}^{s}\mu_{11}^{k-2}f(k,s)-\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(s)$$

$$=\Sigma_2'\mu_{11}\alpha(s) - \Sigma_1'[\Sigma_2' - \mu_{22}]\alpha(s)$$
$$=[\Sigma_1'\mu_{22} + \mu_{11}\Sigma_2' - \Sigma_1'\Sigma_2']\alpha(s) \geq 0,$$

where the last inequality follows from (84). Therefore (80) is non-negative.

For (81), $s \leq N-1$ yields

$$(N-1)\theta \sum_{k=2}^{s} \mu_{11}^{k-2} f(k+N-s, N) - (s-1)\theta\alpha(s)$$
$$\geq (s-1)\theta\left[\sum_{k=2}^{s} \mu_{11}^{k-2} f(k+N-s, N) - \alpha(s)\right] \geq 0,$$

where the last inequality follows from (75).

The non-negativity of (82) follows as $f(N+1-s, N) \geq f(1, s)$ for $s = 1, \ldots, N-1$.

Finally, for (83), $s \leq N-1$ yields

$$\Sigma_1'\mu_{22}(N-1)\theta f(N+1-s, N) + (\mu_{11} - \Sigma_1')(s-1)\theta\Sigma_2' f(1, s)$$
$$\geq \Sigma_1'\mu_{22}(s-1)\theta f(1, s) + (\mu_{11} - \Sigma_1')(s-1)\theta\Sigma_2' f(1, s)$$
$$= (s-1)\theta f(1, s)\left[\Sigma_1'\mu_{22} + (\mu_{11} - \Sigma_1')\Sigma_2'\right]$$
$$= (s-1)\theta f(1, s)\left[\Sigma_1'\mu_{22} + \mu_{11}\Sigma_2' - \Sigma_1'\Sigma_2'\right] \geq 0,$$

where the last inequality again follows from (84). Hence we have shown $\Gamma''(s, a_{11}) \geq 0$ and therefore $\Gamma(s, a_{11}) \geq 0$.

Case 3: $s = N, \ldots, B+2$.

In this case we have $d_N(s) = a_{22}$.

Case 3, Part (i) : $a = a_{12}$.

We consider the two sub-cases (a) $s = N \ldots, B+1$ and (b) $s = B+2$ separately.

(a) $s = N \ldots, B+1$. We have

$$\Gamma(s, a_{12}) = \Sigma_2' + \frac{\Sigma_2' + (s-1)\theta}{q} h_N(s-1) + \left(1 - \frac{\Sigma_2' + (s-1)\theta}{q}\right) h_N(s)$$
$$- \mu_{22} - \frac{\mu_{11}}{q} h_N(s+1) - \frac{\mu_{22} + (s-1)\theta}{q} h_N(s-1) - \left(1 - \frac{\mu_{11} + \mu_{22} + (s-1)\theta}{q}\right) h_N(s).$$

That is

$$\Gamma(s, a_{12}) = \Sigma_2' - \mu_{22} - \frac{\mu_{11}}{q} h_N(s+1) + \frac{\Sigma_2' - \mu_{22}}{q} h_N(s-1) + \frac{\mu_{11}}{q} h_N(s) - \frac{\Sigma_2' - \mu_{22}}{q} h_N(s)$$
$$= \Sigma_2' - \mu_{22} + \frac{\Sigma_2' - \mu_{22}}{q}[h_N(s-1) - h_N(s)] + \frac{\mu_{11}}{q}[h_N(s) - h_N(s+1)]. \tag{85}$$

It follows from Lemma 5 that

$$\Gamma(s, a_{12}) = \Sigma_2' - \mu_{22} + \frac{\Sigma_2' - \mu_{22}}{\Sigma_2' + (s-1)\theta}[g_N - \Sigma_2'] + \frac{\mu_{11}}{\Sigma_2' + s\theta}[g_N - \Sigma_2']. \tag{86}$$

In particular for $s = N$,

$$\Gamma(N, a_{12}) = \Sigma_2' - \mu_{22} + \frac{\Sigma_2' - \mu_{22}}{\Sigma_2' + (N-1)\theta}[g_N - \Sigma_2'] + \frac{\mu_{11}}{\Sigma_2' + N\theta}[g_N - \Sigma_2'].$$

Using Proposition 2, we have

$$g_N - \Sigma_2' = \frac{-[\Sigma_2' + (N-1)\theta]\left(\Sigma_2' f(1, N) + \Sigma_1'[\Sigma_2' - \mu_{22}]\alpha(N)\right)}{\Delta(N)}. \tag{87}$$

Hence

$$\begin{aligned}\Gamma(N,a_{12}) =&\Sigma_2' - \mu_{22} - \frac{\Sigma_2' - \mu_{22}}{\Delta(N)}\Big(\Sigma_2' f(1,N) + \Sigma_1'[\Sigma_2' - \mu_{22}]\alpha(N)\Big)\\ &- \frac{\mu_{11}}{\Sigma_2' + N\theta}\left[\frac{[\Sigma_2' + (N-1)\theta]\Big(\Sigma_2' f(1,N) + \Sigma_1'[\Sigma_2' - \mu_{22}]\alpha(N)\Big)}{\Delta(N)}\right].\end{aligned}$$

That is

$$\begin{aligned}&[\Sigma_2' + N\theta]\Delta(N)\Gamma(N,a_{12})\\ =&[\Sigma_2' + N\theta]\left[(\Sigma_2' - \mu_{22})\Delta(N) - (\Sigma_2' - \mu_{22})\Big(\Sigma_2' f(1,N) + \Sigma_1'[\Sigma_2' - \mu_{22}]\alpha(N)\Big)\right]\\ &- \mu_{11}[\Sigma_2' + (N-1)\theta]\Big(\Sigma_2' f(1,N) + \Sigma_1'[\Sigma_2' - \mu_{22}]\alpha(N)\Big).\end{aligned}$$

Recalling (19), we have

$$\begin{aligned}&[\Sigma_2' + N\theta]\Delta(N)\Gamma(N,a_{12})\\ =&[\Sigma_2' + N\theta]\Bigg[(\Sigma_2' - \mu_{22})[\Sigma_2' + (N-1)\theta][f(1,N) + \Sigma_1'\alpha(N)] + (\Sigma_2' - \mu_{22})\Sigma_1'\mu_{11}^{N-1}\\ &- (\Sigma_2' - \mu_{22})\Big(\Sigma_2' f(1,N) + \Sigma_1'[\Sigma_2' - \mu_{22}]\alpha(N)\Big)\Bigg]\\ &- \mu_{11}[\Sigma_2' + (N-1)\theta]\Big(\Sigma_2' f(1,N) + \Sigma_1'[\Sigma_2' - \mu_{22}]\alpha(N)\Big)\\ =&[\Sigma_2' + N\theta]\Bigg[(\Sigma_2' - \mu_{22})[\Sigma_2' + (N-1)\theta]f(1,N) + (\Sigma_2' - \mu_{22})[(\Sigma_2' - \mu_{22}) + \mu_{22} + (N-1)\theta]\Sigma_1'\alpha(N)\\ &+ (\Sigma_2' - \mu_{22})\Sigma_1'\mu_{11}^{N-1} - (\Sigma_2' - \mu_{22})\Big(\Sigma_2' f(1,N) + \Sigma_1'[\Sigma_2' - \mu_{22}]\alpha(N)\Big)\Bigg]\\ &- \mu_{11}[\Sigma_2' + (N-1)\theta]\Big(\Sigma_2' f(1,N) + \Sigma_1'[\Sigma_2' - \mu_{22}]\alpha(N)\Big)\\ =&[\Sigma_2' + N\theta]\Bigg[(\Sigma_2' - \mu_{22})[\Sigma_2' + (N-1)\theta]f(1,N) + (\Sigma_2' - \mu_{22})^2\Sigma_1'\alpha(N)\\ &+ (\Sigma_2' - \mu_{22})\Sigma_1'\Big((\mu_{22} + (N-1)\theta)\alpha(N) + \mu_{11}^{N-1}\Big) - (\Sigma_2' - \mu_{22})\Sigma_2' f(1,N) - \Sigma_1'(\Sigma_2' - \mu_{22})^2\alpha(N)\Bigg]\\ &- \mu_{11}[\Sigma_2' + (N-1)\theta]\Big(\Sigma_2' f(1,N) + \Sigma_1'[\Sigma_2' - \mu_{22}]\alpha(N)\Big)\\ \overset{(*)}{=}&[\Sigma_2' + N\theta]\Big[(\Sigma_2' - \mu_{22})[\Sigma_2' + (N-1)\theta]f(1,N) + (\Sigma_2' - \mu_{22})\Sigma_1'\alpha(N+1) - (\Sigma_2' - \mu_{22})\Sigma_2' f(1,N)\Big]\\ &- \mu_{11}[\Sigma_2' + (N-1)\theta]\Big(\Sigma_2' f(1,N) + \Sigma_1'[\Sigma_2' - \mu_{22}]\alpha(N)\Big)\\ =&[\Sigma_2' + N\theta]\Big[- \mu_{22}[\Sigma_2' + (N-1)\theta]f(1,N) + \Sigma_2'[\Sigma_2' + (N-1)\theta]f(1,N) - (\Sigma_2' - \mu_{22})\Sigma_2' f(1,N)\end{aligned}$$

$$
\begin{aligned}
&+(\Sigma_2'-\mu_{22})\Sigma_1'\alpha(N+1)\Big]-\mu_{11}[\Sigma_2'+(N-1)\theta]\Big(\Sigma_2'f(1,N)+\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(N)\Big)\\
=&[\Sigma_2'+N\theta]\Big[-\mu_{22}[\Sigma_2'+(N-1)\theta]f(1,N)+\Sigma_2'[\mu_{22}+(N-1)\theta]f(1,N)+(\Sigma_2'-\mu_{22})\Sigma_1'\alpha(N+1)\Big]\\
&-\mu_{11}[\Sigma_2'+(N-1)\theta]\Big(\Sigma_2'f(1,N)+\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(N)\Big)\\
=&[\Sigma_2'+N\theta]\Big[-\mu_{22}[\Sigma_2'+(N-1)\theta]f(1,N)+\Sigma_2'f(1,N+1)+(\Sigma_2'-\mu_{22})\Sigma_1'\alpha(N+1)\Big]\\
&-\mu_{11}[\Sigma_2'+(N-1)\theta]\Big(\Sigma_2'f(1,N)+\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(N)\Big)\\
=&-[\Sigma_2'+N\theta][\Sigma_2'+(N-1)\theta]\mu_{22}f(1,N)+[\Sigma_2'+N\theta][\Sigma_1'\alpha(N+1)(\Sigma_2'-\mu_{22})+\Sigma_2'f(1,N+1)]\\
&-[\Sigma_2'+(N-1)\theta]\mu_{11}[\Sigma_2'f(1,N)+\Sigma_1'\alpha(N)(\Sigma_2'-\mu_{22})]\\
=&-\tau'(N+1),
\end{aligned}
$$

where $(*)$ follows from (23) and the last equality follows from (25). That is we have shown that

$$[\Sigma_2'+N\theta]\Delta(N)\Gamma(N,a_{12})=-\tau'(N+1)\geq 0,$$

where the last inequality follows from the definition of N.

Now for $s=N,\ldots,B$, (86) yields

$$
\begin{aligned}
&\Gamma(s+1,a_{12})-\Gamma(s,a_{12})\\
=&\Sigma_2'-\mu_{22}+\frac{\Sigma_2'-\mu_{22}}{\Sigma_2'+s\theta}[g_N-\Sigma_2']+\frac{\mu_{11}}{\Sigma_2'+(s+1)\theta}[g_N-\Sigma_2']\\
&-(\Sigma_2'-\mu_{22})-\frac{\Sigma_2'-\mu_{22}}{\Sigma_2'+(s-1)\theta}[g_N-\Sigma_2']-\frac{\mu_{11}}{\Sigma_2'+s\theta}[g_N-\Sigma_2']\\
=&(g_N-\Sigma_2')\left[(\Sigma_2'-\mu_{22})\left(\frac{1}{\Sigma_2'+s\theta}-\frac{1}{\Sigma_2'+(s-1)\theta}\right)+\mu_{11}\left(\frac{1}{\Sigma_2'+(s+1)\theta}-\frac{1}{\Sigma_2'+s\theta}\right)\right]\geq 0,
\end{aligned}
$$

where the last inequality follows as $(g_N-\Sigma_2')<0$ from (87). This shows $\Gamma(s,a_{12})\geq 0$ for for $s=N.\ldots,B+1$.

(b) $s=B+2$. We have

$$
\begin{aligned}
\Gamma(B+2,a_{12})=&\Sigma_2'+\frac{\Sigma_2'+(B+1)\theta}{q}h_N(B+1)+\left(1-\frac{\Sigma_2'+(B+1)\theta}{q}\right)h_N(B+2)\\
&-\mu_{22}-\frac{\mu_{22}+(B+1)\theta}{q}h_N(B+1)-\left(1-\frac{\mu_{22}+(B+1)\theta}{q}\right)h_N(B+2)\\
=&\Sigma_2'-\mu_{22}+\frac{\Sigma_2'-\mu_{22}}{q}[h_N(B+1)-h_N(B+2)]\\
=&\Sigma_2'-\mu_{22}+\frac{\Sigma_2'-\mu_{22}}{\Sigma_2'+(B+1)\theta}(g_N-\Sigma_2'),
\end{aligned}
$$

where the last equality follows from Lemma 5. Moreover, from (86) we have

$$\Gamma(B+1,a_{12})=\Sigma_2'-\mu_{22}+\frac{\Sigma_2'-\mu_{22}}{\Sigma_2'+B\theta}[g_N-\Sigma_2']+\frac{\mu_{11}}{\Sigma_2'+(B+1)\theta}[g_N-\Sigma_2'].$$

Therefore,

$$\begin{aligned}
&\Gamma(B+2,a_{12})-\Gamma(B+1,a_{12})\\
&=(g_N-\Sigma_2')\left[\frac{\Sigma_2'-\mu_{22}}{\Sigma_2'+(B+1)\theta}-\frac{\mu_{11}}{\Sigma_2'+(B+1)\theta}-\frac{\Sigma_2'-\mu_{22}}{\Sigma_2'+B\theta}\right]\\
&=(g_N-\Sigma_2')\left[\frac{\Sigma_2'-\mu_{22}-\mu_{11}}{\Sigma_2'+(B+1)\theta}-\frac{\Sigma_2'-\mu_{22}}{\Sigma_2'+B\theta}\right]\geq 0.
\end{aligned}$$

The inequality follows as $\Sigma_2'-\mu_{22}-\mu_{11}\leq\Sigma_2'-\mu_{22}$, $\Sigma_2'+(B+1)\theta\geq\Sigma_2'+B\theta$ and $g_N-\Sigma_2'\leq 0$ from (87). Therefore $\Gamma(s,a_{12})\geq 0$ for for $s=N.\ldots,B+2$.

Case 3, Part (ii) : For $a=a_{21}$.

Again, we consider the two sub-cases (a) $s=N.\ldots,B+1$ and (b) $s=B+2$ separately. (a) $s=N,\ldots,B+1$. Using an argument similar to that in (86) (replace μ_{11} with μ_{21} and μ_{22} with μ_{12}) we obtain

$$\Gamma(s,a_{21})=\Sigma_2'-\mu_{12}+\frac{\Sigma_2'-\mu_{12}}{\Sigma_2'+(s-1)\theta}(g_N-\Sigma_2')+\frac{\mu_{21}}{\Sigma_2'+s\theta}(g_N-\Sigma_2').$$

Letting $s=N$ and using (87) we obtain

$$\begin{aligned}
\Gamma(N,a_{21})=&\Sigma_2'-\mu_{12}-\frac{\Sigma_2'-\mu_{12}}{\Delta(N)}\Big(\Sigma_2'f(1,N)+\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(N)\Big)\\
&-\frac{\mu_{21}}{\Sigma_2'+N\theta}\frac{[\Sigma_2'+(N-1)\theta]\Big(\Sigma_2'f(1,N)+\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(N)\Big)}{\Delta(N)}.
\end{aligned}$$

That is

$$\begin{aligned}
&\Gamma(N,a_{21})\Delta(N)(\Sigma_2'+N\theta)\\
=&(\Sigma_2'+N\theta)\Delta(N)(\Sigma_2'-\mu_{12})-(\Sigma_2'+N\theta)\Big(\Sigma_2'f(1,N)+\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(N)\Big)(\Sigma_2'-\mu_{12})\\
&-\mu_{21}[\Sigma_2'+(N-1)\theta]\Big(\Sigma_2'f(1,N)+\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(N)\Big).
\end{aligned}$$

We consider the three terms in the above separately. The first term satisfies

$$\begin{aligned}
&(\Sigma_2'+N\theta)\Delta(N)(\Sigma_2'-\mu_{12})\\
=&(\Sigma_2'-\mu_{12})\Big[[\Sigma_2'+(N-1)\theta][f(1,N)+\Sigma_1'\alpha(N)]+\Sigma_1'\mu_{11}^{N-1}\Big](\Sigma_2'+N\theta)\\
=&(\Sigma_2'-\mu_{12})\Sigma_2'f(1,N)(\Sigma_2'+N\theta)+\Sigma_1'(\Sigma_2'-\mu_{12})\Sigma_2'\alpha(N)(\Sigma_2'+N\theta)+\Sigma_1'\mu_{11}^{N-1}(\Sigma_2'-\mu_{12})(\Sigma_2'+N\theta)\\
&+(\Sigma_2'+N\theta)(\Sigma_2'-\mu_{12})(N-1)\theta f(1,N)+\Sigma_1'(\Sigma_2'-\mu_{12})(N-1)\theta\alpha(N)(\Sigma_2'+N\theta)\\
=&(\Sigma_2'-\mu_{12})\Sigma_2'f(1,N)(\Sigma_2'+N\theta)+\Sigma_1'(\Sigma_2'-\mu_{12})\Sigma_2'\alpha(N)(\Sigma_2'+N\theta)+\Sigma_1'\mu_{11}^{N-1}(\Sigma_2'-\mu_{12})(\Sigma_2'+N\theta)\\
&+(\Sigma_2'+N\theta)(\Sigma_2'-\mu_{12})(N-1)\theta f(1,N)+\Sigma_1'(\Sigma_2'-\mu_{12})(N-2)\theta\alpha(N)(\Sigma_2'+N\theta)\\
&+\Sigma_1'(\Sigma_2'-\mu_{12})\theta\alpha(N)(\Sigma_2'+N\theta).
\end{aligned}\tag{88}$$

The second term satisfies

$$\begin{aligned}
&-(\Sigma_2'+N\theta)\Big(\Sigma_2'f(1,N)+\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(N)\Big)(\Sigma_2'-\mu_{12})\\
=&(\Sigma_2'+N\theta)\Big(-\Sigma_2'f(1,N)+\Sigma_1'[\mu_{22}-\Sigma_2']\alpha(N)\Big)(\Sigma_2'-\mu_{12})
\end{aligned}$$

$$
\begin{aligned}
&= -(\Sigma_2'-\mu_{12})\Sigma_2' f(1,N)(\Sigma_2'+N\theta)+\Sigma_1'\mu_{22}(\Sigma_2'-\mu_{12})\alpha(N)(\Sigma_2'+N\theta)-\Sigma_1'\Sigma_2'(\Sigma_2'-\mu_{12})\alpha(N)(\Sigma_2'+N\theta)\\
&= -(\Sigma_2'-\mu_{12})\Sigma_2' f(1,N)(\Sigma_2'+N\theta)-\Sigma_1'\Sigma_2'(\Sigma_2'-\mu_{12})\alpha(N)(\Sigma_2'+N\theta)\\
&\quad+\Sigma_1'\mu_{22}(\Sigma_2'-\mu_{12})(\Sigma_2'+N\theta)\Big[\sum_{k=3}^{N}\mu_{11}^{k-2}f(k,N)+f(2,N)\Big]\\
&= -(\Sigma_2'-\mu_{12})\Sigma_2' f(1,N)(\Sigma_2'+N\theta)-\Sigma_1'\Sigma_2'(\Sigma_2'-\mu_{12})\alpha(N)(\Sigma_2'+N\theta)\\
&\quad+\Sigma_1'\mu_{22}(\Sigma_2'-\mu_{12})(\Sigma_2'+N\theta)\sum_{k=3}^{N}\mu_{11}^{k-2}f(k,N)+\Sigma_1'(\Sigma_2'-\mu_{12})(\Sigma_2'+N\theta)f(1,N)\\
&= -(\Sigma_2'-\mu_{12})\Sigma_2' f(1,N)(\Sigma_2'+N\theta)-\Sigma_1'\Sigma_2'(\Sigma_2'-\mu_{12})\alpha(N)(\Sigma_2'+N\theta)\\
&\quad+\Sigma_1'\mu_{22}(\Sigma_2'-\mu_{12})(\Sigma_2'+N\theta)\sum_{k=3}^{N}\mu_{11}^{k-2}f(k,N)+\Sigma_1'(\Sigma_2'-\mu_{12})(\Sigma_2'+(N-1)\theta)f(1,N)\\
&\quad+\Sigma_1'(\Sigma_2'-\mu_{12})\theta f(1,N). \qquad (89)
\end{aligned}
$$

And the third term satisfies

$$
\begin{aligned}
&-\mu_{21}[\Sigma_2'+(N-1)\theta]\Big(\Sigma_2' f(1,N)+\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(N)\Big)\\
&= -\mu_{21}\Sigma_2'[\Sigma_2'+(N-1)\theta]f(1,N)-\mu_{21}\Sigma_1'[\Sigma_2'-\mu_{22}][\Sigma_2'+(N-1)\theta]\alpha(N)\\
&= -\mu_{21}\Sigma_2'[\Sigma_2'+(N-1)\theta]f(1,N)+\Big(\Sigma_1'(\Sigma_2'-\mu_{12})\mu_{11}-\mu_{21}\Sigma_1'[\Sigma_2'-\mu_{22}]\Big)[\Sigma_2'+(N-1)\theta]\alpha(N)\\
&\quad-\Sigma_1'(\Sigma_2'-\mu_{12})\mu_{11}[\Sigma_2'+(N-1)\theta]\alpha(N). \qquad (90)
\end{aligned}
$$

Adding (88)-(90) yields

$$
\begin{aligned}
&\Gamma(N,a_{21})\Delta(N)(\Sigma_2'+N\theta)\\
=&(\Sigma_2'-\mu_{12})\Sigma_2' f(1,N)(\Sigma_2'+N\theta)+\Sigma_1'(\Sigma_2'-\mu_{12})\Sigma_2'\alpha(N)(\Sigma_2'+N\theta)+\Sigma_1'\mu_{11}^{N-1}(\Sigma_2'-\mu_{12})(\Sigma_2'+N\theta)\\
&+(\Sigma_2'+N\theta)(\Sigma_2'-\mu_{12})(N-1)\theta f(1,N)+\Sigma_1'(\Sigma_2'-\mu_{12})(N-2)\theta\alpha(N)(\Sigma_2'+N\theta)\\
&+\Sigma_1'(\Sigma_2'-\mu_{12})\theta\alpha(N)(\Sigma_2'+N\theta)\\
&-(\Sigma_2'-\mu_{12})\Sigma_2' f(1,N)(\Sigma_2'+N\theta)-\Sigma_1'\Sigma_2'(\Sigma_2'-\mu_{12})\alpha(N)(\Sigma_2'+N\theta)\\
&+\Sigma_1'\mu_{22}(\Sigma_2'-\mu_{12})(\Sigma_2'+N\theta)\sum_{k=3}^{N}\mu_{11}^{k-2}f(k,N)+\Sigma_1'(\Sigma_2'-\mu_{12})(\Sigma_2'+(N-1)\theta)f(1,N)\\
&+\Sigma_1'(\Sigma_2'-\mu_{12})\theta f(1,N)\\
&-\mu_{21}\Sigma_2'[\Sigma_2'+(N-1)\theta]f(1,N)+\Big(\Sigma_1'(\Sigma_2'-\mu_{12})\mu_{11}-\mu_{21}\Sigma_1'[\Sigma_2'-\mu_{22}]\Big)[\Sigma_2'+(N-1)\theta]\alpha(N)\\
&-\Sigma_1'(\Sigma_2'-\mu_{12})\mu_{11}[\Sigma_2'+(N-1)\theta]\alpha(N)\\
=&\Sigma_1'\mu_{11}^{N-1}(\Sigma_2'-\mu_{12})(\Sigma_2'+N\theta)\\
&+(\Sigma_2'+N\theta)(\Sigma_2'-\mu_{12})(N-1)\theta f(1,N)+\Sigma_1'(\Sigma_2'-\mu_{12})(N-2)\theta\alpha(N)(\Sigma_2'+N\theta)\\
&+\Sigma_1'(\Sigma_2'-\mu_{12})\theta\alpha(N)(\Sigma_2'+N\theta)\\
&+\Sigma_1'\mu_{22}(\Sigma_2'-\mu_{12})(\Sigma_2'+N\theta)\sum_{k=3}^{N}\mu_{11}^{k-2}f(k,N)+\Sigma_1'(\Sigma_2'-\mu_{12})(\Sigma_2'+(N-1)\theta)f(1,N)\\
&+\Sigma_1'(\Sigma_2'-\mu_{12})\theta f(1,N)
\end{aligned}
$$

$$
\begin{aligned}
&- \mu_{21}\Sigma_2'[\Sigma_2' + (N-1)\theta]f(1,N) + \Big(\Sigma_1'(\Sigma_2' - \mu_{12})\mu_{11} - \mu_{21}\Sigma_1'[\Sigma_2' - \mu_{22}]\Big)[\Sigma_2' + (N-1)\theta]\alpha(N) \\
&- \Sigma_1'(\Sigma_2' - \mu_{12})\mu_{11}[\Sigma_2' + (N-1)\theta]\alpha(N) \\
=&(\Sigma_2' + N\theta)(\Sigma_2' - \mu_{12})(N-1)\theta f(1,N) + \Sigma_1'(\Sigma_2' - \mu_{12})\theta f(1,N) + \Sigma_1'(\Sigma_2' - \mu_{12})(\Sigma_2' + (N-1)\theta)f(1,N) \\
&- \mu_{21}\Sigma_2'[\Sigma_2' + (N-1)\theta]f(1,N) + \Big(\Sigma_1'(\Sigma_2' - \mu_{12})\mu_{11} - \mu_{21}\Sigma_1'[\Sigma_2' - \mu_{22}]\Big)[\Sigma_2' + (N-1)\theta]\alpha(N) \\
&+ \Sigma_1'(\Sigma_2' - \mu_{12})(N-2)\theta\alpha(N)(\Sigma_2' + N\theta) + \Sigma_1'\mu_{11}^{N-1}(\Sigma_2' - \mu_{12})(\Sigma_2' + N\theta) \\
&+ \Sigma_1'(\Sigma_2' - \mu_{12})\theta\alpha(N)(\Sigma_2' + N\theta) + \Sigma_1'\mu_{22}(\Sigma_2' - \mu_{12})(\Sigma_2' + N\theta)\sum_{k=3}^{N}\mu_{11}^{k-2}f(k,N) \\
&- \Sigma_1'(\Sigma_2' - \mu_{12})\mu_{11}[\Sigma_2' + (N-1)\theta]\alpha(N) \\
=&(\Sigma_2' + N\theta)(\Sigma_2' - \mu_{12})(N-1)\theta f(1,N) + \Sigma_1'(\Sigma_2' - \mu_{12})\theta f(1,N) && (91) \\
&+ \Big[\Sigma_1'(\Sigma_2' - \mu_{12}) - \mu_{21}\Sigma_2'\Big][\Sigma_2' + (N-1)\theta]f(1,N) && (92) \\
&+ \Sigma_1'\Big((\Sigma_2' - \mu_{12})\mu_{11} - \mu_{21}[\Sigma_2' - \mu_{22}]\Big)[\Sigma_2' + (N-1)\theta]\alpha(N) && (93) \\
&+ \Sigma_1'(\Sigma_2' - \mu_{12})\Big[(N-2)\theta\alpha(N)(\Sigma_2' + N\theta) + \mu_{11}^{N-1}(\Sigma_2' + N\theta) + \theta\alpha(N)(\Sigma_2' + N\theta)\Big] && (94) \\
&+ \Sigma_1'(\Sigma_2' - \mu_{12})\Big[\mu_{22}(\Sigma_2' + N\theta)\sum_{k=3}^{N}\mu_{11}^{k-2}f(k,N) - \mu_{11}[\Sigma_2' + (N-1)\theta]\alpha(N)\Big]. && (95)
\end{aligned}
$$

We show that each of the term (91)-(95) is non-negative. (91) is clearly positive.
For (92) we have

$$
\begin{aligned}
&\Sigma_1'(\Sigma_2' - \mu_{12}) - \mu_{21}\Sigma_2' \\
=&(1+\lambda)(\mu_{11} + \mu_{21})(\lambda\mu_{12} + (1+\lambda)\mu_{22}) - \mu_{21}(1+\lambda)(\mu_{22} + \mu_{12}) \\
=&(1+\lambda)\Big[\lambda\mu_{11}\mu_{12} + \lambda\mu_{21}\mu_{12} + (1+\lambda)\mu_{11}\mu_{22} + (1+\lambda)\mu_{21}\mu_{22} - \mu_{21}\mu_{22} - \mu_{21}\mu_{12}\Big] \\
=&(1+\lambda)\Big[\lambda\Sigma_1\Sigma_2 + (\mu_{11}\mu_{22} - \mu_{21}\mu_{12})\Big] \\
\geq&0,
\end{aligned}
$$

where the inequality follows as $\mu_{11}\mu_{22} - \mu_{21}\mu_{12} \geq 0$.

For (93), note that $\alpha(1) = 0$ and hence (93) equals zero when $N = 1$. For $N \geq 2$, we have

$$
\begin{aligned}
(\Sigma_2' - \mu_{12})\mu_{11} - \mu_{21}[\Sigma_2' - \mu_{22}] =&(\mu_{22} + \lambda\Sigma_2)\mu_{11} - \mu_{21}(\mu_{12} + \lambda\Sigma_2) \\
=&(\mu_{11}\mu_{22} - \mu_{21}\mu_{12}) + \lambda\mu_{11}\Sigma_2 - \lambda\mu_{21}\Sigma_2 \\
\overset{(*)}{\geq}&(\mu_{11}\mu_{22} - \mu_{21}\mu_{12}) + \lambda\mu_{11}\Sigma_2 - \frac{\mu_{21}}{\Sigma_1}(\mu_{11}\mu_{22} - \mu_{21}\mu_{12}) \\
=&(\mu_{11}\mu_{22} - \mu_{21}\mu_{12})\Big[1 - \frac{\mu_{21}}{\Sigma_1}\Big] + \lambda\mu_{11}\Sigma_2 \\
=&(\mu_{11}\mu_{22} - \mu_{21}\mu_{12})\frac{\mu_{11}}{\Sigma_1} + \lambda\mu_{11}\Sigma_2 \\
\overset{(**)}{\geq}&0,
\end{aligned}
$$

where in $(**)$ we have used $\mu_{11}\mu_{22}-\mu_{21}\mu_{12}\geq 0$ and in $(*)$ we have used (73) since we have $\tau'(2)\geq 0$ when $N\geq 2$. It follows that (93) is non-negative for all possible values of N.

Finally, we show that (94)+(95) is non-negative. We have

$$
\begin{aligned}
&(94)+(95)\\
=&\Sigma_1'(\Sigma_2'-\mu_{12})\Big[(N-2)\theta\alpha(N)(\Sigma_2'+N\theta)+\mu_{11}^{N-1}(\Sigma_2'+N\theta)+\theta\alpha(N)(\Sigma_2'+N\theta)\\
&+\mu_{22}(\Sigma_2'+N\theta)\sum_{k=3}^{N}\mu_{11}^{k-2}f(k,N)-\mu_{11}[\Sigma_2'+(N-1)\theta]\alpha(N)\Big]\\
=&\Sigma_1'(\Sigma_2'-\mu_{12})\Big[(N-2)\theta\alpha(N)(\Sigma_2'+N\theta)+\mu_{11}^{N-1}(\Sigma_2'+N\theta)+\theta\alpha(N)(\Sigma_2'+N\theta)\\
&+\mu_{22}(\Sigma_2'+N\theta)\sum_{k=3}^{N}\mu_{11}^{k-2}f(k,N)-\mu_{11}[\Sigma_2'+N\theta]\alpha(N)+\mu_{11}\theta\alpha(N)\Big]\\
=&\Sigma_1'(\Sigma_2'-\mu_{12})\Big[(N-2)\theta\alpha(N)(\Sigma_2'+N\theta)+(\Sigma_2'+N\theta)\Big(\mu_{11}^{N-1}+\theta\alpha(N)+\mu_{22}\sum_{k=3}^{N}\mu_{11}^{k-2}f(k,N)\\
&-\mu_{11}\alpha(N)\Big)+\mu_{11}\theta\alpha(N)\Big]\\
=&\Sigma_1'(\Sigma_2'-\mu_{12})\Big[(N-2)\theta\alpha(N)(\Sigma_2'+N\theta)+(\Sigma_2'+N\theta)\Big(\mu_{11}^{N-1}-\sum_{k=2}^{N}\mu_{11}^{k-1}f(k,N)+\theta\alpha(N)\\
&+\mu_{22}\sum_{k=3}^{N}\mu_{11}^{k-2}f(k,N)\Big)+\mu_{11}\theta\alpha(N)\Big]\\
=&\Sigma_1'(\Sigma_2'-\mu_{12})\Big[(N-2)\theta\alpha(N)(\Sigma_2'+N\theta)+(\Sigma_2'+N\theta)\Big(-\sum_{k=2}^{N-1}\mu_{11}^{k-1}f(k,N)+\theta\alpha(N)\\
&+\mu_{22}\sum_{k=3}^{N}\mu_{11}^{k-2}f(k,N)\Big)+\mu_{11}\theta\alpha(N)\Big].
\end{aligned}
$$

Let $k'=k-1$. Then

$$
\begin{aligned}
&(94)+(95)\\
=&\Sigma_1'(\Sigma_2'-\mu_{12})\Big[(N-2)\theta\alpha(N)(\Sigma_2'+N\theta)\\
&+(\Sigma_2'+N\theta)\Big(-\sum_{k=2}^{N-1}\mu_{11}^{k-1}f(k,N)+\theta\alpha(N)+\mu_{22}\sum_{k'=2}^{N-1}\mu_{11}^{k'-1}f(k'+1,N)\Big)+\mu_{11}\theta\alpha(N)\Big]\\
=&\Sigma_1'(\Sigma_2'-\mu_{12})\Big[(N-2)\theta\alpha(N)(\Sigma_2'+N\theta)\\
&+(\Sigma_2'+N\theta)\Big(-\sum_{k=2}^{N-1}\mu_{11}^{k-1}[\mu_{22}+(k-1)\theta]f(k+1,N)+\theta\alpha(N)+\mu_{22}\sum_{k'=2}^{N-1}\mu_{11}^{k'-1}f(k'+1,N)\Big)\\
&+\mu_{11}\theta\alpha(N)\Big]
\end{aligned}
$$

$$
\begin{aligned}
=&\Sigma_1'(\Sigma_2'-\mu_{12})\Big[(N-2)\theta\alpha(N)(\Sigma_2'+N\theta)+(\Sigma_2'+N\theta)\Big(-\sum_{k=2}^{N-1}\mu_{11}^{k-1}(k-1)\theta f(k+1,N)+\theta\alpha(N)\Big)\\
&+\mu_{11}\theta\alpha(N)\Big]\\
=&\Sigma_1'(\Sigma_2'-\mu_{12})\Big[(\Sigma_2'+N\theta)\Big((N-2)\theta\alpha(N)-\sum_{k=2}^{N-1}\mu_{11}^{k-1}(k-1)\theta f(k+1,N)+\theta\alpha(N)\Big)+\mu_{11}\theta\alpha(N)\Big]\\
\geq&\Sigma_1'(\Sigma_2'-\mu_{12})\Big[(\Sigma_2'+N\theta)\Big((N-1)\theta\alpha(N)-(N-1)\theta\sum_{k=2}^{N-1}\mu_{11}^{k-1}f(k+1,N)\Big)+\mu_{11}\theta\alpha(N)\Big]\\
=&\Sigma_1'(\Sigma_2'-\mu_{12})\Big[(\Sigma_2'+N\theta)(N-1)\theta\Big(\sum_{k=2}^{N}\mu_{11}^{k-2}f(k,N)-\sum_{k=2}^{N-1}\mu_{11}^{k-1}f(k+1,N)\Big)+\mu_{11}\theta\alpha(N)\Big].
\end{aligned}
$$

Let $k''=k+1$, We have

$$
\begin{aligned}
(94)+(95)=&\Sigma_1'(\Sigma_2'-\mu_{12})\Big[(\Sigma_2'+N\theta)(N-1)\theta\Big(\sum_{k=2}^{N}\mu_{11}^{k-2}f(k,N)-\sum_{k''=3}^{N}\mu_{11}^{k''-2}f(k'',N)\Big)+\mu_{11}\theta\alpha(N)\Big]\\
=&\Sigma_1'(\Sigma_2'-\mu_{12})\Big[(\Sigma_2'+N\theta)(N-1)\theta f(2,N)+\mu_{11}\theta\alpha(N)\Big]\\
\geq&0.
\end{aligned}
$$

This proves $\Gamma(N,a_{21})\geq 0$.

Now for $s=N,\ldots,B$,

$$
\begin{aligned}
&\Gamma(s+1,a_{21})-\Gamma(s,a_{21})\\
=&(g_N-\Sigma_2')\Big[\mu_{21}\Big(\frac{1}{\Sigma_2'+(s+1)\theta}-\frac{1}{\Sigma_2'+s\theta}\Big)+(\Sigma_2'-\mu_{12})\Big(\frac{1}{\Sigma_2'+s\theta}-\frac{1}{\Sigma_2'+(s-1)\theta}\Big)\Big]>0
\end{aligned}
$$

as $(g_N-\Sigma_2')\leq 0$ from (87) . This proves $\Gamma(s,a_{21})\geq 0$ for $s=N,\ldots,B+1$.

(b) For $s=B+2$, we have

$$
\begin{aligned}
\Gamma(B+2,a_{21})=&\Sigma_2'+\frac{\Sigma_2'+(B+1)\theta}{q}h_N(B+1)+\Big(1-\frac{\Sigma_2'+(B+1)\theta}{q}\Big)h_N(B+2)\\
&-\mu_{12}-\frac{\mu_{12}+(B+1)\theta}{q}h_N(B+1)-\Big(1-\frac{\mu_{12}+(B+1)\theta}{q}\Big)h_N(B+2)\\
=&\Sigma_2'-\mu_{12}+\frac{\Sigma_2'-\mu_{12}}{q}[h_N(B+1)-h_N(B+2)]\\
=&\Sigma_2'-\mu_{12}+\frac{\Sigma_2'-\mu_{12}}{\Sigma_2'+(B+1)\theta}(g_N-\Sigma_2')\\
=&\frac{(\Sigma_2'-\mu_{12})[(B+1)\theta+g_N]}{\Sigma_2'+(B+1)\theta}>0,
\end{aligned}
$$

where the third equality follows from Lemma 5. Thus $\Gamma(s,a_{21})\geq 0$ for all $s=N,\ldots,B+2$.

Case 3, Part (iii) : $a=a_{11}$.

We consider the two sub-cases (a) $s=B+2$ and (b) $s=N.\ldots,B+1$ separately.

(a) $s=B+2$. We have from Lemma 5

$$
\Gamma(B+2,a_{11})=\Sigma_2'+\frac{\Sigma_2'+(B+1)\theta}{q}h_N(B+1)+\Big(1-\frac{\Sigma_2'+(B+1)\theta}{q}\Big)h_N(B+2)
$$

$$
\begin{aligned}
&- \frac{(B+2)\theta}{q} h_N(B+1) - \left(1 - \frac{(B+2)\theta}{q}\right) h_N(B+2) \\
=&\Sigma_2' + \frac{\Sigma_2'}{q}[h_N(B+1) - h_N(B+2)] + \frac{\theta}{q}[h_N(B+2) - h_N(B+1)] \\
=&\Sigma_2' + \frac{\Sigma_2'}{\Sigma_2' + (B+1)\theta}(g_N - \Sigma_2') + \frac{\theta}{q}[h_N(B+2) - h_N(B+1)] \\
=&\frac{\Sigma_2'(B+1)\theta + \Sigma_2' g_N}{\Sigma_2' + (B+1)\theta} + \frac{\theta}{q}[h_N(B+2) - h_N(B+1)] \geq 0, \qquad (96)
\end{aligned}
$$

where the third equality follows from Lemma 5 and the last inequality follows from Lemma 6.

(*b*) $s = N, \ldots, B+1$. We have

$$
\begin{aligned}
\Gamma(s, a_{11}) =&\Sigma_2' + \frac{\Sigma_2' + (s-1)\theta}{q} h_N(s-1) + \left(1 - \frac{\Sigma_2' + (s-1)\theta}{q}\right) h_N(s) \\
&- \frac{\Sigma_1'}{q} h_N(s+1) - \frac{s\theta}{q} h_N(s-1) - \left(1 - \frac{\Sigma_1' + s\theta}{q}\right) h_N(s) \\
=&\Sigma_2' + \frac{\Sigma_2' - \theta}{q} h_N(s-1) - \frac{\Sigma_1'}{q} h_N(s+1) - \frac{\Sigma_2' - \theta}{q} h_N(s) + \frac{\Sigma_1'}{q} h_N(s) \\
=&\Sigma_2' + \frac{\Sigma_2'}{q}[h_N(s-1) - h_N(s)] + \frac{\Sigma_1'}{q}[h_N(s) - h_N(s+1)] + \frac{\theta}{q}[h_N(s) - h_N(s-1)]. \qquad (97)
\end{aligned}
$$

In order to show that $\Gamma(s, a_{11})$ is non-negative, we consider two scenarios; (1)$N = 1$ and (2)$N \geq 2$ separately.

Scenario 1: $N = 1$.

We have for $s = N, \ldots, B+1$, using Lemma 5

$$
\begin{aligned}
\Gamma(s, a_{11}) =&\Sigma_2' + \frac{\Sigma_2'}{\Sigma_2' + (s-1)\theta}(g_N - \Sigma_2') + \frac{\Sigma_1'}{\Sigma_2' + s\theta}(g_N - \Sigma_2') + \frac{\theta}{q}[h_N(s) - h_N(s-1)] \\
=&\Gamma_1(s, a_{11}) + \frac{\theta}{q}[h_N(s) - h_N(s-1)],
\end{aligned}
$$

where

$$
\Gamma_1(s, a_{11}) = \Sigma_2' + \frac{\Sigma_2'}{\Sigma_2' + (s-1)\theta}(g_N - \Sigma_2') + \frac{\Sigma_1'}{\Sigma_2' + s\theta}(g_N - \Sigma_2').
$$

Since $h_N(s) - h_N(s-1) \geq 0$ from Lemma 6, it is sufficient to show $\Gamma_1(s, a_{11}) \geq 0$ for $s = N, \ldots, B+1$. Now

$$
g_N - \Sigma_2' = g_1 - \Sigma_2' = \frac{\beta(1)}{\Delta(1)} - \Sigma_2' = \frac{\Sigma_1'\Sigma_2'}{\Sigma_2' + \Sigma_1'} - \Sigma_2' = \frac{-(\Sigma_2')^2}{\Sigma_1' + \Sigma_2'}.
$$

Hence

$$
\begin{aligned}
\Gamma_1(N, a_{11}) = \Gamma_1(1, a_{11}) =&\Sigma_2' - \frac{(\Sigma_2')^2}{\Sigma_1' + \Sigma_2'} - \frac{\Sigma_1'(\Sigma_2')^2}{(\Sigma_2' + \theta)(\Sigma_1' + \Sigma_2')} \\
=&\frac{\Sigma_1'\Sigma_2'}{\Sigma_1' + \Sigma_2'} - \frac{\Sigma_1'(\Sigma_2')^2}{(\Sigma_2' + \theta)(\Sigma_1' + \Sigma_2')} \\
=&\frac{\Sigma_1'\Sigma_2'\theta}{(\Sigma_2' + \theta)(\Sigma_1' + \Sigma_2')} \geq 0.
\end{aligned}
$$

Furthermore for $s = N, \ldots, B$, we have

$$\begin{aligned}
&\Gamma_1(s+1, a_{11}) - \Gamma_1(s, a_{11}) \\
=&(g_N - \Sigma_2')\left[\Sigma_1'\left(\frac{1}{\Sigma_2' + (s+1)\theta} - \frac{1}{\Sigma_2' + s\theta}\right) + \Sigma_2'\left(\frac{1}{\Sigma_2' + s\theta} - \frac{1}{\Sigma_2' + (s-1)\theta}\right)\right] \geq 0,
\end{aligned}$$

where the last inequality follows as $g_N - \Sigma_2 \leq 0$ from (87). Therefore $\Gamma_1(s, a_{11}) \geq 0$ for $s = N, \ldots, B+1$ and hence $\Gamma(s, a_{11}) \geq 0$ for $s = N, \ldots, B+1$ (if $N = 1$).

Scenario 2: $N \geq 2$.

We have from (97) for $s = N, \ldots, B+1$

$$\begin{aligned}
\Gamma(s, a_{11}) =&\Sigma_2' + \frac{\Sigma_2'}{q}[h_N(s-1) - h_N(s)] + \frac{\Sigma_1'}{q}[h_N(s) - h_N(s+1)] + \frac{\theta}{q}[h_N(s) - h_N(s-1)] \\
=&\Sigma_2' - \mu_{22} + \frac{\Sigma_2' - \mu_{22}}{q}[h_N(s-1) - h_N(s)] + \frac{\mu_{11}}{q}[h_N(s) - h_N(s+1)] \\
&+ \mu_{22} + \frac{\mu_{22}}{q}[h_N(s-1) - h_N(s)] + \frac{\Sigma_1' - \mu_{11}}{q}[h_N(s) - h_N(s+1)] \\
&+ \frac{\theta}{q}[h_N(s) - h_N(s-1)] \\
=&\Gamma(s, a_{12}) + \Gamma_2(s, a_{11}) + \frac{\theta}{q}[h_N(s) - h_N(s-1)],
\end{aligned}$$

where $\Gamma_2(s, a_{11}) = \mu_{22} + \frac{\mu_{22}}{q}[h_N(s-1) - h_N(s)] + \frac{\Sigma_1' - \mu_{11}}{q}[h_N(s) - h_N(s+1)]$ and $\Gamma(s, a_{12})$ as defined in (85). Since we have shown earlier $\Gamma(s, a_{12}) \geq 0$ and we have $h_N(s) - h_N(s-1) \geq 0$ from Lemma 6, it is sufficient to show $\Gamma_2(s, a_{11}) \geq 0$ for $s = N, \ldots, B+1$. Using Lemma 5,

$$\Gamma_2(s, a_{11}) = \mu_{22} + \frac{\mu_{22}}{\Sigma_2' + (s-1)\theta}(g_N - \Sigma_2') + \frac{\Sigma_1' - \mu_{11}}{\Sigma_2' + s\theta}(g_N - \Sigma_2').$$

From (87) we obtain

$$\begin{aligned}
\Gamma_2(N, a_{11}) =&\mu_{22} - \frac{\mu_{22}}{\Delta(N)}\Big(\Sigma_2' f(1, N) + \Sigma_1'[\Sigma_2' - \mu_{22}]\alpha(N)\Big) \\
&- \frac{\Sigma_1' - \mu_{11}}{\Sigma_2' + N\theta} \times \frac{[\Sigma_2' + (N-1)\theta]\Big(\Sigma_2' f(1, N) + \Sigma_1'[\Sigma_2' - \mu_{22}]\alpha(N)\Big)}{\Delta(N)},
\end{aligned}$$

and hence

$$\begin{aligned}
&\Gamma_2(N, a_{11})\Delta(N)(\Sigma_2' + N\theta) \\
&= (\Sigma_2' + N\theta)\Delta(N)\mu_{22} - (\Sigma_2' + N\theta)\Big(\Sigma_2' f(1, N) + \Sigma_1'[\Sigma_2' - \mu_{22}]\alpha(N)\Big)\mu_{22} \\
&- (\Sigma_1' - \mu_{11})[\Sigma_2' + (N-1)\theta]\Big(\Sigma_2' f(1, N) + \Sigma_1'[\Sigma_2' - \mu_{22}]\alpha(N)\Big).
\end{aligned}$$

We consider the three terms in the above separately. The first term satisfies

$$\begin{aligned}
&(\Sigma_2' + N\theta)\Delta(N)\mu_{22} \\
=&\mu_{22}\Big[[\Sigma_2' + (N-1)\theta][f(1, N) + \Sigma_1'\alpha(N)] + \Sigma_1'\mu_{11}^{N-1}\Big](\Sigma_2' + N\theta) \\
=&\mu_{22}\Sigma_2' f(1, N)(\Sigma_2' + N\theta) + \Sigma_1'\mu_{22}\Sigma_2'\alpha(N)(\Sigma_2' + N\theta) + \Sigma_1'\mu_{11}^{N-1}\mu_{22}(\Sigma_2' + N\theta) \\
&+ (\Sigma_2' + N\theta)\mu_{22}(N-1)\theta f(1, N) + \Sigma_1'\mu_{22}(N-1)\theta\alpha(N)(\Sigma_2' + N\theta)
\end{aligned}$$

$$
\begin{aligned}
=&\mu_{22}\Sigma_2' f(1,N)(\Sigma_2'+N\theta)+\Sigma_1'\mu_{22}\Sigma_2'\alpha(N)(\Sigma_2'+N\theta)+\Sigma_1'\mu_{11}^{N-1}\mu_{22}(\Sigma_2'+N\theta)\\
&+(\Sigma_2'+N\theta)\mu_{22}(N-1)\theta f(1,N)+\Sigma_1'\mu_{22}(N-2)\theta\alpha(N)(\Sigma_2'+N\theta)+\Sigma_1'\mu_{22}\theta\alpha(N)(\Sigma_2'+N\theta).
\end{aligned}
\tag{98}
$$

The second term satisfies

$$
\begin{aligned}
&-(\Sigma_2'+N\theta)\Big(\Sigma_2' f(1,N)+\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(N)\Big)\mu_{22}\\
=&(\Sigma_2'+N\theta)\Big(-\Sigma_2' f(1,N)+\Sigma_1'[\mu_{22}-\Sigma_2']\alpha(N)\Big)\mu_{22}\\
=&-\mu_{22}\Sigma_2' f(1,N)(\Sigma_2'+N\theta)+\Sigma_1'\mu_{22}^2\alpha(N)(\Sigma_2'+N\theta)-\Sigma_1'\Sigma_2'\mu_{22}\alpha(N)(\Sigma_2'+N\theta)\\
=&-\mu_{22}\Sigma_2' f(1,N)(\Sigma_2'+N\theta)-\Sigma_1'\Sigma_2'\mu_{22}\alpha(N)(\Sigma_2'+N\theta)+\Sigma_1'\mu_{22}^2(\Sigma_2'+N\theta)\sum_{k=2}^{N}\mu_{11}^{k-2}f(k,N)\\
=&-\mu_{22}\Sigma_2' f(1,N)(\Sigma_2'+N\theta)-\Sigma_1'\Sigma_2'\mu_{22}\alpha(N)(\Sigma_2'+N\theta)+\Sigma_1'\mu_{22}^2(\Sigma_2'+N\theta)\sum_{k=3}^{N}\mu_{11}^{k-2}f(k,N)\\
&+\Sigma_1'\mu_{22}(\Sigma_2'+N\theta)f(1,N)\\
=&-\mu_{22}\Sigma_2' f(1,N)(\Sigma_2'+N\theta)-\Sigma_1'\Sigma_2'\mu_{22}\alpha(N)(\Sigma_2'+N\theta)+\Sigma_1'\mu_{22}^2(\Sigma_2'+N\theta)\sum_{k=3}^{N}\mu_{11}^{k-2}f(k,N)\\
&+\Sigma_1'\mu_{22}(\Sigma_2'+(N-1)\theta)f(1,N)+\Sigma_1'\mu_{22}\theta f(1,N).
\end{aligned}
\tag{99}
$$

And the third term satisfies

$$
\begin{aligned}
&-(\Sigma_1'-\mu_{11})[\Sigma_2'+(N-1)\theta]\Big(\Sigma_2' f(1,N)+\Sigma_1'[\Sigma_2'-\mu_{22}]\alpha(N)\Big)\\
=&(\Sigma_1'-\mu_{11})[\Sigma_2'+(N-1)\theta]\Big(-\Sigma_2' f(1,N)+\Sigma_1'[\mu_{22}-\Sigma_2']\alpha(N)\Big)\\
=&(\mu_{11}\Sigma_2'-\Sigma_1'\Sigma_2')[\Sigma_2'+(N-1)\theta]f(1,N)+[\Sigma_2'+(N-1)\theta](\Sigma_1'-\mu_{11})(\Sigma_1'\mu_{22}-\Sigma_1'\Sigma_2')\alpha(N)\\
=&(\mu_{11}\Sigma_2'-\Sigma_1'\Sigma_2')[\Sigma_2'+(N-1)\theta]f(1,N)+[\Sigma_2'+(N-1)\theta](\Sigma_1'\mu_{22}-\Sigma_1'\Sigma_2')\Sigma_1'\alpha(N)\\
&-\mu_{11}[\Sigma_2'+(N-1)\theta](\Sigma_1'\mu_{22}-\Sigma_1'\Sigma_2')\alpha(N)\\
=&(\mu_{11}\Sigma_2'-\Sigma_1'\Sigma_2')[\Sigma_2'+(N-1)\theta]f(1,N)+[\Sigma_2'+(N-1)\theta](\Sigma_1'\mu_{22}+\mu_{11}\Sigma_2'-\Sigma_1'\Sigma_2')\Sigma_1'\alpha(N)\\
&-\mu_{11}[\Sigma_2'+(N-1)\theta](\Sigma_1'\mu_{22}-\Sigma_1'\Sigma_2')\alpha(N)-\mu_{11}\Sigma_1'\Sigma_2'[\Sigma_2'+(N-1)\theta]\alpha(N)\\
=&(\mu_{11}\Sigma_2'-\Sigma_1'\Sigma_2')[\Sigma_2'+(N-1)\theta]f(1,N)+[\Sigma_2'+(N-1)\theta](\Sigma_1'\mu_{22}+\mu_{11}\Sigma_2'-\Sigma_1'\Sigma_2')\Sigma_1'\alpha(N)\\
&-\mu_{11}\Sigma_1'\mu_{22}[\Sigma_2'+(N-1)\theta]\alpha(N).
\end{aligned}
\tag{100}
$$

Adding (98)-(100) yields

$$
\begin{aligned}
&\Gamma_2(N,a_{11})\Delta(N)(\Sigma_2'+N\theta)\\
=&\mu_{22}\Sigma_2' f(1,N)(\Sigma_2'+N\theta)+\Sigma_1'\mu_{22}\Sigma_2'\alpha(N)(\Sigma_2'+N\theta)+\Sigma_1'\mu_{11}^{N-1}\mu_{22}(\Sigma_2'+N\theta)\\
&+(\Sigma_2'+N\theta)\mu_{22}(N-1)\theta f(1,N)+\Sigma_1'\mu_{22}(N-2)\theta\alpha(N)(\Sigma_2'+N\theta)+\Sigma_1'\mu_{22}\theta\alpha(N)(\Sigma_2'+N\theta)\\
&-\mu_{22}\Sigma_2' f(1,N)(\Sigma_2'+N\theta)-\Sigma_1'\Sigma_2'\mu_{22}\alpha(N)(\Sigma_2'+N\theta)+\Sigma_1'\mu_{22}^2(\Sigma_2'+N\theta)\sum_{k=3}^{N}\mu_{11}^{k-2}f(k,N)\\
&+\Sigma_1'\mu_{22}(\Sigma_2'+(N-1)\theta)f(1,N)+\Sigma_1'\mu_{22}\theta f(1,N)\\
&+(\mu_{11}\Sigma_2'-\Sigma_1'\Sigma_2')[\Sigma_2'+(N-1)\theta]f(1,N)+[\Sigma_2'+(N-1)\theta](\Sigma_1'\mu_{22}+\mu_{11}\Sigma_2'-\Sigma_1'\Sigma_2')\Sigma_1'\alpha(N)
\end{aligned}
$$

$$
\begin{aligned}
&- \mu_{11}\Sigma_1'\mu_{22}[\Sigma_2' + (N-1)\theta]\alpha(N) \\
=&\Sigma_1'\mu_{11}^{N-1}\mu_{22}(\Sigma_2' + N\theta) \\
&+ (\Sigma_2' + N\theta)\mu_{22}(N-1)\theta f(1,N) + \Sigma_1'\mu_{22}(N-2)\theta\alpha(N)(\Sigma_2' + N\theta) + \Sigma_1'\mu_{22}\theta\alpha(N)(\Sigma_2' + N\theta) \\
&+ \Sigma_1'\mu_{22}^2(\Sigma_2' + N\theta)\sum_{k=3}^{N}\mu_{11}^{k-2}f(k,N) \\
&+ \Sigma_1'\mu_{22}(\Sigma_2' + (N-1)\theta)f(1,N) + \Sigma_1'\mu_{22}\theta f(1,N) \\
&+ (\mu_{11}\Sigma_2' - \Sigma_1'\Sigma_2')[\Sigma_2' + (N-1)\theta]f(1,N) + [\Sigma_2' + (N-1)\theta](\Sigma_1'\mu_{22} + \mu_{11}\Sigma_2' - \Sigma_1'\Sigma_2')\Sigma_1'\alpha(N) \\
&- \mu_{11}\Sigma_1'\mu_{22}[\Sigma_2' + (N-1)\theta]\alpha(N) \\
=&(\Sigma_2' + N\theta)\mu_{22}(N-1)\theta f(1,N) + \Sigma_1'\mu_{22}\theta f(1,N) && (101) \\
&+ (\Sigma_1'\mu_{22} + \mu_{11}\Sigma_2' - \Sigma_1'\Sigma_2')(\Sigma_2' + (N-1)\theta)\Big(f(1,N) + \Sigma_1'\alpha(N)\Big) && (102) \\
&+ \Sigma_1'\mu_{22}(N-2)\theta\alpha(N)(\Sigma_2' + N\theta) + \Sigma_1'\mu_{11}^{N-1}\mu_{22}(\Sigma_2' + N\theta) + \Sigma_1'\mu_{22}\theta\alpha(N)(\Sigma_2' + N\theta) && (103) \\
&+ \Sigma_1'\mu_{22}^2(\Sigma_2' + N\theta)\sum_{k=3}^{N}\mu_{11}^{k-2}f(k,N) - \mu_{11}\Sigma_1'\mu_{22}[\Sigma_2' + (N-1)\theta]\alpha(N). && (104)
\end{aligned}
$$

We show all the terms (101)-(104) is non-negative. (101) is clearly non-negative.

We have (102) is non negative as $\Sigma_1'\mu_{22} + \mu_{11}\Sigma_2' - \Sigma_1'\Sigma_2' \geq 0$ from (84) since $N \geq 2$ and hence $\tau'(2) \geq 0$.

Finally we show (103)+(104) is non-negative.

$$
\begin{aligned}
&(103) + (104) \\
=&\Sigma_1'\mu_{22}(N-2)\theta\alpha(N)(\Sigma_2' + N\theta) + \Sigma_1'\mu_{11}^{N-1}\mu_{22}(\Sigma_2' + N\theta) + \Sigma_1'\mu_{22}\theta\alpha(N)(\Sigma_2' + N\theta) \\
&+ \Sigma_1'\mu_{22}^2(\Sigma_2' + N\theta)\sum_{k=3}^{N}\mu_{11}^{k-2}f(k,N) - \mu_{11}\Sigma_1'\mu_{22}[\Sigma_2' + N\theta]\alpha(N) + \mu_{11}\Sigma_1'\mu_{22}\theta\alpha(N) \\
=&\Sigma_1'\mu_{22}(N-2)\theta\alpha(N)(\Sigma_2' + N\theta) \\
&+ \Sigma_1'\mu_{22}(\Sigma_2' + N\theta)\Big[\mu_{11}^{N-1} + \theta\alpha(N) + \mu_{22}\sum_{k=3}^{N}\mu_{11}^{k-2}f(k,N) - \mu_{11}\alpha(N)\Big] + \mu_{11}\Sigma_1'\mu_{22}\theta\alpha(N) \\
=&\Sigma_1'\mu_{22}(N-2)\theta\alpha(N)(\Sigma_2' + N\theta) \\
&+ \Sigma_1'\mu_{22}(\Sigma_2' + N\theta)\Big[\mu_{11}^{N-1} - \sum_{k=2}^{N}\mu_{11}^{k-1}f(k,N) + \theta\alpha(N) + \mu_{22}\sum_{k=3}^{N}\mu_{11}^{k-2}f(k,N)\Big] + \mu_{11}\Sigma_1'\mu_{22}\theta\alpha(N) \\
=&\Sigma_1'\mu_{22}(N-2)\theta\alpha(N)(\Sigma_2' + N\theta) \\
&+ \Sigma_1'\mu_{22}(\Sigma_2' + N\theta)\Big[-\sum_{k=2}^{N-1}\mu_{11}^{k-1}f(k,N) + \theta\alpha(N) + \mu_{22}\sum_{k=3}^{N}\mu_{11}^{k-2}f(k,N)\Big] + \mu_{11}\Sigma_1'\mu_{22}\theta\alpha(N).
\end{aligned}
$$

Let $k' = k - 1$. Then

$$
\begin{aligned}
&(103) + (104) \\
=&\Sigma_1'\mu_{22}(N-2)\theta\alpha(N)(\Sigma_2' + N\theta)
\end{aligned}
$$

$$
\begin{aligned}
&+\Sigma_1'\mu_{22}(\Sigma_2'+N\theta)\Big[-\sum_{k=2}^{N-1}\mu_{11}^{k-1}f(k,N)+\theta\alpha(N)+\mu_{22}\sum_{k'=2}^{N-1}\mu_{11}^{k'-1}f(k'+1,N)\Big]\\
&+\mu_{11}\Sigma_1'\mu_{22}\theta\alpha(N)\\
=&\Sigma_1'\mu_{22}(N-2)\theta\alpha(N)(\Sigma_2'+N\theta)\\
&+\Sigma_1'\mu_{22}(\Sigma_2'+N\theta)\Big[-\sum_{k=2}^{N-1}\mu_{11}^{k-1}[\mu_{22}+(k-1)\theta]f(k+1,N)+\theta\alpha(N)\\
&+\mu_{22}\sum_{k'=2}^{N-1}\mu_{11}^{k'-1}f(k'+1,N)\Big]+\mu_{11}\Sigma_1'\mu_{22}\theta\alpha(N)\\
=&\Sigma_1'\mu_{22}(N-2)\theta\alpha(N)(\Sigma_2'+N\theta)\\
&+\Sigma_1'\mu_{22}(\Sigma_2'+N\theta)\Big[-\sum_{k=2}^{N-1}\mu_{11}^{k-1}(k-1)\theta f(k+1,N)+\theta\alpha(N)\Big]+\mu_{11}\Sigma_1'\mu_{22}\theta\alpha(N)\\
=&\Sigma_1'\mu_{22}(\Sigma_2'+N\theta)\Big[(N-2)\theta\alpha(N)-\sum_{k=2}^{N-1}\mu_{11}^{k-1}(k-1)\theta f(k+1,N)+\theta\alpha(N)\Big]+\mu_{11}\Sigma_1'\mu_{22}\theta\alpha(N)\\
\geq&\Sigma_1'\mu_{22}(\Sigma_2'+N\theta)\Big[(N-1)\theta\alpha(N)-(N-1)\theta\sum_{k=2}^{N-1}\mu_{11}^{k-1}f(k+1,N)\Big]+\mu_{11}\Sigma_1'\mu_{22}\theta\alpha(N)\\
=&\Sigma_1'\mu_{22}(\Sigma_2'+N\theta)(N-1)\theta\Big[\sum_{k=2}^{N}\mu_{11}^{k-2}f(k,N)-\sum_{k=2}^{N-1}\mu_{11}^{k-1}f(k+1,N)\Big]+\mu_{11}\Sigma_1'\mu_{22}\theta\alpha(N).
\end{aligned}
$$

Let $k''=k+1$. Then

$$
\begin{aligned}
&(103)+(104)\\
=&\Sigma_1'\mu_{22}(\Sigma_2'+N\theta)(N-1)\theta\Big[\sum_{k=2}^{N}\mu_{11}^{k-2}f(k,N)-\sum_{k''=3}^{N}\mu_{11}^{k''-2}f(k'',N)\Big]+\mu_{11}\Sigma_1'\mu_{22}\theta\alpha(N)\\
=&\Sigma_1'\mu_{22}(\Sigma_2'+N\theta)(N-1)\theta f(2,N)+\mu_{11}\Sigma_1'\mu_{22}\theta\alpha(N)\geq 0.
\end{aligned}
$$

This proves $\Gamma_2(N,a_{11})\geq 0$. Furthermore for $s=N,\ldots,B$

$$
\begin{aligned}
&\Gamma_2(s+1,a_{11})-\Gamma_2(s,a_{11})\\
&=(g_N-\Sigma_2')\Big[(\Sigma_1'-\mu_{11})\Big(\frac{1}{\Sigma_2'+(s+1)\theta}-\frac{1}{\Sigma_2'+s\theta}\Big)+\mu_{22}\Big(\frac{1}{\Sigma_2'+s\theta}-\frac{1}{\Sigma_2'+(s-1)\theta}\Big)\Big]\geq 0,
\end{aligned}
$$

where the inequality follows as $(g_N-\Sigma_2')\leq 0$ from (87). Therefore $\Gamma_2(s,a_{11})\geq 0$ for $s=N,\ldots,B+1$. And hence $\Gamma(s,a_{11})\geq 0$ for $s=N,\ldots,B+1$ (if $N\geq 2$).

Hence sub-case (a) and the two scenarios for sub-case (b) prove that $\Gamma(s,a_{11})\geq 0$ for $s=N,\ldots,B+2$. This completes the proof of the theorem. □

A.4 Details of the proof of Proposition 11

We have

$$
\begin{aligned}
\tau_\gamma'(n)=&[\Sigma_2'+(n-1)\theta][\Sigma_2'+(n-2)\theta]\mu_{22}f(1,n-1)-[\Sigma_2'+(n-1)\theta][\Sigma_1'\alpha(n)(\Sigma_2'-\mu_{22})+\Sigma_2'f(1,n)]\\
&+[\Sigma_2'+(n-2)\theta]\mu_{11}[\Sigma_2'f(1,n-1)+\Sigma_1'\alpha(n-1)(\Sigma_2'-\mu_{22})]
\end{aligned}
$$

$$
\begin{aligned}
=&[\Sigma_2' + (n-1)\theta][\Sigma_2' + (n-2)\theta]\mu_{22} f(1,n-1) - (1+\lambda)\Big[[\Sigma_2' + (n-1)\theta][\Sigma_1\alpha(n)(\Sigma_2' - \mu_{22}) + \Sigma_2 f(1,n)] \\
&\quad + [\Sigma_2' + (n-2)\theta]\mu_{11}[\Sigma_2 f(1,n-1) + \Sigma_1\alpha(n-1)(\Sigma_2' - \mu_{22})]\Big] \\
=&(1+\lambda)\Big[[\Sigma_2 + (n-1)\theta][\Sigma_2 + (n-2)\theta]\mu_{22} f(1,n-1) - [\Sigma_2' + (n-1)\theta][\Sigma_1\alpha(n)(\Sigma_2' - \mu_{22}) + \Sigma_2 f(1,n)] \\
&\quad + [\Sigma_2' + (n-2)\theta]\mu_{11}[\Sigma_2 f(1,n-1) + \Sigma_1\alpha(n-1)(\Sigma_2' - \mu_{22})]\Big] \\
&+ [\Sigma_2' + (n-1)\theta][\Sigma_2' + (n-2)\theta]\mu_{22} f(1,n-1) - (1+\lambda)[\Sigma_2 + (n-1)\theta][\Sigma_2 + (n-2)\theta]\mu_{22} f(1,n-1) \\
=&(1+\lambda)\Big[[\Sigma_2 + (n-1)\theta][\Sigma_2 + (n-2)\theta]\mu_{22} f(1,n-1) - [\Sigma_2' + (n-1)\theta][\Sigma_1\alpha(n)(\Sigma_2 - \mu_{22}) + \Sigma_2 f(1,n)] \\
&\quad + [\Sigma_2' + (n-2)\theta]\mu_{11}[\Sigma_2 f(1,n-1) + \Sigma_1\alpha(n-1)(\Sigma_2 - \mu_{22})]\Big] \\
&- (1+\lambda)\Big[[\Sigma_2' + (n-1)\theta]\Sigma_1\alpha(n)\lambda\Sigma_2\Big] + (1+\lambda)\Big[[\Sigma_2' + (n-2)\theta]\mu_{11}\Sigma_1\alpha(n-1)\lambda\Sigma_2\Big] \\
&+ [\Sigma_2' + (n-1)\theta][\Sigma_2' + (n-2)\theta]\mu_{22} f(1,n-1) - (1+\lambda)[\Sigma_2 + (n-1)\theta][\Sigma_2 + (n-2)\theta]\mu_{22} f(1,n-1) \\
=&(1+\lambda)\Big[[\Sigma_2 + (n-1)\theta][\Sigma_2 + (n-2)\theta]\mu_{22} f(1,n-1) - [\Sigma_2 + (n-1)\theta][\Sigma_1\alpha(n)(\Sigma_2 - \mu_{22}) + \Sigma_2 f(1,n)] \\
&\quad + [\Sigma_2 + (n-2)\theta]\mu_{11}[\Sigma_2 f(1,n-1) + \Sigma_1\alpha(n-1)(\Sigma_2 - \mu_{22})]\Big] \\
&- (1+\lambda)\lambda\Sigma_2[\Sigma_1\alpha(n)(\Sigma_2 - \mu_{22}) + \Sigma_2 f(1,n)] + (1+\lambda)\lambda\Sigma_2\mu_{11}[\Sigma_2 f(1,n-1) + \Sigma_1\alpha(n-1)(\Sigma_2 - \mu_{22})] \\
&- (1+\lambda)\Big[[\Sigma_2' + (n-1)\theta]\Sigma_1\alpha(n)\lambda\Sigma_2\Big] + (1+\lambda)\Big[[\Sigma_2' + (n-2)\theta]\mu_{11}\Sigma_1\alpha(n-1)\lambda\Sigma_2\Big] \\
&+ [\Sigma_2' + (n-1)\theta][\Sigma_2' + (n-2)\theta]\mu_{22} f(1,n-1) - (1+\lambda)[\Sigma_2 + (n-1)\theta][\Sigma_2 + (n-2)\theta]\mu_{22} f(1,n-1) \\
=&(1+\lambda)\tau_1'(n) \\
&- (1+\lambda)\lambda\Sigma_2[\Sigma_1\alpha(n)(\Sigma_2 - \mu_{22}) + \Sigma_2 f(1,n)] + (1+\lambda)\lambda\Sigma_2\mu_{11}[\Sigma_2 f(1,n-1) + \Sigma_1\alpha(n-1)(\Sigma_2 - \mu_{22})] \\
&- (1+\lambda)\Big[[\Sigma_2' + (n-1)\theta]\Sigma_1\alpha(n)\lambda\Sigma_2\Big] + (1+\lambda)\Big[[\Sigma_2' + (n-2)\theta]\mu_{11}\Sigma_1\alpha(n-1)\lambda\Sigma_2\Big] \\
&+ [\Sigma_2' + (n-1)\theta][\Sigma_2' + (n-2)\theta]\mu_{22} f(1,n-1) - (1+\lambda)[\Sigma_2 + (n-1)\theta][\Sigma_2 + (n-2)\theta]\mu_{22} f(1,n-1) \\
=&(1+\lambda)\tau_1'(n) \\
&- (1+\lambda)\lambda\Sigma_2[\Sigma_1\alpha(n)(\Sigma_2 - \mu_{22}) + \Sigma_2 f(1,n)] + (1+\lambda)\lambda\Sigma_2\mu_{11}[\Sigma_2 f(1,n-1) + \Sigma_1\alpha(n-1)(\Sigma_2 - \mu_{22})] \\
&- (1+\lambda)\Big[[\Sigma_2' + (n-1)\theta]\Sigma_1\alpha(n)\lambda\Sigma_2\Big] + (1+\lambda)\Big[[\Sigma_2' + (n-2)\theta]\mu_{11}\Sigma_1\alpha(n-1)\lambda\Sigma_2\Big] \\
&+ \mu_{22} f(1,n-1)\Big[[\Sigma_2' + (n-1)\theta][\Sigma_2' + (n-2)\theta] - (1+\lambda)[\Sigma_2 + (n-1)\theta][\Sigma_2 + (n-2)\theta]\Big] \\
=&(1+\lambda)\tau_1'(n) \\
&- (1+\lambda)\lambda\Sigma_2[\Sigma_1\alpha(n)(\Sigma_2 - \mu_{22}) + \Sigma_2 f(1,n)] + (1+\lambda)\lambda\Sigma_2\mu_{11}[\Sigma_2 f(1,n-1) + \Sigma_1\alpha(n-1)(\Sigma_2 - \mu_{22})]
\end{aligned}
$$

$$
\begin{aligned}
&- (1+\lambda)\Big[[\Sigma_2' + (n-1)\theta]\Sigma_1\alpha(n)\lambda\Sigma_2\Big] + (1+\lambda)\Big[[\Sigma_2' + (n-2)\theta]\mu_{11}\Sigma_1\alpha(n-1)\lambda\Sigma_2\Big] \\
&+ \mu_{22}f(1,n-1)\Big[\Sigma_2'\Sigma_2' + \Sigma_2'(n-2)\theta + \Sigma_2'(n-1)\theta + (n-1)(n-2)\theta^2 - (1+\lambda)\Sigma_2\Sigma_2 \\
&- (1+\lambda)\Sigma_2(n-1)\theta - (1+\lambda)\Sigma_2(n-2)\theta - (1+\lambda)(n-1)(n-2)\theta^2\Big] \\
=&(1+\lambda)\tau_1'(n) \\
&- (1+\lambda)\lambda\Sigma_2[\Sigma_1\alpha(n)(\Sigma_2 - \mu_{22}) + \Sigma_2 f(1,n)] + (1+\lambda)\lambda\Sigma_2\mu_{11}[\Sigma_2 f(1,n-1) + \Sigma_1\alpha(n-1)(\Sigma_2 - \mu_{22})] \\
&- (1+\lambda)\Big[[\Sigma_2' + (n-1)\theta]\Sigma_1\alpha(n)\lambda\Sigma_2\Big] + (1+\lambda)\Big[[\Sigma_2' + (n-2)\theta]\mu_{11}\Sigma_1\alpha(n-1)\lambda\Sigma_2\Big] \\
&+ \mu_{22}f(1,n-1)\Big[\Sigma_2'\Sigma_2' - (1+\lambda)\Sigma_2\Sigma_2 - \lambda(n-1)(n-2)\theta^2\Big] \\
=&(1+\lambda)\tau_1'(n) \\
&- (1+\lambda)\lambda\Sigma_2[\Sigma_1\alpha(n)(\Sigma_2 - \mu_{22}) + \Sigma_2 f(1,n)] + (1+\lambda)\lambda\Sigma_2\mu_{11}[\Sigma_2 f(1,n-1) + \Sigma_1\alpha(n-1)(\Sigma_2 - \mu_{22})] \\
&- (1+\lambda)\Big[[\Sigma_2' + (n-1)\theta]\Sigma_1\alpha(n)\lambda\Sigma_2\Big] + (1+\lambda)\Big[[\Sigma_2' + (n-2)\theta]\mu_{11}\Sigma_1\alpha(n-1)\lambda\Sigma_2\Big] \\
&+ \lambda\Sigma_2\Sigma_2'\mu_{22}f(1,n-1) - \lambda(n-1)(n-2)\theta^2\mu_{22}f(1,n-1) \\
=&(1+\lambda)\tau_1'(n) \\
&- (1+\lambda)\lambda\Sigma_2\Big[\Sigma_1\alpha(n)\mu_{12} + \Sigma_2 f(1,n) - \mu_{11}\Sigma_2 f(1,n-1) - \mu_{11}\Sigma_1\alpha(n-1)\mu_{12} + [\Sigma_2' + (n-1)\theta]\Sigma_1\alpha(n) \\
&- [\Sigma_2' + (n-2)\theta]\mu_{11}\Sigma_1\alpha(n-1) - \Sigma_2\mu_{22}f(1,n-1)\Big] - \lambda(n-1)(n-2)\theta^2\mu_{22}f(1,n-1) \\
=&(1+\lambda)\tau_1'(n) \\
&- (1+\lambda)\lambda\Sigma_2\Big[\Sigma_1\alpha(n)\mu_{12} + \Sigma_2 f(1,n) - \mu_{11}\Sigma_2 f(1,n-1) - \mu_{11}\Sigma_1\alpha(n-1)\mu_{12} + \Sigma_2\Sigma_1\alpha(n) \\
&+ [\lambda\Sigma_2 + (n-1)\theta]\Sigma_1\alpha(n) - \Sigma_2\Sigma_1\mu_{11}\alpha(n-1) \\
&- [\lambda\Sigma_2 + (n-2)\theta]\mu_{11}\Sigma_1\alpha(n-1) - \Sigma_2\mu_{22}f(1,n-1)\Big] - \lambda(n-1)(n-2)\theta^2\mu_{22}f(1,n-1) \\
=&(1+\lambda)\tau_1'(n) \\
&- (1+\lambda)\lambda\Sigma_2\Big[\Sigma_1\alpha(n)\mu_{12} + \Sigma_2 f(1,n) - \mu_{11}\Sigma_2 f(1,n-1) - \mu_{11}\Sigma_1\alpha(n-1)\mu_{12} + \Sigma_2\Sigma_1\alpha(n) \\
&- \Sigma_2\Sigma_1\mu_{11}\alpha(n-1) - \Sigma_2\mu_{22}f(1,n-1)\Big] - (1+\lambda)\lambda\Sigma_2\Big[[\lambda\Sigma_2 + (n-1)\theta]\Sigma_1\alpha(n) \\
&- [\lambda\Sigma_2 + (n-2)\theta]\mu_{11}\Sigma_1\alpha(n-1)\Big] - \lambda(n-1)(n-2)\theta^2\mu_{22}f(1,n-1) \\
=&(1+\lambda)\tau_1'(n) \\
&- (1+\lambda)\lambda\Sigma_2\Big[\Sigma_1\alpha(n)\mu_{12} + \Sigma_2(\mu_{22} + (n-2)\theta)f(1,n-1) - \mu_{11}\Sigma_2 f(1,n-1) - \mu_{11}\Sigma_1\alpha(n-1)\mu_{12}
\end{aligned}
$$

$$
\begin{aligned}
&+\Sigma_2\Sigma_1\alpha(n)-\Sigma_2\Sigma_1\mu_{11}\alpha(n-1)-\Sigma_2\mu_{22}f(1,n-1)\Big]-(1+\lambda)\lambda\Sigma_2\Big[[\lambda\Sigma_2+(n-2)\theta]\Sigma_1\alpha(n)\\
&-[\lambda\Sigma_2+(n-2)\theta]\mu_{11}\Sigma_1\alpha(n-1)\Big]-(1+\lambda)\lambda\Sigma_2\theta\Sigma_1\alpha(n)-\lambda(n-1)(n-2)\theta^2\mu_{22}f(1,n-1)\\
=&(1+\lambda)\tau_1'(n)\\
&-(1+\lambda)\lambda\Sigma_2\Big[\Sigma_1\alpha(n)\mu_{12}-\mu_{11}\Sigma_2 f(1,n-1)-\mu_{11}\Sigma_1\alpha(n-1)\mu_{12}+\Sigma_2\Sigma_1\alpha(n)-\Sigma_2\Sigma_1\mu_{11}\alpha(n-1)\Big]\\
&-(1+\lambda)\lambda\Sigma_2^2(n-2)\theta f(1,n-1)-(1+\lambda)\lambda\Sigma_2[\lambda\Sigma_2+(n-2)\theta]\Sigma_1(\alpha(n)-\mu_{11}\alpha(n-1))\\
&-(1+\lambda)\lambda\Sigma_2\theta\Sigma_1\alpha(n)-\lambda(n-1)(n-2)\theta^2\mu_{22}f(1,n-1).
\end{aligned}
$$

A.5 Table and Figure for $\bar{N}$ Frequency Under Task Dependent Synergy

$\bar{N}$ value	Frequency	$\bar{N}$ Value	Frequency	$\bar{N}$ Value	Frequency
1	7702	31	6	61	2
2	326	32	4	62	2
3	339	33	8	63	3
4	296	34	8	65	1
5	242	35	7	66	1
6	176	36	11	67	3
7	139	37	4	70	1
8	101	38	5	71	1
9	73	39	1	72	3
10	70	40	3	73	1
11	51	41	5	74	1
12	36	42	3	75	2
13	41	43	3	76	4
14	41	44	4	78	1
15	22	45	3	80	2
16	24	46	2	83	2
17	21	47	2	84	2
18	21	48	2	85	1
19	9	49	4	86	1
20	15	50	4	87	2
21	18	51	3	88	2
22	20	52	2	89	1
23	10	53	4	90	1
24	11	54	5	92	1
25	11	55	1	93	1
26	4	56	2	94	1
27	7	57	2	96	1
28	4	58	1	98	1
29	5	59	2	99	1
30	6	60	1	101	1

Table 5: Frequency of $\bar{N}$ values

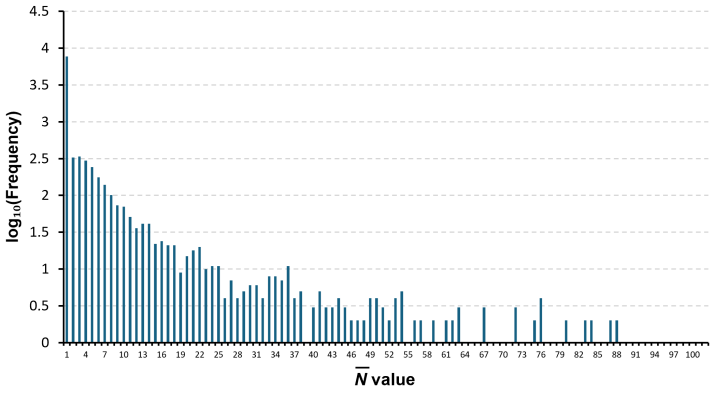


Fig. 5: Frequency distribution of $\bar{N}$ (log scale).