

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
NAAS Rating (2026): 4.49
Maths 2026; 11(9): 82-90
© 2026 Stats & Maths
<https://www.mathsjournal.com>
Received: 20-07-2026
Accepted: 24-08-2026

Naveen Kumar
Department of Mathematics,
Baba Mastnath University,
Asthal Bohar, Rohtak, Haryana,
India

Anu Rathee
Department of Mathematics,
Baba Mastnath University,
Asthal Bohar, Rohtak, Haryana,
India

Indu Rohilla
Department of Mathematics,
Baba Mastnath University,
Asthal Bohar, Rohtak, Haryana,
India

Corresponding Author:
Anu Rathee
Department of Mathematics,
Baba Mastnath University,
Asthal Bohar, Rohtak, Haryana,
India

Analysing Customer Service Efficiency in Commercial Banks Using a Multi-Server Queueing Model

Naveen Kumar, Anu Rathee and Indu Rohilla

DOI: <https://www.doi.org/10.22271/maths.2026.v11.i9a.2690>

Abstract

The multi-server queueing model was used in this study to compare the effectiveness of Bank customer service. Primary goal of current research work is to evaluate the effectiveness of customer service, system usage, average wait time and service capacity of both banks under various arrival rates. The data from both Banks, including arrival and service rates, during five consecutive days of week in (10:00am to 1:00pm) peak hours were used for this study. These results indicate that the Bank A ($\rho = 0.061$) uses the system on average less than Bank B ($\rho = 0.108$). Since the server is not overloaded in Bank A, it can easily handle heavy load during peak hours. Also, it requires a shorter average queue waiting time to expresses a relaxed service environment. However, the bank B is managing more customer requests with fewer servers and a comparatively shorter total processing time, indicating improved resource allocation with operational effectiveness. The outcome demonstrated that improving server allocation and putting request management techniques into practice led to an improved user experience.

Keywords: Multi-server Queueing Model, Bank operations, Arrival Rate, Service Delivery, Server Utilisation, Waiting Lines Management

1. Introduction

In the highly competitive retail Banking sector, service delivery speed and customer satisfaction are major urges of institutional growth. Long waiting times are a primary source of customer dissatisfaction, leading to decreased brand loyalty and capacity of customer defection. As a result, commercial banks face the continuous challenges of managing arrival spikes while maintaining operational efficiency. While expanding infrastructure or adding service counters can reduce customer waiting time, its synchronously administrative and employee overhead costs. To achieve this balance, quantitative frameworks like queueing theory provide a defensible, data-driven approach to analysing corporate resource management. Active research is therefore focused on queue analysis to optimise their operations and reduce waiting times for customers. Queueing theory aims to strike a balance between the two extremes of maximizing customer satisfaction and minimizing service rates. Due to inadequate queue management, most commercial banks experience queues on a daily basis. It converts unprocessed operational logs into predictable performance indicators by modelling customer behavioural patterns and service flow framework using probability distributions. This research paper focused on comparing and analysing the queue framework of two well-known commercial banks in India. Among these frameworks, the multi-server queueing model is particularly useful for modelling modern bank layouts, where a single, integrated waiting line feeds into multiple parallel teller windows. A comparative operational analysis of service delivery between two commercial banks is presented in this research article. Raimi and Nenuwa (2015) have identified optimal efficiency in Nigerian banks as one of the most important issues. They have modified the multi-server queue model to analyse data from two banks. The objective of this study project is to optimise customer satisfaction and service delivery. Farayibi and Oladapo (2016) evaluated the UBA branch in Agbor, Delta State, and reported that a careful analysis of service characteristics and operating costs has led to a significant reduction in waiting times and costs for customers. Bereket Tessema Zewude

(2016) ^[2] this article uses the multi-server queue model to analyse the arrival and service rate as well as the average line wait time. Two days of the week were observed simultaneously in order to collect data for both banks. The studies finding indicate that the bank offers the greatest number of waiting lines and achieves optimal performance when compared to the commercial bank Ethiopia Tona branch. Asogwa et. al., (2019) ^[4] study was the use of queue modelling by the Edeaja Group in the Nigerian banking sector Data were collected from the banking hall withdrawal area over four days for one month. The Markovian birth and death process has been used for data analysis. This study indicated that having more servers helps to reduce client waiting times and related costs. Al-Jazeera and Kilani (2020) examine link between service as long as the interested party fully believes that the service is being provided. Throughout the service provision process, the customer must have a clear presumption of the service demand. Banks attach to a set of service models to guarantee that services are rendered with the goal of increasing customers loyalty to their offerings. Aronu (2021) ^[16] evaluated the Agbor branch of the United Bank for Africa (UBA) in Delta State and reported that a careful analysis of service characteristics and operational costs improves efficiency as well as customer satisfaction.

Smith (2021) ^[6] a framework known as service culture centres on the values of offering safe services, the bank's standards, mission and vision, as well as the set of value propositions that the bank specifies. Rao and Singh (2021) ^[7] investigated the connection between Indian bank customers preferences and the availability of digital services. According to their research, customers favour banks with robust online and mobile banking features. One of the primary factors influencing customer loyalty and preferences is the ease of managing finance online without physically visiting a branch. The study made clear that in order to stay competitive, digital services must constantly be innovative. Gupta and Sharma (2022) ^[9] ^[8] studied the availability of banking services for the general public has increased due to positive reforms and competition in the Indian banking industry. Customers now have more options for high-quality, effective banking services due to increased bank competition and branch availability. Akande et. al., (2022) confirming that mathematical multi-server models allow financial institutions to evaluate queue dynamics precisely. Ultimately, expanding these methodologies across all self-service touchpoints will significantly improve institutional operational layout while drastically cutting customer transit times. Nair and Bhat (2022) ^[10] complemented these findings by establishing a connection between improved customer loyalty and high service quality in the Indian banking sector, stressing the need for prompt service, courteous staff and effective resolution of problems, highlighting the importance of timely service, polite employees, efficient problem-solving and ongoing staff training to uphold high standards of service. Mehta and Kapoor (2023) investigated how mobile banking apps affected customer engagement and discovered feature-rich, user-friendly apps with functions like bill payment, instant money transfers and mobile applications that are updated to incorporate new features. Yifter (2023) ^[11] used simulation-based queue models at the Commercial Bank of Ethiopia to predict and manage the arrival of clients, and demonstrated that predictive modelling improves the perceived quality of service. Ogunleye and Musa (2023) ^[12] show that structured queueing systems, including single- and multi-server models, can be successfully adapted to the African banking environment, resulting in improved efficiency and customer satisfaction. This technological integration turns queue management from a routine operational task into a strategic tool to increase competitiveness and the quality of services. Lolemo and Pandya (2024) ^[13] reminds us that virtual service touchpoints dictate modern customer perception. Merging both physical efficiency and seamless digital interfaces provides the ultimate framework for maximizing institutional service quality. The main factors were ease of use, affordability and speed of transaction. The study recommended that banks should continue investing in digital infrastructure and user-friendly interfaces to sustain and increase customer loyalty. Daramola (2024) ^[18] used multi-server queueing models to evaluate access banks and found that structured resource allocation is key to reducing wait times and increasing service efficiency. Bassey (2025) ^[17] studied some of the banks in Ogoja that effective waiting line management has a direct impact on customer satisfaction, demonstrating that operational interventions can have a tangible impact on service experiences. Chandra et al., (2025) ^[14] review the service delivery practices followed by three selected public sector banks and assess the use of queue model for the provision of services and highlights the banks strategic focus of on digital transformation to improve customer satisfaction and operational efficiency. Srivastava and Gupta (2025) ^[15] Study conducted to benchmark the efficiency of customer service in HDFC and ICICI Bank using the multi-server queue model. The main role of this research document is to assess the effectiveness of customer service, the use of the system, the average length for customers and the service capacity of both banks in different arrival frequencies.

2. Objectives

1. Review the provision of services in selected banks.
2. Compare the provision of services in selected banks using the Queueing model.
3. Analyse key determinants of service delivery using factor analysis.

3. Methodology

To extract data for the queue model, the observational method is used to analyse the average of five working days per week in banks. Observations shall be held between 10:00 AM and 1:00 PM, which is a working day. The average waiting time, the average service rate and the utilisation factor can be easily estimated from the statistics provided. In compiling this research paper from secondary sources, journals, articles, internet source, annual reports used the personal observations of selected banks.

3.1. Method of Data Analysis

To mathematically evaluate the operational efficiency of selected commercial banks, a standard multi-server queueing model, specifically the $(M/M/C):(\infty - FCFS)$ system is employed. In this framework, customer arrivals follow a Poisson distribution with an average arrival rate of λ , while service times follows an exponential distribution with an average service rate of μ per server. The system operates under ' c ' identically parallel servers on a FCFS basis.

3.2. Mathematical notation

m : Number of customers in the system (waiting time + service facility) at time t
 $\lambda(t)$: Mean customer arrival rate per unit time (Arrival intensity),
 $\mu(t)$: Mean customers service rate per unit of time (Server capacity),
 $P_m(t)$: Steady state probability of finding exactly n customers in the system at time t ,
 $L_q(t)$: Expected (average) line length (number of customers waiting in the queue),
 $L_s(t)$: Expected number of customers in the system (waiting + being served),
 $W_q(t)$: Expected time spent per customers waiting strictly in the queue,
 $W_s(t)$: Expected time a customer spends inside the system (in waiting + being served),
 ρ : System utilisation,
 $P_0(t)$: Stationary probability of an entirely empty, idle system ($1 - \rho$),
 $P_w(t)$: The probability that a customer has to wait,
 c : The number of servers available to provide service

The main features of the Multi-server Queueing system are the following:

A. Traffic Intensity/ Server Utilisation(ρ)

The system utilisation factor, denoted by ρ , represents the average traffic intensity and signifies the probability that a specific server is busy. It is mathematically formulated as:

$$\rho = \frac{\lambda}{c \cdot \mu} \quad (1)$$

Where $\rho < 1$ is the strict steady-state condition required to prevent the queue from growing infinitely, λ is the mean arrival rate of customers, c is the number of parallel server and μ is the service rate per hour.

B. Probability of zero customers in the system(ρ_0)

The probability that the queueing system is entirely idle, designated as (ρ_0), is derived using the steady- state probability distribution as follows:

$$P_0(t) = \left\{ \sum_{n=0}^{c-1} \frac{1}{n!} \left(\frac{\lambda}{\mu} \right)^n + \frac{1}{c!} \left(\frac{\lambda}{\mu} \right)^c \frac{c\mu}{c\mu - \lambda} \right\}^{-1} \quad (2)$$

$$P_m(t) = \begin{cases} \left(\frac{\rho^m}{m!} \right) P_0, & \text{if } m \leq c \\ \frac{\rho^m}{(c! c^{m-c}) P_0}, & \text{if } m > c \end{cases} \quad (3)$$

Additional customers must wait if $m \geq c$, indicating that the number of customers exceeds the number of servers.

$$C(c, \rho) = \sum_{m=c}^{\infty} P_m(t) = \frac{\rho^c}{c! (1 - \rho c)} P_0(t) \quad (4)$$

C. Expected queue length/ Average customers in queue (L_q)

Expected number of customers queueing before receiving service, denoted by L_q , is formulated as:

$$L_q(t) = \left\{ \frac{1}{(c-1)!} \left(\frac{\lambda}{\mu} \right)^c \frac{\mu\lambda}{(c\mu - \lambda)^2} \right\} P_0(t). \quad (5)$$

D. Average number of customers in the system(L_s)

At the same time, the expected total number of customers on the Bank establishment (both waiting and served), denoted by L_s , is given by:

$$L_s(t) = L_q(t) + \frac{\lambda(t)}{\mu(t)} = \left\{ \frac{1}{(c-1)!} \left(\frac{\lambda(t)}{\mu(t)} \right)^c \frac{\mu\lambda}{(c\mu - \lambda)^2} \right\} P_0(t) + \frac{\lambda(t)}{\mu(t)} \quad (6)$$

E. Performance waiting metrics (W_q and W_s)

Utilising Little's formulas, average time the customer spends waiting in the queue (W_q) and overall average time spent in the system (W_s) which includes both waiting and service time, are expressed as:

$$W_q(t) = \frac{L_q(t)}{\lambda(t)} = \left\{ \frac{1}{(c-1)!} \left(\frac{\lambda}{\mu} \right)^c \frac{\mu}{(c\mu - \lambda)^2} \right\} P_0(t) \quad (7)$$

$$W_s(t) = \frac{L_s(t)}{\lambda(t)} = \left\{ \frac{1}{(c-1)!} \left(\frac{\lambda(t)}{\mu(t)} \right)^c \frac{\mu\lambda}{(c\mu - \lambda)^2} \right\} P_0(t) + \frac{1}{\mu} \quad (8)$$

Where $L_q(t)$ is the average number of customers in the queue, $L_s(t)$ is the average number of customers in the system and $\lambda(t)$ is the average arrival rate.

3.2. Data Analysis

The empirical data collected over a five-day observation period for Bank A and Bank B were processed using the $(M/M/C)$ queueing model. To analyse the operational characteristics of both financial institutions, the empirical data encompassing customer arrival rates and service rates were documented simultaneously over a five-day observation window. Table 1 presents the comparative baseline data for Bank A and Bank B. The computed system performance metrics are condensed and presented in table 1.

Table 1: Customer arrival rate $\{\lambda(t)\}$ and service rate $\{\mu(t)\}$ for Bank A and Bank B

Days	Institution	Arrival Rate $\{\lambda(t)\}$	Service Rate $\{\mu(t)\}$
Monday	Bank A	60	56
	Bank B	62	54
Tuesday	Bank A	48	42
	Bank B	55	50
Wednesday	Bank A	43	30
	Bank B	40	36
Thursday	Bank A	34	32
	Bank B	36	30
Friday	Bank A	30	25
	Bank B	32	22

From the data in Table 1, Bank A and Bank b exhibits a significantly higher server utilisation rate (ρ) on Bank A compared to Bank B is lowest server utilisation rate (ρ). The spike indicated a critical congestion level, suggesting that the current number of active servers (c) is insufficient during peak opening days in Bank B. To, visually demonstrates the daily fluctuations and behavioural trends of customer queues, a comparative graphical representation:

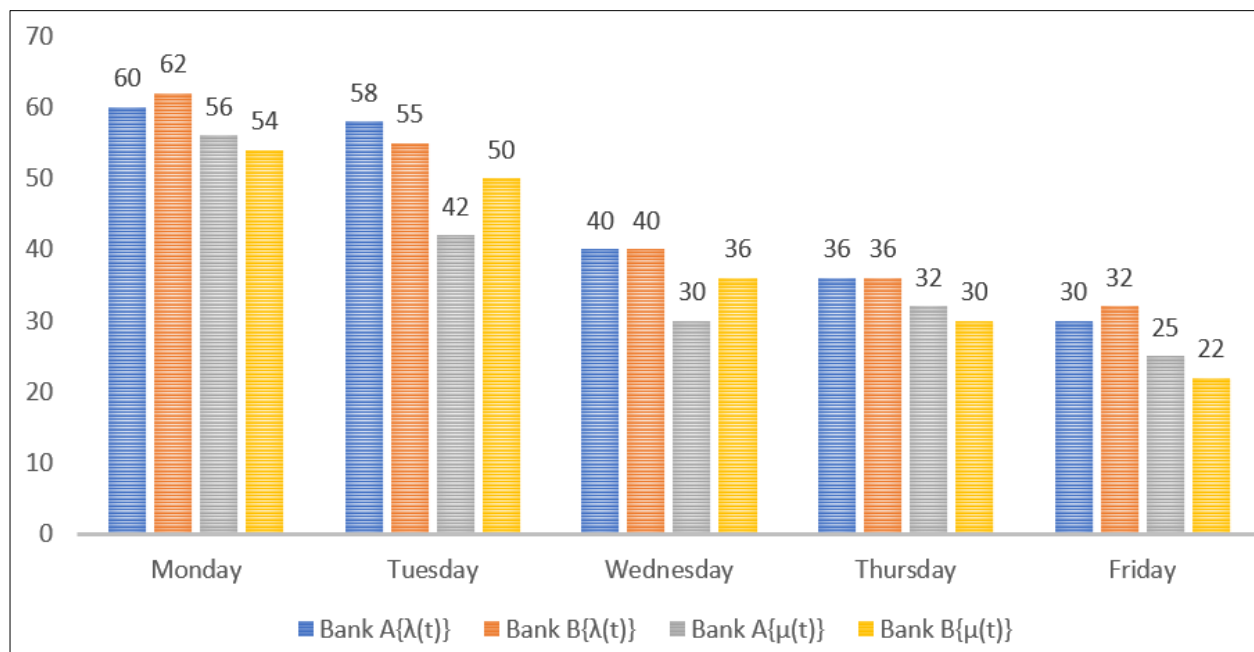


Fig 1: Daily variation of customer arrival rate $\{\lambda(t)\}$ and service rate $\{\mu(t)\}$ of Bank A and Bank B

The graphical trajectory in Figure 1, both banks experience a summarized escalation in arrival rates on Monday, clearly highlights a distinct peak in Monday for both institutions, which statistically confirms the beginning of the week rush.

3.3. Computing Bank A Queueing Framework:

By applying the steady-state formulations of the $M/M/C$ multi-server queueing model in detail, the secondary performance indicators were computed.

When a line exists let $c = 7$, $\lambda = 215$, $\mu = 185$, $\rho = \frac{\lambda}{\mu} = 1.1621$,

$$P_0(t) = \left\{ \sum_{m=0}^{c-1} \frac{1}{m!} \left(\frac{\lambda}{\mu} \right)^m + \frac{1}{c!} \left(\frac{\lambda}{\mu} \right)^c \frac{c\mu}{c\mu - \lambda} \right\}^{-1}$$

$$\begin{aligned} \text{Now, } \sum_{n=0}^{c-1} \frac{1}{n!} \left(\frac{\lambda}{\mu} \right)^n &= \sum_{n=0}^6 \frac{1}{n!} \left(\frac{215}{185} \right)^n \\ &= \frac{1}{0!} \left(\frac{215}{185} \right)^0 + \frac{1}{1!} \left(\frac{215}{185} \right)^1 + \frac{1}{2!} \left(\frac{215}{185} \right)^2 + \frac{1}{3!} \left(\frac{215}{185} \right)^3 + \frac{1}{4!} \left(\frac{215}{185} \right)^4 + \frac{1}{5!} \left(\frac{215}{185} \right)^5 + \frac{1}{6!} \left(\frac{215}{185} \right)^6 \\ &= 3.1957 \end{aligned}$$

Therefore, we have

$$P_0(t) = \left\{ 3.1957 + \frac{1}{7!} \left(\frac{215}{190} \right)^7 \frac{7 * 185}{7 * 185 - 215} \right\}^{-1} = (3.1967)^{-1} = 0.3128$$

$$L_q(t) = \left\{ \frac{1}{(c-1)!} \left(\frac{\lambda}{\mu} \right)^c \frac{\mu\lambda}{(c\mu - \lambda)^2} \right\} P_0 = \left\{ \frac{1}{(7-1)!} \left(\frac{215}{185} \right)^7 \frac{215 * 185}{(5 * 185 - 215)^2} \right\} 0.312 = 0.082$$

$$L_s(t) = L_q + \frac{\lambda}{\mu} = 1.245$$

$$W_q(t) = \frac{L_q}{\lambda} = 0.00038$$

$$W_s(t) = \frac{L_s}{\lambda} = 0.0057$$

Now, we calculate for lines 2 to 7 using the formula above, i.e.,

$$L_s(t) = \frac{\rho}{1 - \rho} = 1.3866,$$

$$L_q(t) = \frac{\rho^2}{1 - \rho} = 0.8054,$$

$$W_s(t) = \frac{1}{\mu - \lambda} = 0.0129,$$

$$W_q(t) = \frac{\rho}{\mu - \lambda} = 0.00749.$$

In a similar way, the analysis remains the same if the lines are 3 to 7.

Table 3: Estimated queueing framework for Bank A (7 server)

Queueing Parameter	Mean Arrival rate	Mean Service rate	Server utilisation	Average queue length	Average number of customers	Average time in system	Average time in Queue
Symbol	$\lambda(t)$	$\mu(t)$	ρ	$L_s(t)$	$L_q(t)$	$W_s(t)$	$W_q(t)$
1	215	185	1.1621	1.2450	0.0829	0.0057	0.00038
2	107.5	185	0.5810	1.3866	0.8054	0.0129	0.00749
3	71.66	185	0.3873	0.6361	0.2448	0.0088	0.00341
4	53.75	185	0.2905	0.4094	0.1188	0.0076	0.00221
5	43	185	0.2334	0.3027	0.0703	0.0070	0.00163
6	35.83	185	0.1936	0.2400	0.0301	0.0067	0.000129
7	30.71	185	0.1660	0.1990	0.0329	0.0064	0.00107

Table 3 shows that as the line sequence increase, the customers arrival rates decrease from 215 to 30.71 customers per hour. This systematic change reduces traffic intensity (ρ) from a saturated 1.1621 to an optimized 0.1660. The results confirm that strategic load distribution effectively eliminates operational traffic congestion.

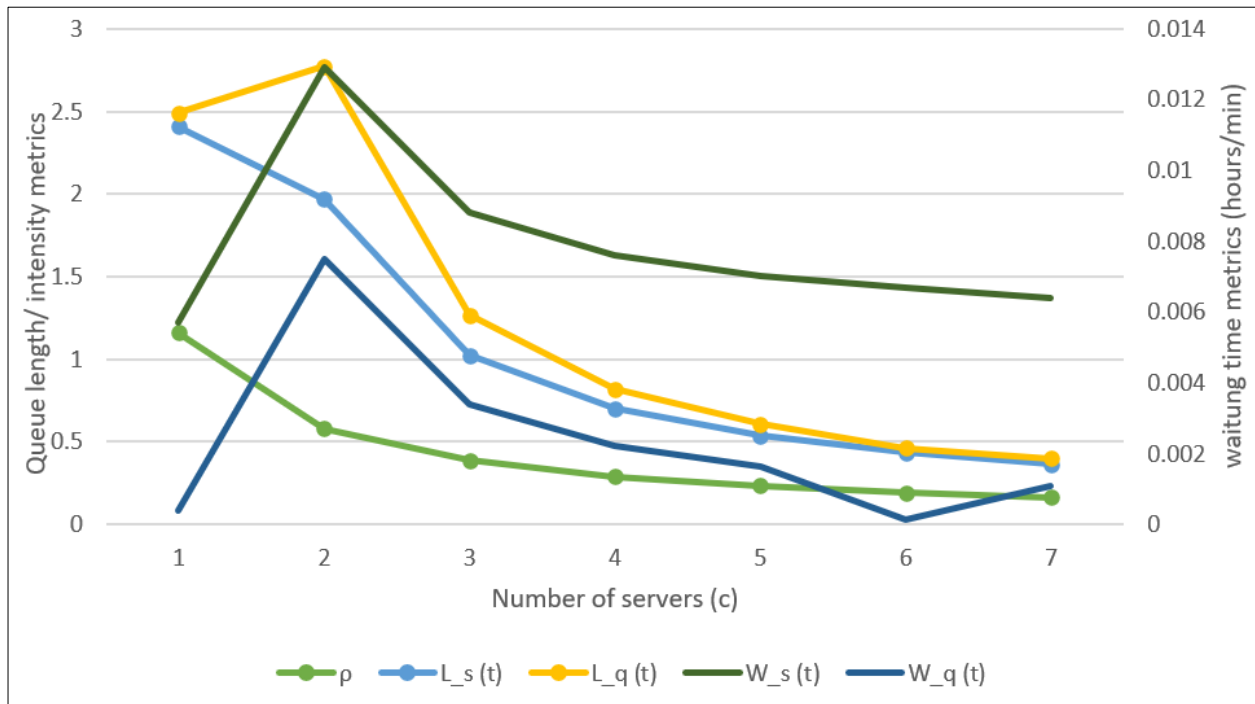


Fig 3: Sensitivity analysis of queueing framework $\{\rho, L_s(t), L_q(t), W_s(t), W_q(t)\}$ against varying number of waiting lines/servers ($c = 7$)

Mathematical method gives that the customer who arrives at Bank A must wait for service

$$P_W(t) = \left(\frac{\lambda}{\mu}\right)^c \frac{P_0}{c! \left(1 - \frac{\lambda}{c\mu}\right)} = \left(\frac{215}{185}\right)^7 \frac{0.3128}{7! \left(1 - \frac{215}{7 \cdot 185}\right)} = 0.0021\%.$$

3.4. Computing Bank B Queueing Framework

There are five servers in the Bank B, whenever there is a line $c = 5$,

$$\lambda(t) = 225, \mu(t) = 190, \rho = \frac{\lambda}{\mu} = \frac{225}{190} = 1.1842.$$

$$P_0(t) = \left\{ \sum_{m=0}^{c-1} \frac{1}{m!} \left(\frac{\lambda}{\mu}\right)^m + \frac{1}{c!} \left(\frac{\lambda}{\mu}\right)^c \frac{c\mu}{c\mu - \lambda} \right\}^{-1}$$

$$\begin{aligned} \text{Now, } \sum_{m=0}^{c-1} \frac{1}{m!} \left(\frac{\lambda}{\mu}\right)^m &= \sum_{n=0}^4 \frac{1}{n!} \left(\frac{225}{185}\right)^n \\ &= \frac{1}{0!} \left(\frac{225}{185}\right)^0 + \frac{1}{1!} \left(\frac{225}{185}\right)^1 + \frac{1}{2!} \left(\frac{225}{185}\right)^2 + \frac{1}{3!} \left(\frac{225}{185}\right)^3 + \frac{1}{4!} \left(\frac{225}{185}\right)^4 \\ &= 3.2439 \end{aligned}$$

Therefore,

$$\begin{aligned} P_0(t) &= \left\{ 3.2439 + \frac{1}{5!} \left(\frac{225}{185}\right)^5 \frac{5 \cdot 190}{5 \cdot 190 - 225} \right\}^{-1} \\ &= (3.2693)^{-1} \\ &= 0.3058 \end{aligned}$$

$$\begin{aligned} L_q(t) &= \left\{ \frac{1}{(c-1)!} \left(\frac{\lambda}{\mu}\right)^c \frac{\mu\lambda}{(c\mu - \lambda)^2} \right\} P_0 \\ &= \left\{ \frac{1}{(5-1)!} \left(\frac{225}{190}\right)^5 \frac{225 \cdot 190}{(5 \cdot 190 - 225)^2} \right\} 0.3058 \\ &= 0.00229 \end{aligned}$$

$$L_s(t) = L_q + \frac{\lambda}{\mu} = 0.1864$$

$$W_q(t) = \frac{L_q}{\lambda} = 0.000010$$

$$W_s(t) = \frac{L_s}{\lambda} = 0.00082.$$

The method described above allows us to sequentially for lines 2 to 5.

Table 4: Estimated queueing framework values for Bank B (5 server)

Waiting Line(c)	$\lambda(t)$	$\mu(t)$	ρ	$L_s(t)$	$L_q(t)$	$W_s(t)$	$W_q(t)$
1	225	190	1.1842	1.1864	0.0029	0.00082	0.000010
2	112.5	190	0.5921	1.4515	0.8592	0.0129	0.00764
3	75	190	0.3947	0.6520	0.2572	0.0086	0.00343
4	56.25	190	0.2960	0.4204	0.1244	0.0074	0.00221
5	45	190	0.2368	0.3102	0.0733	0.0068	0.00163

Table 4 shows that as the line sequence increase, the customers arrival rates decrease from 225 to 45 customers per hour. This systematic change reduces traffic intensity (ρ) from a saturated 1.1842 to an optimized 0.2368. The results confirm that strategic load distribution effectively eliminates operational traffic congestion.

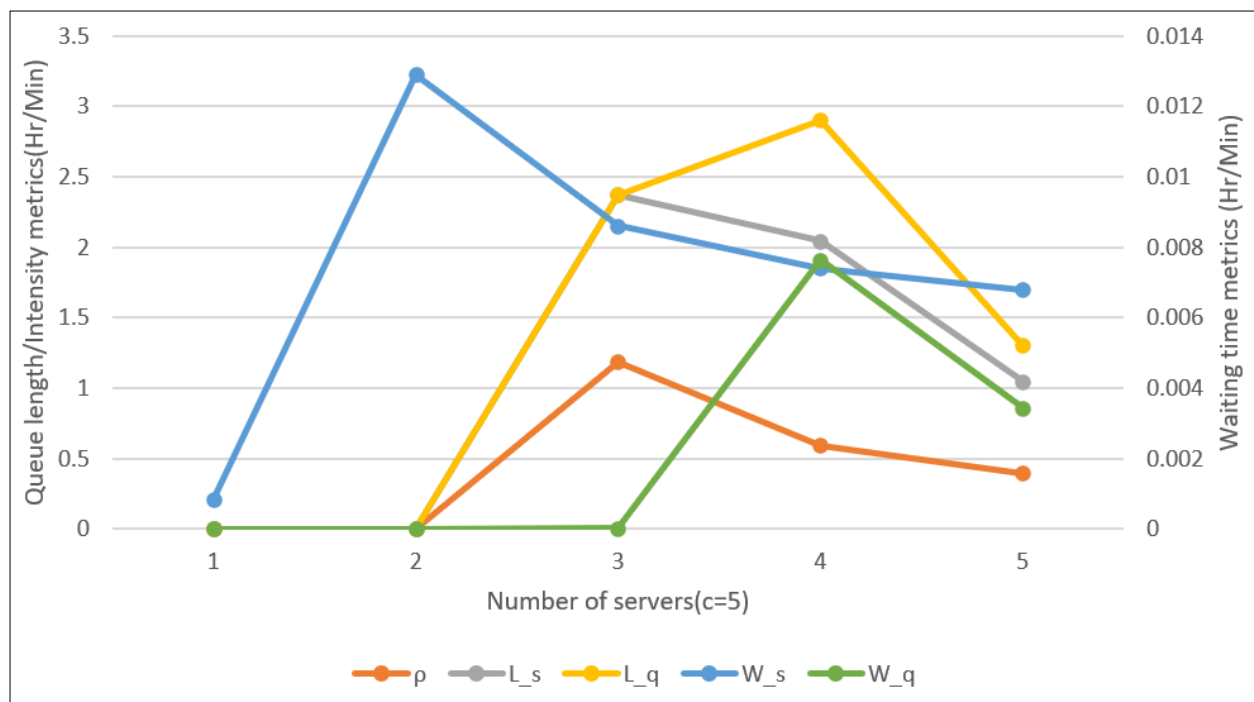


Fig 4: Sensitivity analysis of queueing framework $\{\rho, L_s(t), L_q(t), W_s(t), W_q(t)\}$ against varying number of waiting lines/servers ($c = 5$)

This formula gives the probability that a new customer will experience a wait for service at the bank.

$$P_w(t) = \left(\frac{\lambda}{\mu}\right)^c \left\{ \frac{P_0}{c! \left(1 - \frac{\lambda}{c\mu}\right)} \right\} = \left(\frac{225}{190}\right)^7 \left\{ \frac{0.3058}{5! \left(1 - \frac{225}{5 \times 190}\right)} \right\} = 0.90\%.$$

To evaluate and compare the operational performance of selected commercial banking institutions, a multi-server queueing model was successfully simulated. The primary empirical data gathered regarding customer service frequencies and arrival densities were proceed using standard stochastic formulations. The comprehensive statistical output, illustrating a comparative layout of server utilisation framework, waiting probabilities and dynamic time factors for both Banks is systematically structured in Table 5.

Table 5: Comparative Operational performance and Queueing Metrics of Bank A and Bank B Using $M/M/C$ Model

Parameter	Operational Metric/Description	Bank A	Bank B
$\lambda(t)$	Average Arrival Rate (Customer/Hour)	79.63	102.75
$\mu(t)$	Average Service Rate per Server (Customer/Hour)	185	190
c	No of Active channels Servers	7	5
ρ	Average Server Utilisation Factor	0.0614	0.1081
$P_W(t)$	Probability that a Customer Waits in Queue	0.021%(Very Low)	0.90% (Moderate)
$W_q(t)$	Average Time a customer waits in Queue	0.00103 min	0.063 min
$W_s(t)$	Average Total Time in System	0.0064 min	0.068 min
Commentary	System Queue Operational Status	Optimized Traffic; Low congestion	Higher Arrival Density; Needs Monitoring

4. Results

The empirical find is structured in Table 5 revealing critical operational dimensions regarding the service on both Banks. A comparative review indicates that Bank B experiences a higher average customer arrival density ($\lambda = 102.75$ customers/hr) compared to Bank A ($\lambda = 79.63$ customers/hr). Despite operating with fewer active servers ($c = 5$), Bank B maintains a slightly lower total time in the system ($W_s = 0.068$ min) due to its marginally faster average transaction processing speed ($\mu = 190$ customers/hr per counter). Conversely, Bank A exhibits an optimized traffic safety margin. Utilizing seven servers ($c = 7$) keeps its average server utilisation factor ($\rho \approx 0.061$) lower than Bank B traffic density ($\rho \approx 0.108$). As a result, the chance of waiting (P_W) in Bank A is almost zero (0.021%) and a customer's wait only (0.00103 min) in the queue. In conclusion, both Banks are running safely and efficiently because their workload factor is well below 100% ($\rho < 1$). There is a greater possibility of waiting for customers, which highlights the need for optimisation of staff or scheduling in order to manage the high flow of customers efficiently.

5. Conclusion

The primary goal of this study is to use a multi-server queue model to compare the queue frameworks of Bank A and Bank B. The study's conclusions suggest that Bank B in particular should think about expanding the number of service windows or optimizing staffing during peak hours in order to lower the possibility of system overload and waiting in both Banks. The analysis indicates that Bank A utilizes a larger counter capacity to achieve near-zero waiting times, whereas Bank B effectively manages higher customer density through faster transaction processing speeds, offering actionable insights for optimizing branch operations. This study concluded that both banks could improve the quality of their services and customer satisfaction through efficient queue management and better resource planning. Future research can extend this study by incorporating a queue based on arrival predictions.

References

1. Abiodun RO, Omosule I. Queuing model for banking system: a comparative study of selected banks in Owo Local Government Area of Ondo State, Nigeria. *Int J Eng Res Technol*. 2015;4(4):191-195.
2. Zewude BT. Queuing modeling for comparative study of banking system on Commercial Bank of Ethiopia Tona Branch and Dashen Bank: the case of Wolaita Zone, Ethiopia. *Res J Finance Account*. 2016;7(21):11-16.
3. Farayibi O. Optimizing service levels in Nigerian banks using queuing models: a case study of GT Bank and Ecobank. *J Financ Oper*. 2016;8(1):33-50.
4. Asogwa OC, Eze CM, Edeaja MT. On the application of queuing model in Nigeria banking sector, a case study of Zenith Bank Plc, Abakaliki Branch. *Int J Curr Res*. 2019;19(2):1-8.
5. Alhijawi B, Kilani Y. A collaborative filtering recommender system using genetic algorithm. *Inf Process Manage*. 2020;57(6):192310. doi:10.1016/j.ipm.2020.102310.
6. Smith JA. Understanding service culture in banking: principles, norms, and quality. *Financ Serv J*. 2021;15(3):45-67.
7. Rao P, Singh R. Exploring the relationship between digital service delivery and customer preferences in Indian banks. *J Financ Serv Mark*. 2021;26(3):210-225.
8. Gupta R, Sharma P. Impact of competition and reforms on the Indian banking industry. *J Bank Finance*. 2022;34(1):56-78.
9. Akande SA. Application of multi-server queue model (M/M/c) for waiting lines management in banking system. *Asian J Math Sci*. 2022;6(1):1-8.
10. Nair S, Bhat R. Linking high service quality with increased customer loyalty in the Indian banking sector. *J Financ Serv Mark*. 2022;27(2):194-210.
11. Yifter G. Simulation-based queuing model for enhancing service quality in Ethiopian banks. *J Afr Financ Stud*. 2023;11(2):77-94.
12. Ogunleye O, Musa K. Technology-enhanced queuing systems in Nigerian deposit money banks: a desk review. *J Oper Manage Emerg Econ*. 2023;9(2):101-120.
13. Lolemo S, Pandya H. The impact of digital banking on customer satisfaction and loyalty in commercial banks: a systematic literature review. *Int J Manage Econ Comm*. 2024;1(1):69-75.
14. Chandra Ch S, Kumar AP, Samisetty S, Kumar P. Service delivery in select public sector banks using queuing model - a comparative study. *Acad Mark Stud J*. 2025;29(1):1-17.
15. Srivastava RK, Gupta H. A comparative study between two commercial banks using multi server queuing model. *Int J Res Sci Technol*. 2025;15:21-31.
16. Aronu C. Analysis of queuing systems in United Bank for Africa: implications for service efficiency. *Int J Manage Econ*. 2021;6(1):112-130.

17. Bassey I. Managing waiting lines in Nigerian banks: an empirical study of selected branches in Ogoja. J Bank Oper. 2025;7(2):56-70.
18. Daramola A. Application of multi-server queuing models to optimize service efficiency: case study of Access Bank Plc, Anyigba, Kogi State. Int J Oper Res. 2024;15(4):210-230.