Investigating the Application of Queue Theory in the Nigerian Banking System

FARAYIBI, Adesoji

Centre for Allied Research and Economic Development, Ibadan, Oyo State, Nigeria.

6 September 2016

Investigating the Application of Queue Theory in the Nigerian Banking System

FARAYIBI, Adesoji Oladapo
Centre for Allied Research and Economic Development, Ibadan, Oyo State, Nigeria.
E-mail: afarayibi2000@gmail.com

ABSTRACT
This study examined the application of queue theory in the banking system in Nigeria, with particular reference to GTBank and Ecobank Idumota branch, Lagos, Lagos state. The queuing characteristics of the banks were analyzed using a Multi-Server Queuing Model. The performance measures analysis including the waiting and operation costs for the banks were computed with a view to determining the optimal service level. Findings revealed that the traffic intensity was higher in GTbank with p =0.98 than in Ecobank with p= 0.78. Also, the potential utilization showed that Ecobank was far below efficiency compared to GTBank. Looking at the waiting time of customers in line and the time spent in the system, that is (Wq + Ws), we discovered that customers in Ecobank spent more time before being served both on queue and in the system than that of GTBank bank. The study concluded by emphasizing the relevance of queuing theory to the effective service delivery of the banking sector in Nigeria and strongly recommends that for efficiency and quality of service delivery to customers, the management of GTBank and Ecobank should adopt a 13-server model and 10-server model respectively to reduce total expected costs and increase customer satisfaction.
Keywords: Queue Theory, Banking System, Multi-Server Model, Traffic Intensity, Waiting Cost

1. INTRODUCTION
Today's era of globalization is characterized by an intense and ever increasing competition both within and across industries, and the financial institutions are not exempted. Bank operation managers continue to experience wrenching challenges, which they must keep up with for survival. Banks being major component of the financial system, intermediating between the surplus and deficit sectors of the economy, are always the center of attraction to many customers that want to carry out one transaction or the other through the services provided by these banks.
A common feature of Nigerian banks is overcrowded banking halls. This often leads to poor level of customer satisfaction and movement of customers from one bank to the other, seeking for better banking services without much delay. The modern day automation of bank services (such as Online Real Time, Automated Teller Machine (ATM), POS etc.), with the intention of minimizing queue problem has not yielded the desired results due to frequent network breakdowns and inadequate high level professionals to manage the server. Hence, long queue had persisted in all Nigerian banks. Since the time available to any economic agent has alternative uses, for leisure or for work, incessant queues often lead to economic wastage.

Also, the number of hours devoted to work determines the individual's wage. Apart from situation of work or leisure, economic agents sometime commit considerable amounts of their time when they come into service stations for service. Typical of Nigeria, customers wait for hours on queue to get service without the service providers feeling that there is an opportunity cost for the wasted time. Many customers have thus wondered why banks cannot employ additional staffs or increase their branch network by opening more branches in order to provide better and efficient service and boost the level of customer satisfaction (Ogunsakin, et al, 2013).

However, capacity management is an important aspect in the design of service operations. These decisions involve a trade-off between the costs of sustaining a service level standard and the value that customers attach to it. This trade-off analysis is very important in order to meet the profitability objectives of the stakeholders of the bank. In the light of the level of service in the Nigerian financial sector, particularly as it relates to labour and customer turnover, it is important to understand and identify the effect of queues on customer satisfaction in banks in the country. This study therefore investigates the application of queue theory in the banking system in Lagos. The study aims at determining the following queue characteristics:

I.   Average number of customer in a queue
II.   Average waiting time of customers in the queue
III.   Average waiting time of customers in the system
IV.   Average number of customers in the system
V.   The probability that an arriving customer has to wait for service
VI.   The probability of 'n' customer being in the system

The rest of the paper is as follows: section two is the literature review. Section three presents the theoretical framework and research methodology while section four presents the empirical analysis and results. The last section five covers the conclusion and recommendations.

2. EMPIRICAL LITERATURE

This section provides a review of empirical studies on the effect of waiting on customer satisfaction and its implications for the managing queues in service delivery in Nigerian banks. This review will help to identify the relevant variables for the analysis.

Ogunsakin et al 2013 did a comparative analysis of Service Delivery by ATM in Two Banks with the application of Queuing Theory. From the empirical analysis, the study found that the average arrival rate, average service rate, average time spent in the queue for Access bank as 2.01, 1.65, 0.5 respectively and UBA as 3.28, 1.75, 1.67minutes, respectively. The study concluded that the average number of idle time obtained for the two banks were 3minutes and 7minutes respectively.

In examining the Queuing Process and its Application to Customer Service Delivery in Fidelity Bank Plc, Maiduguri by Bakari, et al (2014), the study obtained the value of the traffic intensity, otherwise known as the utilization factor to be less the one (i.e. $\rho<1$). The study concluded that the system operates under steady-state condition. Thus, the value of the traffic intensity, which is the probability that the system is busy, implies that 95% of the time period considered during data collection the system was busy as against 4% idle time. This indicates high utilization of the system.

In offering the Queuing model as a Technique of Queue solution in Nigeria Banking Industry, Anichebe (2013) found that, using a three-server system was better than a 2-server or 4-server systems in terms of the performance criteria. The study recommended that the management should adopt a three-server model to reduce total expected costs and increase customer satisfaction. Ogbadu and Usman (2012) studied the Imperatives of Customer Relationship Management in Nigerian Banking Industry. Findings from the study revealed that there is a direct relationship between customer relationship management and customer loyalty as well as

banks profitability. The study recommended that management of banks should pursue customer relationship management programmes with rigour to achieve the business objectives of the bank. Mandelbaum and Zeltyn (2004) used analytical queuing models with customer impatience to explain nonlinear relationships between waiting time and customer abandonment. They found that, in the context of call-center outsourcing, the common use of service level agreements based on delay thresholds at the upper-tail of the distribution (e.g. 95% of the customers wait less than 2 minutes) was consistent with non-linear effects of waiting on customer behavior. Taylor (1994) studies a survey of delayed airline passengers and finds that delay decreases service evaluations by invoking uncertainty and anger affective reactions. Deacon and Sonstelie (1985) evaluated customers' time value of waiting based on a survey on gasoline purchases. They found that although surveys are useful to uncover the behavioral process by which waiting affects customer behaviour and the factors that mediate this effect; they suffer from some disadvantages.

Forbes (2008) analyzed the impact of airline delays on customer complaints, showing that customer expectations play an important role mediating this effect. Campbell and Frei (2010) studied multiple branches of bank, provided empirical evidence that teller waiting times affect customer satisfaction and retention. Their study revealed significant heterogeneity in customer sensitivity to waiting time, some of which could be explained through demographics and the intensity of competition faced by the branch.

Aksin-Karaesmen et al, (2011) modelled callers' abandonment decision as an optimal stopping problem in a call center context, and found heterogeneity in caller's waiting behavior. The study also looked at customer heterogeneity in waiting sensitivity and related this sensitivity to customers' price sensitivity. They found that association between price and waiting sensitivity has important managerial implications. Also Afeche and Mendelson (2004) and Afanasyev and Mendelson (2010) showed that it plays an important role for setting priorities in queue and it affects the level of competition among service providers.

Odirichukwu et al, (2014) examined the banking queue system in Nigeria. The study uncovered the applicability and extent of usage of queuing models in achieving customer satisfaction at the lowest cost. The study recommended that, if First-in-First-out Queue Method is implemented, the design achieves an orderly service delivery. Also customers who have successful gotten the queue number should be attended to first based on FIFO-Queue Model already programmed.

Odunukwe (2013) examined the application of queuing models to customers management in the banking system using United Bank for Africa, Okpara Avenue Branch Enugu, as a case study. The results obtained from the study showed that the arrival pattern follows a poisson distribution and that the service pattern follows an exponential distribution. The study recommended that the Bank management should increase the number of servers to three so as to help reduce the time customers spend on queue and also reduce cost incurred from waiting.

Despite these studies, there is still a gap in the literature as regards studies that empirically investigate queue theory parameters in the area of study selected. Therefore, as a contribution to the existing literature on queue theory and service delivery, the present study makes use of a mathematical model to harness the complexities of queuing situations in the banking halls in Nigeria.

3. THEORETICAL FRAMEWORK AND METHODOLOGY

3.1 Theoretical Framework

Comparatively, Economics as a social science describes the factors that determine the production, distribution and consumption of goods and services. While queue theory is a branch of operations research which aids decision making process about how such goods and services can be property administered. Therefore, queue theory has a theoretical linkage with economic theory since it is applicable in the theory of the firm. For instance, production is the lifeblood of any company since a lot goes into making it happen. Queue theory is the mathematics which is applicable to the company as it strives to match product –generation with service delivery. Queue theory determines the quality of the distribution of the company(s) goods and services. However, the distribution of goods and services may be cost focused and not customers focus. Theoretically, distribution center is often perceived as a cost center since it hardly provides revenue for firm. Proper application of queue theory not only helps the firm(s) to solve the problems associated with distribution or service delivery but also gives the firm a competitive edge over those that are not applying it. How a customer experience waiting and the effect of waiting on commerce may in all cases affect the profitability of the firm. Once a customer wait more than necessary, the perceived waiting time multiplies with each passing minute (Glassen, 2007).

Also, queues are source of dissatisfaction to customers on one hand and economics cost to organizations on the other hand. The act of waiting has significant impact on customers' dissatisfaction. The amount of time customers must spend waiting not only affects customers' satisfaction but also customers' expectations or attribution for the causes of the waiting. (Talyor, 1994; Obaniro, 2010). Waiting time has been cited as the most important cause of dissatisfaction of customers when involved in one banking transaction or the other (trout et al 2000). Delays have also been associated with adverse outcomes (Derlet and Richards, 2000) and increase violence in the banking hall (Stirling et al, 2001). Queue theory, as a mathematical approach, is thus relevant to the study of waiting in lines since its application minimizes cost by reducing inefficiencies and delays in service delivery and ultimately increasing customer satisfaction (Singli, 2006, Kartz, Larson and Larson, 1991; Davis and Heineke, 1994).

3.2 Queuing Modelling Specification

Queuing theory modeling is classified by using special (or standard) notations described by D.G. Kendall (1953) in the form of (*a/b/c*). Later, A.M. Lee added the symbols *d* and *e* to the Kendell notation. In the literature of queuing theory, the standard format used to describe queuing models is as follows:

$$\{(a/b/c) : (d/e)\}$$

Where *a* = arrival distribution
     b = service time distribution
     c = number of servers (service channels)
     d = capacity of the system (queue plus service)
     e = queue (or service discipline)
In place of notation *a* and *b*, the following descriptive notations are used for the arrival and service times distribution:

M = Markovian (or Exponential) inter-arrival time or service-time distribution.

D= Deterministic (or constraint) inter-arrival time or service topic.

G =General distribution of service time, i.e. no assumption is made about the type of distribution with means and variance

GI=General probability distribution normal or uniform for inter-arrival time.

E_K= Erlang-k distribution for inter-arrival or service time with parameter k (i.e. if k = 1, Erlang is equivalent to exponential   and if k = ∞, Erlang is equivalent to deterministic).

3.3 Research Area

Both the old and new generation banks all have queue problems in some certain respect and degrees. Customers usually have to wait endlessly before they are served. The banks used for this study are GTBank and Ecobank; both are located at Idumota area of Lagos, Lagos state. The choice of these banks was informed by the level of queue in these banks and the fact that the researcher is familiar with the services of these banks.

3.4 Method of Data Collection

Basically, the data used for this study were obtained from primary sources. The method of data collection is through direct observation. A wrist watch, a pen and a notepad were requirements needed for the recording of relevant information such as; number of customers the arrival times of customers, waiting time, and service time. The observation was made during the working hours (8am – 4pm). The recorded information was used to calculate average waiting time, average service time and the utilization factor.

3.5 Method of Analysis

The method of analysis for this study is the multi-server queuing modeling system which follows (M/M/S): (∞/FCFS) specification. In the case, the performance measure analysis including, the arrival time, waiting time service time, priority level, for average customers and the number of servers available were computed using the appropriate tools. Secondly, the graphical representation of the generated performance measure values was done. The third stage shows the computation of the costs involved using queuing analytical models.

3.5.1. Multi-Server Queuing Models: ((M/M/S): (∞/FCFS))

In this case instead of a single server, there are multiple but identical servers in parallel to provide service to customers. It is assumed that queues are formed and customers are served on a first come first serve basis by any of the servers (Bunday, 1996). The service times are distributed exponentially with an average of μ customers per unit of time. If there are n customers in the queuing system at any point in time, then the following cases may arise:

i.   If $n < s$, (number), then there will be no queue of customers in the system is less than the number of servers. However, $(s - n)$ number of servers will not be busy. The combined service rate will then be $\mu_n = n\mu$.

ii.   If $n \geq s$, (number of customers in the system is more than or equal to the number of servers), then all servers will be busy and the maximum number of customers in the queue will be $(n - s)$. The combined service rate will be $\mu_u = s\mu$.

The opening of account in a bank by customers follows this a multi-server queuing modeling because a specific sequence of steps are usually followed: initial contact at the customers service, desk, filling out forms, making identification processing of Account number etc. Because several servers are usually available for this procedure, more than one customer at a time may be processed.

Where:

○ = Person

▢ = Processing point
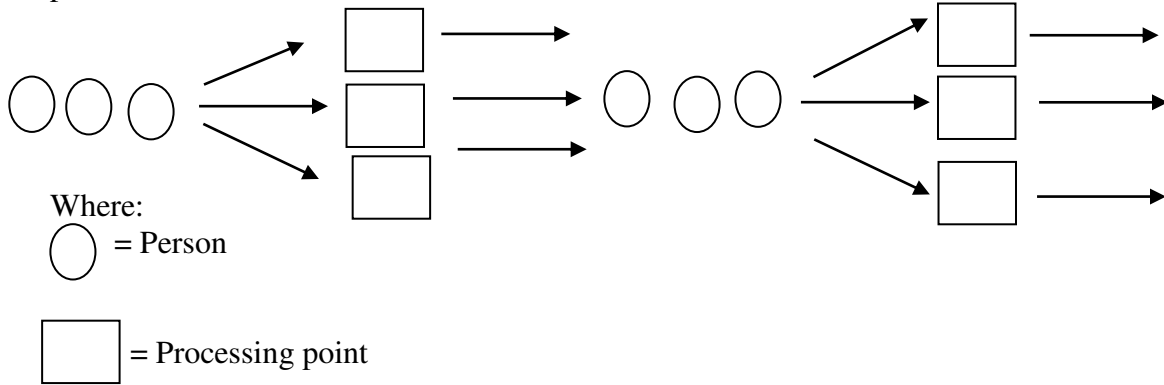
In case a bank customer forms a queue (after leaving the customer service unit or if the customers service unit is congested form a queue to be serviced by the letters (servers) in parallel). The servers have unequal speed which is dependent on the type of transaction the customer presents.

3.5.2 Traffic Intensity and Measures of Effectiveness
The following are the formulae for measuring traffic intensity and effectiveness in the banking industry using queuing modeling (Bunday, 1996):

   1. The probability that a system will be idle. That is there is no queue:

$$P0 = \left(\sum_{n=0}^{s-1} \frac{(s\rho)^n}{n!} + \frac{1}{s!} \frac{(s\rho)^s}{1-\rho}\right)^{-1}$$

$$\rho = \lambda/s\mu$$

$$P0 = \left(\sum_{n=0}^{s-1} \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n + \frac{1}{s!}\left(\frac{\lambda}{\mu}\right)^s \frac{s\mu}{s\mu-\lambda}\right)^{-1}$$

2. The expected number of customers waiting in the queue (queue length)

$$Lq = \left(\frac{1}{s-1}\left(\frac{\lambda}{\mu}\right)^s \left(\frac{s\mu}{s\mu-\lambda}\right)\right) P0$$

3. The expected number of customer in the system.

$$Ls = Lq + \frac{\lambda}{\mu}$$

4. The expected waiting time of a customer in the queue:

$$Wq = \left(\frac{1}{(s-1)!}\left(\frac{\lambda}{\mu}\right)^s \frac{\mu}{(s\mu-\lambda)^2}\right) P0 = Lq/\lambda$$

6

5. The expected waiting time that a customer spends in the system.

$$Ws = Wq + \frac{1}{\mu} = \frac{Lq}{\lambda} + \frac{1}{\mu}$$

6. Probability that all servers are simultaneously busy (or utilization factor), i.e. there will be queue:

$$Pw\,(n \geq s) = \frac{1}{s!}\left(\frac{\lambda}{\mu}\right)^{S} \frac{s\mu}{s\mu - \lambda}\; P0.$$

4. EMPIRICAL ANALYSIS AND RESULTS

4.1 Queue Analysis
Queue analysis involves the mathematical and quantitative representation of the queue theory with a view to determining the value of the average arrival rate in both banks and the average service rate in the two banks. This includes the results for a service pattern of multiple servers in parallel and the output of the M/M/S :(∞/FCFS) queuing modeling system, where S =16, S =15, S =14 and S =13 respectively for GTBank and S=10 and S=9 for Ecobank . This step involves the usage of the data collected during the field study.

4.1.1 Performance Measure Analysis for GTB Bank
Table 4.1 below shows the results of the performance measure analysis of GTBank for different servers. The results for an M/M/S16:(∞/FCFS) shows the average number of customers in the system, L =11.25; The average length of the queue, Lq = 0.25; The average time a customer spends in the system, W = 3.066min; The average time a customer spends in the queue, Wq = 0.0011; The probability that a system will be idle, P_0 = 0.000016; The Probability that all servers are simultaneously busy (utilization factor), Pw(n $\geq$ s) = 0.1124 while the Traffic intensity $\dot{\rho}$ = 69%.

Comparatively, the measures of effectiveness for an M/M/S15:(∞/FCFS) puts the average number of customers in the system, L =11.52; The average length of the queue, Lq = 0.52; The average time a customer spends in the system, W = 0.0524; The average time a customer spends in the queue, Wq = 0.002372; The probability that a system will be idle, P_0 = 0.00001584, The Probability that all servers are simultaneously busy (utilization factor), Pw(n $\geq$ s) = 0.1894 while the traffic intensity is 73%.

Table 4.1: Performance Measure Analysis for GTB bank

Performance Measure	S16	S15	S14	S13	S12
$\rho = \dfrac{\lambda}{s\mu}$	0.6875	0.7333	0.7851	0.8462	0.9167
P0	0.000016	0.00001584	0.00001491	0.00001294	0.0000088
Pw (n ≥ s)	0.1124	0.1894	0.3027	0.4663	0.6939
Lq	0.25	0.52	1.1113	2.5647	7.61
Ls	11.25	11.52	12.1113	13.5647	18.61
Wq	0.0011	0.002372	0.0051	0.0117	0.0346
Ws	0.0511	0.0524	0.0551	0.0617	0.084
Operating cost	5840	5745	5110	4745	4380
Waiting cost	1068	1095	1151	1290	1768
Total cost	6908	6570	6261	6035	6148

Source: Computed with data obtained from field survey

The measures of effectiveness for an M/M/S14:(∞/FCFS) shows the average number of customers in the system, L =12.11; The average length of the queue, Lq = 1.1113; The average time a customer spends in the system, W = 0.0551; The average time a customer spends in the queue, Wq = 0.0051; The probability that a system will be idle, P0 = 0.00001491; The Probability that all servers are simultaneously busy (utilization factor), Pw(n ≥ s) = 0.3037 while the traffic intensity is 79%.

Also, the measures of effectiveness for an M/M/S13:(∞/FCFS) puts the average number of customers in the system, L = 13.56; The average length of the queue, Lq = 2.5647; The average time a customer spends in the system, W = 0.0617; The average time a customer spends in the queue, Wq = 0.0117; The probability that a system will be idle, P_0 = 0.00001294; The Probability that all servers are simultaneously busy (utilization factor), Pw(n ≥ s) = 0.4663 while the traffic intensity is 85%.

4.1.2 Introducing Costs into the GTB Bank Performance Model

In order to evaluate and determine the optimum number of servers in the system, two opposing costs must be considered in making these decisions: (i) Service/operating costs (ii) Waiting time costs of customers. Economic analysis of these costs helps the management to make a trade-off between the increased costs of providing better service and the decreased waiting time costs of customers derived from providing that service.

Results from table 4.1 shows that operating cost and waiting cost for M/M/S16:(∞/FCFS) are N5840 and N1068 while operating cost and service cost for M/M/S15:(∞/FCFS) are N5475 and N1095 respectively. Comparatively, the operating cost and waiting cost for M/M/S14:(∞/FCFS) are N5119 and N1151 while the operating cost and waiting costs for M/M/S13:(∞/FCFS) are 4745 and 1290 respectively. Therefore for a balance between service to customers (short queues implying many servers) and economic considerations (not too many servers), the GTB management would do well to maintain with 13 servers since it has the lowest operating and waiting costs.

4.2:1 Performance Measure Analysis for ECOBANK.

Table 4.2 below shows the results of the performance measure analysis of Ecobank for different servers. The results for an M/M/S9:(∞/FCFS) shows the average number of customers in the system, L =2.11, The average length of the queue, Lq = 0.41, The average time a customer spends in the system, W = 0.062min; The average time a customer spends in the queue, Wq = 0.012, The probability that a system will be idle, P0 = 0.1657; The Probability that all servers are simultaneously busy (utilization factor), Pw(n $\geq$ s) = 0.3132, while the traffic intensity is 56%. Comparatively, the measures of effectiveness for an M/M/S10:(∞/FCFS) puts the average number of customers in the system, L =6.12; The average length of the queue, Lq = 4.42; The average time a customer spends in the system, W = 0.18; The average time a customer spends in the queue, Wq = 0.13, The probability that a system will be idle. That is there is no queue: P0 = 0.08105, The Probability that all servers are simultaneously busy (utilization factor). There will be queue. Pw(n $\geq$ s) = 0.78 while the traffic intensity 85%.

Table 4.2: Performance measure analysis for Ecobank

Performance measure	S9	S10
$\rho = \dfrac{\lambda}{s\mu}$	0.56	0.85
P0	0.1657	0.08105
Pw (n $\geq$ s)	0.3132	0.78
Lq	0.41	4.42
Ls	2.11	6.12
Wq	0.012	0.13
Ws	0.062	0.18
Operating cost	1095	730
Waiting cost	200	581
Total cost	1395	1311

Source: Computed with data obtained from field survey

4.2.2 Introducing Costs into the ECOBANK Performance Model

In order to evaluate and determine the optimum number of servers in the system, two opposing costs must be considered in making these decisions: (i) Service/operating costs (ii) Waiting time costs of customers. Economic analysis of these costs helps the management to make a trade-off between the increased costs of providing better service and the decreased waiting time costs of customers derived from providing that service.

Results from table 4.2 shows that operating cost and waiting cost for M/M/S9:(∞/FCFS) are 1095 and 200 while operating cost and service cost for M/M/S9:(∞/FCFS) are N730 and N581 respectively. The provision of an additional service mechanism may be capital intensive to Ecobank; hence it would pay the bank more to maintain 10 servers since the primary aim of every business organization besides profit making is customer satisfaction.

4.3 Comparative Analysis of GTB Bank And ECOBANK Services
This section presents a comparative analysis of the services of both banks used for the study. Using the optimal number of servers for both banks, GTB bank is better with 13 service points while Ecobank is better with 10 service points. It is upon these parameters that the comparative analysis for both banks is done. The major parameters for this comparison include:

DESCRIPTION	GTBank M/M/S13:(∞/FCFS)	Ecobank M/M/S10:(∞/FCFS)
The average number of customers in the system.	$Ls = 13.56$	$Ls = 6.12$
The average length of the queue	$Lq = 2.5647$	$Lq = 4.42$
The average time a customer spends in the system	$Ws = 0.0617$	$Ws = 0.18$,
The average time a customer spends in the queue	$Wq = 0.0117$	$Wq = 0.13$
The probability that a system will be idle. That is, there is no queue	$P_0 = 0.00001294$	$P_0 = 0.08105$
The Probability that all servers are simultaneously busy (utilization factor). There will be queue.	$Pw(n \geq s) = 0.4663$	$Pw(n \geq s) = 0.78$,
Traffic intensity	56%,	72%,

- **GTBank: Ls = 169 customers Vs Ecobank: Ls =60 customers**.
An average of 169 customers was in the banking hall per time in GTbank while an average of 60 customers was in the system in Ecobank. Hence, more people were in the system in GTBank than in Ecobank.

- **GTBank: Lq = 30 customers Vs Ecobank: Lq =44 customers**
An average of 30 customers was on queue in GTbank while an average of 44customers was on queue in Ecobank. Thus, the queuing situation is better in GTBank than in Ecobank. That is, more people are on queue in Ecobank than in GTBank.

- **GTBank: Ws = 3.07 minute Vs Ecobank: Ws = 10.08 minutes**
The average waiting time for service in GTBank was 3.07 minutes while the average waiting time in Ecobank was 10.08minutes. Therefore, customers spend more time in the banking hall of Ecobank to access banking services than they do in GTBank.

- **GTBank: Wq = 1.2 minute Vs Ecobank: Ws = 7.08 minutes**
The average time customers spent on the queue in Ecobank was 1.2 minutes while customers spent an average of 7.08 minutes on the queue in Ecobank. This implies that customers spend more time on the queue in Ecobank than in GTBank.

- **GTBank: (Wq+ Ws) = 4.27 minutes Vs Ecobank: (Wq+ Ws) = 17.16 minutes**
Looking at the waiting time of customers in line and the time in the banking hall, that is (Wq+ Ws), we discovered that customers in Ecobank spent more time before being served both on queue and in the system than in GTBank. (17.16 minutes compared to 4.27minutes).
- **GTBank: Pw = 47% Vs Ecobank: Pw = 80%**
The capacity utilization in GTBank was 47% while it is 80% in Ecobank. This implies that staffs and service units are not fully utilized in GTBank while staff and service points were tending towards over-utilization in Ecobank.

- **GTBank: Traffic intensity, ρ = 56% Vs Ecobank: Traffic intensity, ρ = 72%**

The traffic intensity for GTBank is 0.56 i.e. 56% and that of Ecobank is 0.72 which amount to 72%. The results show that traffic intensity was higher in Ecobank than in GTBank.

Generally, based on the comparative analysis of the banks on the yardstick of their performance measurement, customers have better banking experience in GTBank than in Ecobank because the former offers improved customer service delivery than the later. Although, both banks were presumed to be working at optimal service levels, there is still room for improvement in their service delivery so that customers can have better banking experience than what is obtainable now.

5. CONCLUSION AND RECOMMENDATIONS

This study has validated that the queuing theory is applicable to the operations of the Nigerian banking industry. The application is most relevant where customers would have to queue in a congested banking hall before they access essential banking services. This situation is true for GTBank and Ecobank as it is for many bank branches in the country. Excessive queue and waste of time in the banking hall would be reduced in GTBank, Ecobank and the entire Nigerian banking sector if service facilities are expanded using best queuing models that consider cost optimization.

Having realized that quality service delivery is a prerequisite for achieving a high level of customer satisfaction in the Nigeria banking industry, the following recommendation are made for the management of GTBank and Ecobank.

- The management of both banks should adopt an optimal service model that reduces total expected costs and increases customer satisfaction. To this end, GTB bank is better with 13 service points while Ecobank is better with 10 service points.
- The queue characteristics of the banks should be viewed from the stand point of customers as to whether the waiting time is reasonable and acceptable by making queue discipline fair and varying the number of service channels according to queue circumstances.
- The management should educate their operation managers and other staff on the application of queuing models to operational problems.
- The management of both banks should entrust their employees, empower them and enrich their skill and job involvement through continuous training so as to improve their ability for better customer service delivery.
- The management of both banks should motivate their employees in order to improve their productivity.
- Additional service channels such as ATM machines, POS etc. can also be deployed within the bank's premises in order to minimize the waiting time and waiting cost of customers and hence reducing the inconveniences and frustrations associated with waiting.

REFERENCE.

Aksin-Karaesmen, Baris, A., Emadi, S. and Che-Lin S. (2011). Structural Estimation of Callers Delay Sensitivity in Call Centers. *International Journal Service & Industrial. Management.* 6 (1), 42 -53

Anichebe, A.(2013). Queuing Model as a Technique of Queue Solution in Nigeria Banking Industry. *Developing Country Studies*, 3(8), pp. 34-42.

Ashley, D.W. (2000): "*Introduction to waiting lime models*" Google http/www.googlescholcom.

Ashley, D.W. (2006), *"Introduction to Waiting Line Models"* Available at Google, http/www.google academic.com.

Bakari, C. and Baba, A. (2014). Queuing Process and its Application to Customer Service Delivery: A Case study of Fidelity Bank Plc, Maiduguri. *International Journal of Mathematics and Statistics Invention,* 2(1), 14-21.

Bunday, B. D. (1996). *An Introduction to Queuing Theory*, New York: Halsted Press, Davis MM.

Campbell, D. and Frei, F. (2010). Market Heterogeneity and Local Capacity Decisions in Services. *Manufacturing & Service Operations Management Research*, 5 (1),31-41.

Churchman, Ackoff and Arnoff (1957). *Introduction to Operation Research* New York, John Willey and SC pp. 8-9.

Cooper, R.B. (1972). *Introduction to Queuing Theory*, New York, Macmillan.

Davis, M.M. and Heineke, J. (2000). Understanding the role of the customer and the operation for better. *www.businessballs.com › leadership/management.*

Deacon and Sonstelie, (1985). Rationing by Waiting and the Value of Time: Results from a Natural Experiment. *Journal of Political Economy* 93(4), 627–647.

Derlet and Richards (2000); Emergency crowding department: *www.systemdynamics.org /conferences/2011/proceed/papers/P1369.pdf.*

Heineke, J. (1994). Understanding the roles of the customer and the operation for better queue management. *International Journal of Operational and Production Management*. 14 (5), 21-34.

Heineke, J. and Davis, M. M. (1998). How disconfirmation, perception and actual waiting times impact customer satisfaction. *Int. J. Serv. Ind. Manage.* 9 (1), 64 -73.

Kendall, (1953). Classification of Queuing Model...*people.revoledu.com/kardi/tutorial/ Queuing/Kendall- Notation.html.*

Obamiro, J.K. (2000). Queuing theory and Patient Satisfaction: An overview of terminology & application in Ante-Natal care unit. http://www.upg-bulletin-se.ro.

Odunukwe Adaora (2013). Application of Queuing Models to Customers Management in the Banking System (A Case Study of United Bank for Africa, Okpara Avenue Branch Enugu): An unpublished undergraduate project, Caritas University, Enugu state, Nigeria.

Ogbadu and Usman, (2012). Imperatives of Customer Relationship Management in Nigeria Banking Industry. *Kuwait Chapter of Arabian Journal of Business and Management Review* 2(1), 123-132.

Stirling, E. T. (2001*)*. Queue theory and applications in health care- increased violence in emergency departments Patient. *www.nets.nihr.ac.uk/_data/assets/pdf_file/0003/.../FR-08-1204-029.pdf*

Taylor, S. (1995). The effects of filled waiting time and service provider control over the delay on evaluations of service. *Journal of Academic Marketing Science*, 23(2), 38-48.

Taylor, S. (1994). Waiting for Service: The Relationship between Delays and Evaluation of Service. *Journal of Marketing*, 58(2): 56-69.

APPENDIX A

Table A1: Calculation of operating cost

Description	N
Machine a cost	150000
Depreciation (Straight line method)	50000/yr
Cost per day = 50000/251	199.2
Cost per hour	22
Labour per month	65000
Cost per day = 65000/21	3095
Cost per hour	343
Operating cost	365/hr

Table A2: Calculation of waiting cost

Description	N
Minimum wage per month	18000
wage per day = 18000/21	857.14
Minimum wage per hr	95
Waiting Cost = λ x Ws x 95	1068

APPENDIX B

Performance Measurement Analysis for GTBank: S_{16}

$$P0 = \left(\sum_{n=0}^{s-1} \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n + \frac{1}{s!}\left(\frac{\lambda}{\mu}\right)^s \frac{s\mu}{s\mu-\lambda}\right)^{-1}$$

$= (1+ 11+ 60.5 + 221.83 + 610 + 1342.09 + 2460.5 + 3866.5

$+ 5316.44 + 6497.87 + 7174.66 + 7174.66 + 6552.02 + 5544.2

$+ 4356.01 + 3194.41 + 2196.16 \times 3.2)^{-1}$

$= (54329.72 + 2196.16 \times 3.2)^{-1} = 0.000016$

$P_w (n \geq s) = \frac{1}{s!}\left(\frac{\lambda}{\mu}\right)^s \frac{s\mu}{s\mu-\lambda} P0 = 2196.16 \times 3.2 \times 0.000016 = 0.1124$

$Lq = \left(\frac{1}{s-1}\left(\frac{\lambda}{\mu}\right)^s \left(\frac{s\mu}{s\mu-\lambda}\right)\right) P0 = 35138.5 \times 0.44 \times 0.000016 = 0.25$

$Ls = Lq + \frac{\lambda}{\mu} = 0.25 + 11 = 11.25$

$Wq = Lq/\lambda = 0.25/220 = 0.0011$

$Ws = Wq + \frac{1}{\mu} = 0.0011 + 0.05 = 0.0511/hr = 0.0511 \times 60 = 3.066min$

Performance Measurement Analysis for GTBank: S_{15}

$P0 = \left(\sum_{n=0}^{s-1} \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n + \frac{1}{s!}\left(\frac{\lambda}{\mu}\right)^s \frac{s\mu}{s\mu-\lambda}\right)^{-1} = (\ldots\ldots + 51135 + 3194.4 \times 3.75)^{-1} = 0.00001584$

$P_w (n \geq s) = \frac{1}{s!}\left(\frac{\lambda}{\mu}\right)^s \frac{s\mu}{s\mu-\lambda} P0 = 3194.41 \times 3.75 \times 0.00001584 = 0.1897$

$Lq = \left(\frac{1}{s-1}\left(\frac{\lambda}{\mu}\right)^s \left(\frac{s\mu}{s\mu-\lambda}\right)\right) P0 = 47916.15 \times 0.6875 \times 0.00001584 = 0.52$

$Ls = Lq + \frac{\lambda}{\mu} = 11.52$

$Wq = Lq/\lambda = 47916.15 \times 0.00001584 = 0.002372$

$Ws = Wq + \frac{1}{\mu} = 0.0524$

Note: The Performance Measurement Analysis for s_{14} and s_{13} (GTBank) and s_9 and s_{10} (Ecobank) follow the same procedure as above.