Modeling the Effect of Population Size on Banking Transaction Channels in Nigeria: Grey Box vs Support Vector Regression

Desmond C. Bartholomew[1,2], **Ngozi P. Olewuezi**[1], **Chrysogonus C. Nwaigwe**[1], **and Felix C. Akanno**[1]

This study investigates the effect of Nigeria's population on four selected banking transaction channels. The Nigerian projected population (2022-2027) was used as an input variable for forecasting future volumes of transactions for each channel. The results show that the Support Vector Regression (SVR) model best fits the ATM, Online, and USSD channels of transaction while the Grey-box was better for POS. The forecast results show that ATM, online, and USSD channels had their highest volume of transactions in 2023, while for POS, the highest volume was recorded in 2027. Further results indicate that online and POS transactions would dominate payment system landscape in the future. To better serve customers in the future, Nigerian banks should expand their capacity for online and POS transactions.

Keywords: Grey-box, banking channels, Nigeria population, support vector regression, POS, USSD.

JEL Classification: C4, C5, C51, C53, G21, G2

DOI: 10.33429/Cjas.15124.1/8

1. Introduction

The banking sector plays an important role in the economy of every country, and everyone wants an easy channel to carry out every needed banking transactions. To offer a simple way to do various banking operations without having to go to a bank, the banking sector introduced different channels of transactions which include the automated teller machine (ATM), point of sale (POS), internet banking (IB), and unstructured supplementary service data (USSD). The ATM was introduced to the Nigerian market in 1989 (Odusina, 2014). The Central Bank of Nigeria (CBN) 2008 Economic Report stated that ATM and electronic payment had an 87% and 90.80%

[1]The authors are staff of the Department of Statistics, School of Physical Sciences, Federal University of Technology Owerri, Imo State, Nigeria. Views expressed on this paper are those of the authors and do not in any way represent the position of the institution where they are employed or that of the Central Bank of Nigeria

[2]desmond.bartholomew@futo.edu.ng

share of transactions respectively (Agbada, 2008). Although ATMs are often patronized, Odusina (2014) found that customers are discouraged due to long queues. In response to this challenge, additional transaction channels such as online banking, USSD, and POS were introduced, creating alternatives that collectively alleviate the burdens associated with traditional ATM queues. These diverse channels have significantly improved the overall efficiency of the payment system. In recent times, the CBN has introduced innovative policy measures, including the cashless policy to improve the payment system landscape with a view to reducing the cost of cash management and promote financial inclusion.

Thus, the cashless policy has made the POS and online banking the preferred channels of transaction (Chukindi, 2023). This had led to reduced deposits to the banks as people are not sure of getting the cash back when needed, leading to less savings which may affect investment and the economy adversely. The surge of transactions through POS and online banking channels and the associated difficulties may be an indication that their developments are not in pace with the growing population. Variability in population undoubtedly affects savings, investments, and the overall economic status of a nation. Kuroda (2018) noted that aging and shrinking populations will affect savings in different ways, with implication for transaction preferences. Other factors such as gender, marital status, religion, income, age, and education levels also have impacts on the adoption of e-banking in Nigeria (Izogo *et al.* 2012). Uzowulu, Anyanwu, & Amakor (2024) analyzed Nigeria's cashless policy's impact on economic growth, using data on ATM, POS, mobile banking, and web transactions with nominal GDP. Key findings indicated that ATM transactions positively but insignificantly impact growth, POS transactions have a significant negative impact, and both web pay and mobile transactions are negatively but insignificantly related to growth. Recommendations included improving ATM infrastructure and strategically placing POS agents to support economic activities.

Bernardo *et al.* (2023) proposed an extension of a model that explains bank profitability based on indicators of operating efficiency and profitability with the incorporation of independent banking access channels. According to them, these channels are independent variables that affect the performance of banks. They found out that, while ATM had a positive but weak relationship with return on asset (ROA), POS had a negative but very weak relationship with ROA.

The rapid growth of Nigeria's population, alongside advancements in banking transaction channels, raises critical questions about the adequacy of current infrastructure to meet future demands. Although it is known that population growth correlates with increased banking customers, few studies have investigated the specific impact of population dynamics on different banking transaction channels. This study seeks to bridge this gap by examining how Nigeria's total population influences the adoption and utilization of four major transaction channels: ATM, POS, internet banking, and USSD.

This research is motivated by two key gaps: (1) a lack of empirical evidence on the relationship between population growth and transaction channel use in Nigeria, and (2) the underutilization of advanced predictive models, such as Support Vector Regression (SVR) and grey-box time series models, in forecasting transaction volumes based on population projections. The study contributes to knowledge by offering an in-depth analysis of this relationship and presenting predictive insights on which channels may dominate in response to population changes.

It is already known that both population and channels of transactions play a role in the growth of the economy (Jajri & Ismail, 2010; Zulu & Banda, 2015; Njoku *et al.*, 2020). Therefore, this study examines the effect of the growing Nigeria's total population on the four selected transaction channels; to identify an appropriate model for the relationship between each of these channels and Nigeria's total population; and to predict the number of transactions for each channel based on the projected population.

The results from this study is beneficial as they guide policymakers and banks on the expected volume of transactions in the future to better expand their capacity for efficient and seamless banking experience for their customers.

The remaining part of this paper is organized as follows: Section 2 is the review of related literature, while Section 3 describes the data and methodology. Section 4 discusses the results, while Section 5 wraps up the study and gives policy recommendations.

2. Literature Review

2.1 Theoretical Literature

According to Maudos & Fernandez de Guevara (2007), economies of scale are one of the most popular theoretical connections between population size and the financial system. Greater accessibility for the populace, cheaper transaction costs, and more varied financial services can all result from a larger banking sector, which can be supported by a larger population. As the population grows, diverse financial services are needed to cater for different income groups, leading to the development of various channels of financial transactions (Demirguc-Kunt & Klapper, 2013). Beck *et al.* (2007) argued that urbanization, often accompanied by larger populations, can alter financial transaction dynamics, leading to more advanced banking infrastructure and electronic payment systems. Also, advancements in technology can significantly impact financial transactions, with a larger population providing larger market for adopting new financial technologies (Allen *et al.*, 2011). Even the regulatory framework in a country may influence the development of banking channels. In larger populations, regulators may need to adapt their policies to ensure stability and consumer protection (Claessens *et al.*, 2001).

From the extant literature, the banking system may be significantly impacted by population size. Larger population typically supports broader and more varied financial services, which could result in economies of scale, greater financial inclusion, and adjustments to improve competition and regulation in the banking industry.

2.2 Empirical Literature

The relationship between population size and banking transaction volumes has been a subject of investigation in many economies, particularly in the context of financial inclusion and the adaptation of banking services to meet rising demands. Nigeria's increasing population has posed unique challenges for the banking sector, particularly in the adequacy of transaction channels to support seamless, high-volume transactions (Izogo *et al.*, 2012). Empirical studies in this domain often explore how transaction channels like ATMs, Point-of-Sale (POS), internet banking, and Unstructured Supplementary Service Data (USSD) respond to demographic and economic factors, including population dynamics.

Over the years researchers have considered a wide range of models to study the different channels of transactions and the impact they have on different sectors. Ighomereho *et al.* (2018) examined different socio-demographic factors to determine the usage of ATMs in Lagos State, Nigeria using the general linear model. Their study showed that income, age, and education level influence ATM usage while occupation and gender have no significant impact. Research has demonstrated that population growth directly influences the demand for banking services. According to Jajri and Ismail (2010), demographic factors, such as a nation's population size and structure, significantly affect savings and transaction volumes. In Nigeria, population-driven demand has led to an increase in the number of ATMs, POS terminals, and mobile banking channels, though studies indicate varying levels of effectiveness across these channels (Njoku *et al.*, 2020). Kuroda (2018) noted that as populations grow, traditional banking infrastructure often struggles to keep pace, resulting in longer transaction times and customer dissatisfaction.

Several studies have highlighted the importance of transaction channel diversity in response to rising population sizes. Agbada (2008) observed that the introduction of the ATM network and subsequent channels in Nigeria, including POS and USSD, were significant strides in accommodating Nigeria's expanding population. However, studies like Odusina (2014) have reported that these channels are often unable to keep pace with population-driven transaction volumes, particularly during peak demand periods, leading to service inefficiencies. In recent years, the Central Bank of Nigeria's (CBN) policies promoting cashless transactions have made POS and USSD dominant channels, as they offer increased accessibility, particularly for underserved populations (Chukindi, 2023).

In terms of forecasting techniques, Support Vector Regression (SVR) and grey-box models have been employed to model various economic and demographic phenomena due to their adaptability and predictive power. SVR is widely used for its capability to handle non-linear relationships in data, a frequent occurrence in population-transaction relationships due to numerous confounding variables like income level, literacy rate, and urbanization (Bernardo *et al.*, 2023). In the Nigerian context, SVR has been applied in forecasting complex datasets with non-linear patterns, such as transaction volumes and economic indicators, demonstrating high predictive accuracy (Zulu & Banda, 2015).

Grey-box modeling, on the other hand, integrates both system knowledge and empirical data, allowing for more interpretable models compared to purely data-driven approaches like SVR. Grey-box models are particularly valuable when some underlying mechanisms of the system (such as transaction behavior) are partially understood but affected by unknown external variables. This is to say that this model recognizes the effects of confounding or mediating variables that could interface the relationship between population growth and volume of transaction, and controls for such effects.

Research by Bernardo *et al.* (2023) extended a profitability model for banks by incorporating grey-box approaches to include independent transaction channels, showing that these models could capture underlying structures that improve interpretability and forecasting accuracy. Studies suggest that grey-box models perform well in economic forecasting where demographic factors and other dynamic variables, like policy changes, influence the data trends (Njoku *et al.*, 2020).

Despite numerous studies on banking transaction channels, there remains limited research explicitly addressing the impact of Nigeria's population growth on each channel's transaction volume. Studies typically focus on overall electronic banking adoption or general transaction trends without delving into the differential effects on each channel (Odusina, 2014; Izogo *et al.*, 2012). Furthermore, while empirical studies have tested predictive models on isolated channels, there is limited application of SVR and grey-box models in a comparative context, particularly for multi-channel forecasting based on population trends. These gaps in literature is handled in this paper.

3. Data and Methodology

This Section discusses the data used for this research, the model specification, and estimation techniques.

3.1 Data

The data were collected from two different sources: the volume of transactions was obtained from the CBN database (www.cbn.gov.ng/Paymentsystem/ePaymentStatistics .asp), while total population data were obtained from the World Development Indicators (www.statista.com), from 2012 to 2021. The projected Nigeria total population from 2022 to 2027 used for the forecasts was also obtained from www.statista.com.

3.2 Model Specification

The SVR is a non-parametric method used when the dependent variable is numerical rather than categorical. Unlike simple linear regression, whose outputs are based on Gauss-Markov assumptions, SVR's output model is independent of the distributions of the underlying dependent and independent variables. Also, SVR is dependent on kernel functions that build a non-linear model without modifying the explanatory variables. The kernel functions often used are linear, polynomial, sigmoid, and radial basis. The user must select the right kernel function when implementing the SVR method. The selection of kernel functions is difficult and necessitates the use of optimization techniques for the best results. A radius basis function (RBF) kernel which converts data from non-linear to linear space was used in this paper as selected by the optimization technique. According to Obite *et al.* (2021), the SVR model is expressed as:

$$Y_i = WK(x_i, x) + b \tag{1}$$

where W and b are the regression parameters and $K(x_i, x)$ is the kernel function.

The SVR technique utilizes a supervised machine learning approach to find an optimal hyperplane as described by Chang and Lin (2011). It handles nonlinearity and flexibility in defining kernel functions. The technique can be employed to classify and predict transaction patterns based on population size. The model has been successfully applied in making accurate predictions of stock market prices (Guo & Guo, 2010), wind speed using hybrid features selection (Adhikari *et al.* 2017), and short-term solar power for optimal energy management (Zeng & Qiao, 2013).

On the other hand, the Grey-Box model combines the advantages of the White-Box and Black-Box models thereby including a system of stochastic differential equations

(SDEs). The White-Box captures the inherent linear and nonlinear behavior of the system under investigation (Sjöberg, 2000). The Black Box uses experimental data and requires input and output variables, as well as their correlation. It accommodates the process noise caused by approximation error and the measurement noise caused by imperfect measurement.

The Grey-Box model combines the advantages of the White-Box and Black-Box by allowing prior physical knowledge and parameter estimation using statistical methods. Due to its natural framework, the Grey-Box model is suitable for use in modeling a dynamic system (Bohlin & Graebe, 1995). The Grey-Box model was further defined as a collection of partially discrete, nonlinear observed SDEs with measurement and process noise. The Grey-Box is the combination of theory-based structures and data-based components (Sohlberg & Jacobese, 2008). The Grey-Box model is given in Equation (2):

$$dX_t = f(X_t, u_t, t, \theta)dt + \sigma(X_t, u_t, t, \theta)d\omega_t \tag{2}$$

$$Y_k = h(X_k, u_k, t_k, \theta) + e_k \tag{3}$$

where X_t denote state variables (Nigerian Total Population, which influences future states based on past behaviors and external factors); u_t are Input variables (lag and lead values of the Nigerian Total Population); θ are parameter (possibly unknown); Y_k are Output variables (Volume of transaction for the kth transaction channel over time); t is Time (Year of transaction); ω_t are Standard Wiener processes; and e_k are White noise with $N(0, S(u_k, t_k, \theta))$.

In a Grey-Box model, the output variable is the primary variable being predicted or observed as the model's outcome, based on the internal state variables and any inputs. It represents the measurable response of the system that the model aims to simulate or forecast. The state variable helps bridge known dynamics (e.g., seasonal trends, speed of innovation, the number of inoperative ATMs, the purchase of cash via POS, payment operator's diversity, among others) with the empirical observations to create a robust model that balances theoretical knowledge with real-world data. It is worth noting that the drift term and diffusion term, respectively, are the first and second terms on the right-hand side of Equation (2) and are referred to as the White-Box. Equation (3) accommodates the noise from the data (it introduces the system of SDEs). This is referred to as the Black-Box. Therefore, a combination of the prior knowledge and the information from the data gives the Grey-Box Model.

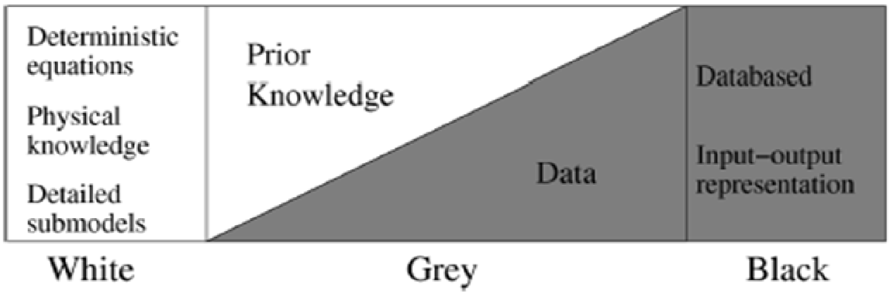

Figure 1: Concept of Grey-Box Model

Source: This picture was first uploaded by Signe Schmidt

The Grey-box modeling is a hybrid modeling technique that combines the strengths of black-box and white-box models to capture complex system dynamics, while maintaining interpretability (Box & Jenkins, 1976). The model is useful for systems with unclear underlying structures or when there is limited data availability as evident in this study. Several studies have utilized the Grey-box technique in cash flow (Panga *et al.* 2015), traffic flow (Chen & Li, 2019), and energy consumption and demand (Kumar & Singh, 2020) to make accurate predictions. The Grey box model and the SVR are effective in handling small sample size and controls the effect of mediating variables, hence, their preference for modeling in this study (Kramer *et al.*, 2009; Akyuz & Bilgil, 2022).

3.3 Estimation Procedure

This study uses Grey-Box and Support Vector Regression (SVR) for modeling the four channels of transactions namely: ATM, POS, Online, and USSD. Three performance measures namely: R-Square, MAPE, and RMSE were used to explain how well the independent variable explains each of the dependent variables. The model with the highest R-square and least MAPE and RMSE is selected as the best model and used to forecast for the four different channels of transaction separately.

We consider four different channels of transactions. Each channel for a given year is explained by the total population for the year and the value two years ago (lag), or the value two years ahead(lead). This means that we need to produce lags and leads for each channels of transaction making this study a multivariate case. The lags represent the values from years ago, while the leads represent the values for years to come. These lags and leads will be created using the xregExpander() function in the R programming language. However, moving several observations of the original data ahead or backward results in missing values at the start or end of the data, R fills these values using estimates from the es() and iss() functions that come in the smooth package.

The step-wise method will be implemented for the selection of the regressors based on information criteria and partial correlations. The most significant variable(s) from the group of input and state variables (the Nigeria total population, the two lags, and the two leads) will be selected as follows:

i. An information criterion is calculated using the regression model of the volume of transactions on the total population and constant term for each channel.

ii. The model in (i) is used to predict the response variable and the correlations of the resulting residuals with all the input variables (lags and leads).

iii. The most significantly correlated variables with the residuals are added to the model in (i) using lm() function.

iv. An information criterion is computed using the regression model in (iii) and the value is compared with the one from (i). The process is stopped, if the value is greater or equal to the previous one else the process is repeated from step (ii).

At each of the four steps illustrated above, the partial correlations are used to ensure that only variable(s) contributing to the model is(are) included and the resulting model will have the lowest information criterion. In conclusion, the analysis fits the

Grey-Box model to the volume of transaction from the selected channels of transaction using the total population size as state variable, its lead and lag values as input variables; then fits the tuned SVR model to the volume of transaction from the selected channels of transaction using the same state and input variables. The predictive powers of the two models (Grey-Box and SVR) is compared using the R-square, MAPE, and RMSE for each channel of transaction; and plots of the forecasted volume of transactions for each channel for 2022 to 2027 are provided.

4. Results and Discussions

4.1 Descriptive Statistics

The descriptive statistics are presented in Table 1. The descriptive statistics reveals that, on average, and with the biggest standard deviation across the ten-year study period, the online transaction channel prevailed. The distribution of all the variables is not symmetrical since all of the skewness values are outside the range of -0.5 and 0.5. Positive kurtosis for all the variables indicates a distribution that is more peaked than the normal distribution.

Table 1: Descriptive Statistics

Statistic	ATM	POS	ONLINE	USSD	Population
Mean	717.864	413.743	1,397.039	149.500	191.376
Standard Error	123.812	263.921	1,047.831	60.130	4.600
Median	695.394	104.991	21.540	47.429	191.081
Standard Deviation	391.528	834.595	3,313.535	190.150	14.549
Kurtosis	1.89	8.97	7.23	0.83	-1.19
Skewness	1.24	2.95	2.67	1.40	0.05
Range	1,303.771	2,740.968	10,319.303	550.613	43.325
Minimum	295.417	2.588	2.276	2.298	170.075
Maximum	1,599.187	2,743.556	10,321.580	552.912	213.401
Sum	7,178.643	4,137.433	13,970.391	1,494.997	1,913.761
No. of Obs.	10	10	10	10	10

Note: All figures are in Billion Naira

The volume of transactions from the four channels of transactions are shown in Figure 2. Figure 2 shows that the volume of transactions from online (mobile transfers) transactions, POS, and USSD were relatively the same from 2012 to 2019, after which a surge was witnessed in online transactions in 2020 and 2021 which can be attributed to COVID-19 lockdown and social distancing.

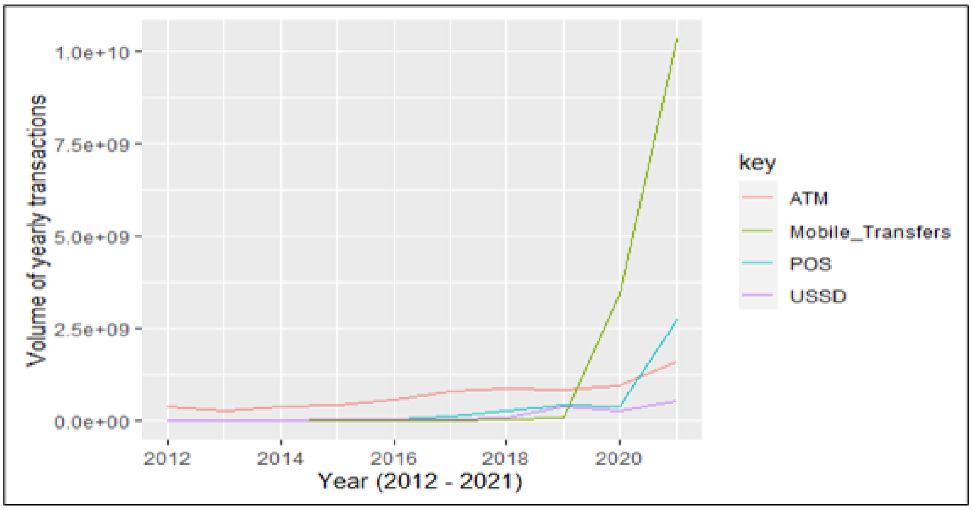

Figure 2: Volume of transactions in Nigerian Banks from 2012 - 2021

During the lockdown, people were forced to rely mostly on the internet, POS, and USSD for their transactions. Interestingly, ATM transactions maintained a higher volume above other channels of transaction between 2012 and 2019 but was overtaken by online transactions from 2019. Other transaction channels except USSD witnessed surge in patronage above the ATM transactions. Having shown and described the study data in Section 4.1, we begin the Grey-Box and SVR modeling of the different channels of transactions in Section 4.2.

4.2 Estimation Results

In Table 2, a simple linear regression of transaction channels on population size were conducted to determine the nature of the relationship between the variables.

Table 2: Simple Regression Analysis of Transaction Channel on Population

Channel	R-Squared	F-Statistics	Functional Form
ATM	0.8395	41.85	Linear
USSD	0.7289	21.51	Linear
POS	0.4542	6.658	Other
ONLINE	0.4373	6.217	Other

The high R-squared values (0.8395 and 0.7289) for ATM and USSD respectively indicate a strong explanatory power of the model and their high F-stat values (41.85 and 21.51) also reinforce the significance of these linear models. The linear functional form suggests that the relationship between these transaction channels and population size is straightforward and proportional.

On the other hand, the moderate R-squared values (0.4542 and 0.4373) for POS and Online transactions respectively indicate a weaker explanatory power and the moderate F-stat values (6.658 and 6.217) confirm that the models are significant but less robust indicating a non-linear relationship. The non-linear functional form suggests that the relationship between POS and Online transactions and population size is more complex and cannot be explored using the linear model. This could mean that changes in population size lead to non-proportional changes in POS and Online transactions, potentially requiring a different model (e.g., polynomial, logarithmic, or machine learning models as used in this study) to accurately describe the relationship.

4.2.1 Grey-Box result for ATM transactions

The results of the Grey-Box model are reported in Table 3. The intercept represents the baseline value of ATM transaction volume when all other variables (population, lagged and lead values of population size) are zero.

Table 3: Summary of the Grey-Box model for the ATM transactions

Coefficients	Estimate	Std. Error	Lower 2.5%	Upper 97.5%
Intercept	-4.0e+09	7.8e+08	-5.8e+09	-2.2e+09*
Population	2.5e+01	4.1e+00	1.5e+01	3.4e+01*
Error standard deviation = 177853476; n = 10; npar = 3; df = 7				
Information criteria: AIC = 410.7414; AICc = 414.7414; BIC = 411.6492; BICc = 416.2544				

Note: df denotes degree of freedom, npar is number of parameters, and n is sample size

Here, it has a large negative value, suggesting that without accounting for population size, the model predicts a highly negative baseline transaction volume. In practical terms, since a negative transaction volume is not realistic, this large negative intercept implies that population size and its lagged/lead effects are crucial to producing realistic transaction volumes. A significant negative intercept indicates that for ATM transaction volumes to be positive and meaningful, a certain minimum level of population (or population-related effects like lag/lead) is necessary. It suggests that ATM usage scales up significantly as population factors are accounted for in the model.

The population parameter indicates that for each unit increase in population size, the ATM transaction volume is predicted to increase by approximately 24.66 units, holding other factors constant. The positive and significant population coefficient confirms that as the population grows, ATM transactions are expected to rise. This strong positive association highlights the demand for ATM services as population size increases, aligning with the model's goal of assessing how population affects transaction volume. The two lags and leads were not correlated with the ATM volume of transactions and were not included in the model. We proceed with the SVR model for ATM volume of transactions in Section 4.2.2.

4.2.2 The SVR Result for ATM transactions

The support vector regression was tuned by trying different values of maximum allowable error and cost parameters. A total of 1,100 SVR models were fitted iteratively using the *tune*() function of the e1071 library and the best was selected. The summary of the best model is shown in Table 4.

The SVR model uses the RBF kernel with carefully chosen parameters (cost, gamma, epsilon) to capture the non-linear relationship between population and ATM transactions. The model was thoroughly validated by training 1,100 models, ensuring robust parameter selection. The use of 6 support vectors indicates an efficient and effective model. The chosen parameters collectively provide a balance between model complexity and predictive accuracy, making this model well-suited for predicting ATM transactions based on population data.

Table 4: SVR Model summary for Effect of Population on ATM transactions

Parameter	Value
Type	eps regression
Kernel	Radial
Cost	3
Gamma	1
Epsilon	0.1
No. of Support vectors	6
No. of models trained	1,100
Weight	6.896
Rho	-0.73

4.3.1 Grey-Box Result for POS

As shown in Table 4, the total population and its first lead were significant in predicting volume of POS transactions. The intercept represents the baseline POS transaction volume when the population and its lead and lag values are zero. Here, it has a large positive value but is not statistically significant (confidence interval includes zero), suggesting that it may not have a strong effect on the model by itself.

Table 5: Summary of the Grey-Box model for the POS transactions

Coefficients	Estimate	Std. Error	Lower 2.5%	Upper 97.5%
Intercept	1.0e+10	5.5e+09	-3.4e+09	2.4e+10
Population	2.6e+03	7.8e+02	7.1e+02	4.5e+03*
Population Lead	-2.5e+03	7.6e+02	-4.4e+03	-6.7e+02*
Error standard deviation =448815302; n = 10; npar = 4; df = 6				
Information criteria: AIC = 429.7130; AICc = 437.7130 ; BIC = 430.9233 ; BICc = 440.1336				

Note: df denotes degree of freedom, npar is number of parameters, and n is sample size

Since this intercept is not significant, the focus should be on the population and lead values as they have a stronger predictive influence. The population coefficient indicates that for each additional person in the population, the POS transaction volume is expected to increase by approximately 2,599 units, all else equal. This coefficient is positive and significant. The positive coefficient for population size indicates that POS transaction volume increases directly with population growth, suggesting that demand for POS transactions scales with the overall population. The population lead coefficient indicates that for each additional unit in the one-step lead population size (i.e., future population), the POS transaction volume is expected to decrease by approximately 2,522 units, all else equal. This coefficient is negative and significant. This negative relationship suggests that projected future increases in population could reduce current POS transaction volumes. This might indicate a shift in transaction preferences or possible infrastructural constraints in POS usage as population grows, prompting policymakers to prepare for higher demand or alternative transaction methods in future.

4.3.2 SVR Results for POS transactions

The support vector regression model for POS transaction volume was tuned by varying the values of the maximum allowable error and cost parameters. A total of 1,100 SVR models were fitted iteratively using the tune() function and the best was selected. The model summary is given in Table 6.

Table 6: SVR Model Summary for Effect of Population on POS transactions

Parameter	Value
Type	eps regression
Kernel	Radial
Cost	1
Gamma	1
Epsilon	0.1
No. of Support vectors	5
No. of models trained	1,100
Weight	6.898
Rho	0.136

The SVR model in Table 6 uses an RBF kernel with carefully selected parameters. The moderate cost and gamma values helped balance the complexity and generalization of the model. The epsilon value provided a margin for minor errors, enhancing the model's robustness. The use of 5 support vectors ensures efficiency and simplicity, while extensive model training ensures optimal parameter selection. Overall, this model is well-suited to capture the potentially complex relationship between population dynamics and POS transactions.

4.4.1 Grey-Box Results for Online Transactions

In Table 7, the intercept is the baseline value of online transaction volume when population size and its lead values are zero. Here, the intercept is significant, suggesting it contributes substantially to the prediction of online transaction volume. The significant positive intercept indicates a foundational level of online transaction volume independent of current population growth, likely due to existing baseline demand for online transactions.

The population coefficient indicates that for each additional unit increase in the population, online transaction volume is expected to increase by approximately 12,064 units, holding other factors constant. This coefficient is positive and statistically significant. The positive relationship shows that as the population grows, online transactions are likely to increase, emphasizing the importance of expanding online transaction infrastructure to accommodate rising demand.

Table 7: Summary of the Grey-Box model for volume of online transactions

Coefficients	Estimate	Std. Error	Lower 2.5%	Upper 97.5%
Intercept	5.2e+10	1.4e+10	1.8e+10	1.7e+04*
Population	1.2e+04	1.9e+03	7.3e+03	4.5e+03*
Population Lead 2	-1.2e+04	1.9e+03	-1.6e+04	-7.0e+03*
Error standard deviation =1131093182; n = 10; npar = 4; df = 6				
Information criteria: AIC = 448.1995; AICc = 456.1995 ; BIC = 449.4099 ; BICc = 458.6202				

Note: df denotes degree of freedom, npar is number of parameters, and n is sample size

For each additional unit increase in the future population (two steps ahead), the online transaction volume is expected to decrease by approximately 11,736 units, all else equal. This coefficient is negative and statistically significant. The negative relationship with future population could suggest that anticipated future population growth might shift demand away from online transactions, possibly due to changes in transaction preferences or limitations in online transaction infrastructure as population increases. This may indicate a need for improvements or alternatives in online services to sustain demand.

4.4.2 SVR Results for Online Transactions

The support vector regression was tuned by varying the values of the maximum allowable error and cost parameters. A total of 1,100 SVR models were fitted iteratively using the tune() and the summary of the best-selected model is shown in Table 8. The results show that, a unit increase in the total population significantly increases the number of online transactions by 7.5 units.

Table 8: SVR Model summary for Effect of Population on Online transactions

Parameter	Value
Type	eps regression
Kernel	Radial
Cost	100
Gamma	1
Epsilon	0.1
No. of Support vectors	7
No. of models trained	1,100
Weight	7.504
Rho	-0.850

The SVR model for the effect of population on online transactions in Table 8 uses an RBF kernel with carefully selected parameters to handle non-linear relationships. The high-cost parameter emphasizes fitting the training data closely, while the gamma and epsilon values help balance the model's complexity and robustness. The use of 7 support vectors ensures the model captures the critical data points efficiently. Extensive training ensures optimal parameter selection, resulting in a well-suited model for predicting online transactions based on population data.

4.5.1 Grey-Box Modeling of the USSD Transactions

The intercept represents the expected baseline of USSD transaction volume when the population is zero. In this case, the intercept is negative and statistically significant, implying that if population size were theoretically zero, USSD transaction volume would be negative. While the intercept itself is not practically meaningful in real terms (since population cannot be zero), its negative significance suggests that USSD transactions are highly dependent on population, with low baseline usage independent of population size.

Table 9: Summary of the Grey-Box model for volume of USSD transactions transactions

Coefficients	Estimate	Std. Error	Lower 2.5%	Upper 97.5%
Intercept	-2.0e+09	4.9e+08	-3.2e+09	-8.2e+08*
Population	1.1e+01	2.6e+00	5.1e+00	1.7e+01*

Error standard deviation =112254781; n = 10; npar = 3; df = 7
Information criteria: AIC = 401.5377; AICc = 405.5377 ; BIC = 402.4454 ;
BICc = 407.0506

Note: df denotes degree of freedom, npar is number of parameters, and n is sample size

The populations coefficient indicates that for each unit increase in population, the USSD transaction volume is expected to increase by approximately 11.16 units, holding other factors constant. This coefficient is positive and statistically significant. The positive relationship indicates that as the population grows, the volume of USSD transactions increases. This implies a growing reliance on USSD services with increasing population, emphasizing the need for robust infrastructure to support this demand.

4.5.2 SVR Results for USSD Transactions

The support vector regression was tuned by varying the values of the maximum allowable error and cost parameters. A total of 1,100 SVR models were fitted iteratively using the *tune*() and the best was selected. The results show, a unit increase in the total population increases the number of USSD transactions by 2.49 units.

Table 10: SVR Model summary for Effect of Population on Online transactions transactions

Parameter	Value
Type	eps regression
Kernel	Radial
Cost	100
Gamma	1
Epsilon	0.1
No. of Support vectors	6
No. of models trained	1,100
Weight	2.496
Rho	0.027

The SVR model analyzing the effect of population on online transactions in Table 10 employs an RBF kernel to manage non-linear relationships. The high-cost parameter ensures a close fit to the training data, while the gamma and epsilon values strike a balance between model complexity and robustness. Using 6 support vectors allows the model to efficiently capture essential data points. Extensive training and cross-validation guarantee the selection of optimal parameters, making this model highly effective for predicting online transactions based on population data.

4.6 The Impact of Population Dynamics on the Four Transaction Channels

The impact of population dynamics on the four transaction channels: ATM, POS, Online, and USSD reveals differences in responsiveness and adaptability. Starting with ATM transactions, the population coefficient is positive and significant, indicating that increases in population size correlate with increased ATM usage. However, the rate of increase per unit of population size is moderate, suggesting that while ATM usage grows with population, the growth rate is stable and not as pronounced as in digital transaction channels. This shows ATM transactions continue to play an enduring role, although their responsiveness is more modest compared to digital options.

POS transactions show a stronger relationship with population changes, as seen by the large positive effect of the population coefficient. However, the lead coefficient for population is negative and significant, suggesting that future expectations of population growth might actually reduce immediate growth in POS usage. This likely points to possible saturation or infrastructure limitations in POS transactions, indicating that sustained population growth may outpace current POS capacity.

Online transactions demonstrate a notably strong positive response to population growth, evidenced by a highly significant population coefficient. This implies a strong and scalable growth in online transaction usage as the population increases. However, like POS transactions, online channels also show a negative lead effect, suggesting potential challenges in sustaining growth to meet future demand unless infrastructure is expanded.

USSD transactions exhibit a positive and significant relationship with population growth, with a moderate but steady increase in usage as the population grows. This steady, stable growth highlights USSD's accessibility and adaptability, particularly for regions with limited internet access, making it well-suited to handle increased demand without signs of saturation.

Overall, online transactions show the strongest response to population growth, followed by POS, then USSD, and lastly ATM transactions. Signs of saturation in POS and online channels suggest a need to prioritize scaling their infrastructure to meet anticipated future demand. The consistent growth in USSD transactions highlights its accessibility and adaptability, while ATM usage, though positively correlated with population growth, may require less immediate investment compared to rapidly expanding digital channels. These findings underscore the importance of focusing on digital infrastructure, particularly for POS and online transactions, to support Nigeria's growing population.

4.7 Comparison of Grey-Box and SVR Results each Transaction Channel

Table 11 provides a comparison of the performance metrics of the Grey-Box and Support Vector Regression (SVR) models across four different transaction channels: ATM, POS, online, and USSD transactions.

Table 11: Comparison of Grey-Box and SVR Results

Model	R-squared	MAPE	RSME
ATM Transactions			
Grey-box	0.8395082	15.65102	14880289
SVR	0.9927389	5.098892	33565159
POS Transactions			
Grey-box	0.8072049	1786.744	347650838
SVR	0.6504067	2254.650	592014068
Online transactions			
Grey-box	0.9223177	12098.030	876141011
SVR	0.9920878	3142.389	284225751
USSD transactions			
Grey-box	0.7289348	500.8410	93919088
SVR	0.8891646	108.5039	82618795

For ATM transactions, the SVR model outperforms the Grey-Box model, achieving an R-squared value of 0.9927 compared to 0.8395, along with a significantly lower Mean Absolute Percentage Error (MAPE) of 5.10% versus 15.65%. However, the Root Mean Square Error (RMSE) for the SVR model is higher at 33,565,159 compared to the Grey-Box model's RMSE of 14,880,289.

In the case of POS transactions, the Grey-Box model has a higher R-squared value of 0.8072, while the SVR model shows a lower value of 0.6504. The MAPE for the Grey-Box model is notably high at 1786.74, compared to the SVR's MAPE of 2254.65. The RMSE also reflects this discrepancy, with the Grey-Box model at 347,650,838 and the SVR model at 592,014,068. For online transactions, the SVR model again shows superior performance with an R-squared value of 0.9921, compared to the Grey-Box's 0.9223. The MAPE for the Grey-Box model is higher at 12,098.03, while the SVR model has a MAPE of 3,142.39. The RMSE for the Grey-Box model is significantly higher at 876,141,011, compared to the SVR model's RMSE of 284,225,751.

Finally, for USSD transactions, the Grey-Box model displays a lower R-squared value of 0.7289 versus 0.8892 for the SVR model. The MAPE indicates better performance for the SVR model at 108.50, compared to the Grey-Box model's 500.84. The RMSE results show that the SVR model has a lower RMSE of 82,618,795, while the Grey-Box model's RMSE is 93,919,088.

In summary, the SVR model generally demonstrates superior predictive capabilities for ATM, online transactions, and USSD while the Grey-Box model performs better for POS transactions, particularly in terms of R-squared and MAPE metrics.

4.8 Forecast Results

The forecast volumes of transactions are presented in Figures 3-6. Figure 3 presents the actual and forecasted ATM transaction volumes based on both Grey-box and SVR model estimates. According to Table 11, the SVR model demonstrates superior predictive accuracy for ATM transactions. The SVR forecast indicates a peak in ATM transaction volumes from February to April 2023, followed by a gradual decline projected from 2024 through 2027.

Figure 4 displays the actual and forecasted POS transaction volumes using both Grey-box and SVR models. As indicated in Table 11, the Grey-box model provides the best predictive accuracy for POS transactions. Its forecast suggests a steady increase in POS transaction volumes, projected to continue through 2027.

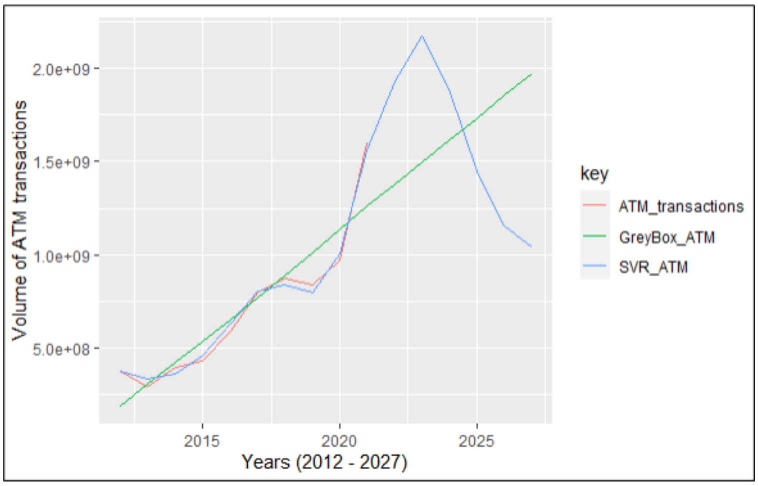

Figure 3: Forecast of volume of ATM Transactions from 2022 - 2027

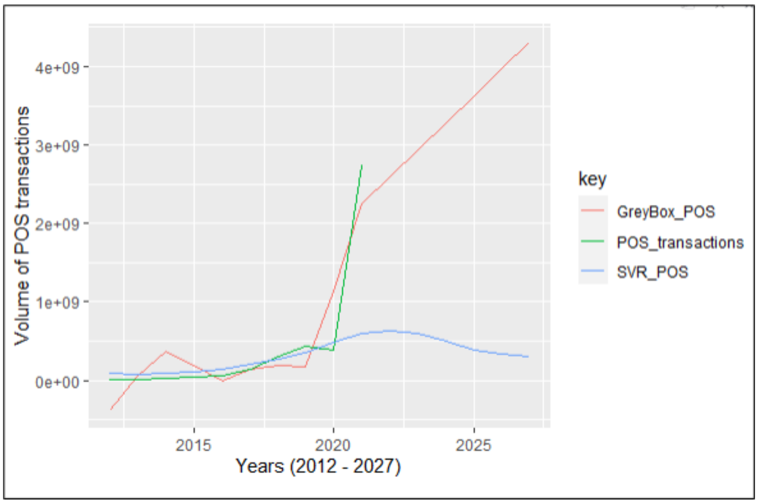

Figure 4: Forecast of volume of POS Transactions from 2022 - 2027

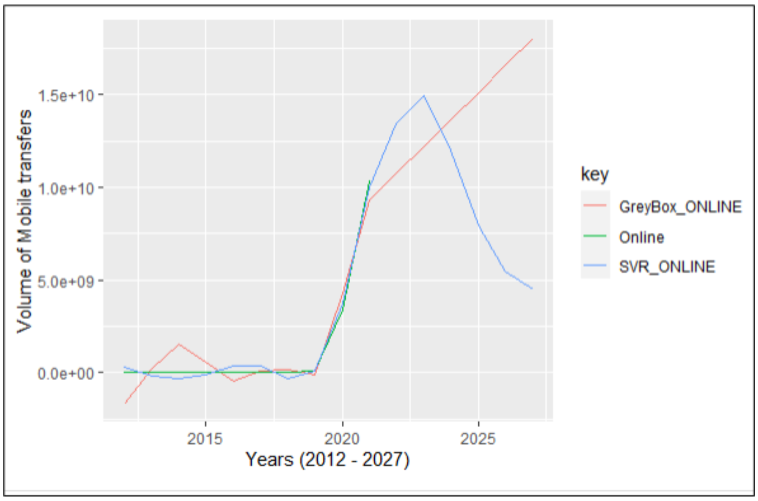

Figure 5: Forecast of volume of Mobile Transfer Transactions from 2022 – 2027

Figure 5 illustrates the actual and forecasted values of online transaction volumes based on both Grey-box and SVR models. According to Table 11, the SVR model provides the most accurate forecasts for online transactions, projecting a decline in transaction volume starting from April 2023 and persisting through 2027. This trend may reflect challenges impacting internet banking in Nigeria, including network issues, internet security concerns, and fraud.

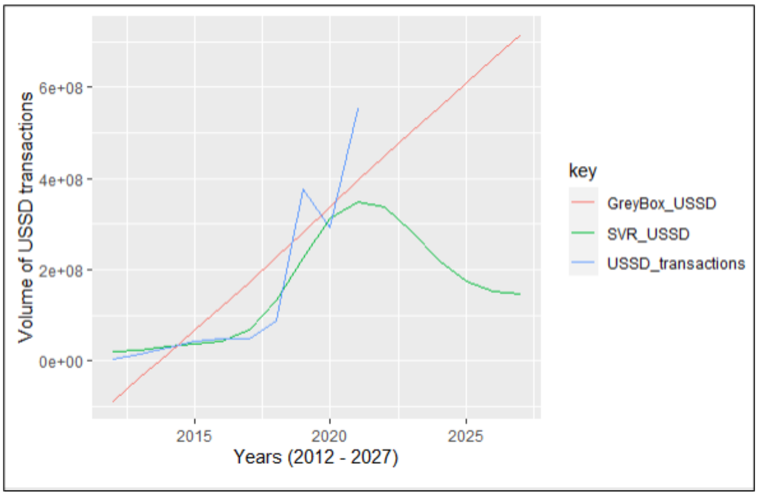

Figure 6: Forecast of volume of USSD Transactions from 2022 - 2027

Figure 6 presents the actual and forecasted values for USSD transaction volumes, as estimated by both Grey-box and SVR models. Based on the results in Table 11, the SVR model provides the most accurate forecast for USSD transactions, indicating that transaction volumes are lower than those of other channels and display a crest-and-trough trend. This fluctuating pattern may be attributed to various challenges affecting USSD banking in Nigeria, including network issues, occasional unavailability of service codes, and limited customer awareness. The actual and predicted values of the four selected channels of transactions are summarized in Table 10.

Table 12 presents the actual and predicted demand for various transaction channels: ATM, POS, online, and USSD, covering the years 2024 to 2027. The table includes columns for the actual transaction volumes and their corresponding predicted values for each channel, allowing for a comparative analysis of the forecasted trends against historical data.

For the year 2024, the predicted ATM transactions are 1,892,724,143, while no actual volume data is available. In contrast, the predicted POS transactions are 3,278,494,703, with no actual figure provided. The online transactions are forecasted at 11,975,581,741, and, similarly, the actual volume is not listed. Lastly, the predicted USSD transactions for 2024 amount to 218,227,161, again without an actual value available. Moving into 2025, the predicted ATM transactions drop

Table 12: Actual and Predicted Demand for Channels of Transaction (2024 – 2027)

Year	ATM	POS	Online	USSD	Pred_ATM	Pred_POS	Pred_Online	Pred_USSD
2012	375.513	2.588	2.276	2.298	375.103	-377.986	335.337	21.350
2013	295.417	9.418	2.900	15.930	334.567	62.932	-149.872	22.413
2014	400.269	20.817	5.567	27.745	361.249	373.740	-325.512	30.981
2015	433.696	33.721	7.981	43.933	456.723	192.657	-103.493	36.295
2016	590.239	63.715	14.088	47.053	629.392	-5.347	345.349	41.457
2017	800.549	146.267	28.991	47.805	802.929	143.345	361.211	66.745
2018	875.519	295.890	50.816	87.086	836.500	187.642	-281.938	131.335
2019	839.820	438.614	103.497	377.265	800.445	168.677	85.111	225.759
2020	968.433	382.846	3,432.693	292.970	1,007.586	1,142.638	3,763.953	312.022
2021	1,599.187	2,743.556	10,321.580	552.912	1,559.994	2,249.138	9,988.963	347.681
2022					1,933.822	2,592.258	13,435.618	337.343
2023					2,171.638	2,935.375	14,979.644	282.777
2024					1,892.724	3,278.495	11,975.582	218.227
2025					1,448.758	3,621.614	7,953.494	173.497
2026					1,153.778	3,964.734	5,446.066	152.710
2027					1,037.363	4,307.853	4,492.870	146.042

Pred_ATM: Predicted ATM, Pred_POS: Predicted POS, Pred_Online: Predicted Online transactions, Pred_USSD: Predicted USSD to 1,448,757,954, with actual data not provided. The POS predicted volume is 3,621,614,182, with no corresponding actual value. Online transactions are forecasted to decrease significantly to 7,953,494,348, while the USSD transactions are predicted to be 173,497,334.

In 2026, the predictions show a continued decline for ATM transactions, estimated at 1,153,778,104. The POS transactions are predicted at 3,964,733,662, while online transactions further decline to 5,446,065,537. USSD transactions are predicted at 152,710,463. Finally, by 2027, the predicted ATM transactions drop to 1,037,363,296. The POS channel is expected to see an increase in predicted transactions, reaching 4,307,853,142, while online transactions decline to 4,492,869,945. The predicted USSD transactions are forecasted at 146,041,624.

Overall, Table 12 illustrates a declining trend in ATM and online transactions over the forecasted years, while POS transactions are expected to increase. These variations are plotted in Figure 7. The absence of actual transaction data in the later years indicates that the data was not available, which could be essential for validating the predictive models used.

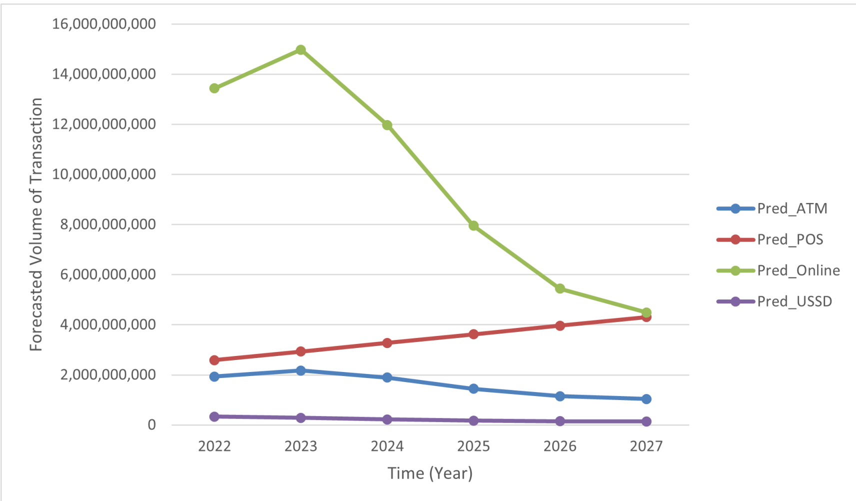

Figure 7: Forecasted Volume of Transaction over Time (Year)

4.9 Discussion of Results

Based on the results, some of the findings are highlighted as follows: the volume of transactions from online transactions, POS, and USSD were relatively the same from 2012 to 2019, after which a surge was witnessed in online transactions from 2020 to 2021. However, ATM transactions maintained a higher volume of transactions than other channels between 2012 and 2019 before online transactions overtook it. The result shows that the relationship between Nigeria's total population and ATM transactions is linear. Further result shows that the linear functional form is a better approximation of the relationship between Nigeria's total population and the volume of USSD transactions, relative to competing functional forms.

The results revealed that the SVR outperformed the Grey-Box model in predicting three out of the four channels of transactions separately because it resulted in the smallest MAPE and RMSE values with a correspondingly higher R-squared value. Hence, the SVR model was used to forecast the future volume of transactions for each of the three channels of transaction from 2022–2027. However, for POS transactions, the Grey-box performed better than the SVR and was used for the forecast. From the plot of the forecast for the volume of transactions, we found that online transactions would have the highest volume of transactions from 2022 – 2027 followed by POS.

According to the findings of this study, ATM transactions will no longer be the channel with the highest volume of transactions in the future, as they were between 2012 and 2019. This finding is consistent with Odusina's position in 2014. According to Chukindi (2023), POS transactions were supposed to be Nigeria's major channel of withdrawing cash owing to the shortage, it has been revealed in this study that POS is the second highest volume of transactions after online transactions, as forecasted by the Grey-box model.

5. Conclusion and Policy Recommendation

We undertook a comprehensive examination of the impact of population size on the channels of bank transactions in Nigerian, employing two modeling approaches: Grey-box and Support Vector Regression (SVR). Our research aimed to shed light on the intricate relationship between population dynamics and the evolution of banking transaction channels in Nigeria. Findings offered valuable insights into this relationship, highlighting the mixed interplay between population size and banking channel preferences. The study concludes that Online transactions will be the dominant channel of banking transactions in Nigeria from 2022 to 2027.

The study contributes to the growing body of knowledge in the field of banking and economics, emphasizing the significance of employing a diversified approach to unravel complex phenomena. We hope that this research catalyzes further investigations into the evolving dynamics of the Nigerian banking sector and, more broadly, for the development of effective policies that align with the ever-changing needs of a dynamic population.

The study recommends the need to improve the payment system infrastructure; promote financial inclusion to enhance access to financial services; and implement policies that encourage the growth of specific transaction channels based on the characteristics of different population sizes.

References

Adhikari, A., Agrawal, R.K. & Majumdar, A. (2017). Support Vector Machine Regression-based Wind Speed Prediction using Hybrid Feature Selection Approach. *Applied Soft Computing, 61,* 425 – 438.

Agbada, A.O. (2008). Electronic Banking in Nigeria: Problems and Problems and Prospects from the Consumer's Perspective. *Bullion, 32*(4), 19-22.

Akyuz, L., & Bilgil, H. (2022). Application of Standard and Exponential Grey Forecasting Models on Turkeys Education Expenditures. *New Trends in Mathematical Science.* https://doi.org/10.20852/ntmsci.2022.484.

Alabar, T.T. (2012). Electronic Banking Services and Customer Satisfaction in the Nigerian Banking Industry. *International Journal of Business Management Tomorrow, 2*(3), 12 - 25.

Allen, F., Carletti, E., & Marquez, R. (2011). Credit Market Competition and Capital Regulation. *Review of Financial Studies, 24*(4), 983-1018.

Beck, T., Demirguc-Kunt, A., & Levine, R. (2007). Finance, Inequality, and Poverty: Cross-Country Evidence. *Journal of Economic Growth, 12*(1), 27-49.

Bernardo H., Gonzalez M., Figueroa E. G. and Torre-Torres O. V. (2023). The Effects of Banking Channels and Efficiency Indicators on Bank Profitability. *Merc. negocios, 24*(48), 3-26.

Bohlin, T., & Graebe, S. F. (1995). Issues in nonlinear stochastic grey-box identification. *International Journal of Adaptive Control and Signal Processing, 9*(6), 465-490. https://doi.org/10.1002/acs.4480090603

Box, G.E.P. & Jenkins, G.M. (1976). Time Series Analysis: Forecasting and Control. Revised Edition Retrieved from https://www.popline.org/node/499376, Holden-Day, San Francisco.

Chang, C.C. & Lin, C.J. (2011). LIBSVMSVR: A Library for Support Vector Machine Regressions. *ACM Transactions on Intelligent Systems and Technology (TIST), 2*(3), 1-27. https://doi.org/10.1145/1961189.1961199

Chen, Q. & Li, X. (2019). Grey-box modeling of Time Series for Traffic Flow Prediction. *IEEE Transactions on Intelligent Transportation Systems, 20*(7), 2395-2402.

Chukindi J. (2023). Naira note scarcity: POS operators dominate the market, devise new strategies. Daily Post. https://dailypost.ng/2023/02/05/naira-note-scarcity-pos-operators-dominate-market-device-new-strategy/

Claessens, S., Demirguc-Kunt, A., & Huizinga, H. (2001). How Does Foreign Entry Affect the Domestic Banking Market? *Journal of Banking & Finance, 25*(5), 891-911.

Demirgüç-Kunt, A., & Klapper, L. (2013). Measuring Financial Inclusion: Explaining Variation in Use of Financial Services across and within Countries. *Brookings Papers on Economic Activity, 44*(1), 279-340.

Guo, Y. & Guo, Y. (2010). Predicting Stock Prices with Support Vector Regression. *Expert Systems with Applications, 37*(12), 7908 -7913.

Ighomereho, S., Ladipo, P., & Dixon-Ogbechi, B. (2018). Determinants of Automated Teller Machine Usage in Lagos State, Nigeria. *Acta Universitatis Sapientiae, Economics and Business, 6*(1), 89–111. https://doi.org/10.1515/eb-2018-0005.

Izogo, E., Ogbuji, C.N., Onuoha, A.O., & Kalu, S.E. (2012). Impact of Demographic Variables on Consumers' Adoption of E-banking in Nigeria: An Empirical Investigation. *European Journal of Business and Management 4*(17), 27-39.

Jajri, I. & Ismail R. (2010). Impact of labour quality on labour productivity and economic growth. *African Journal of Business Management, 4*(4), 486-495.

Kramer, K., Hall, L., Goldgof, D., Remsen, A., & Tong Luo. (2009, August). Fast Support Vector Machines for Continuous Data. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 39*(4), 989–1001. https://doi.org/10.1109/tsmcb.2008.2011645

Kumar, P. & Singh, A. (2020). Greybox Modeling of Energy Consumption for Demand Response Prediction in Smart Grid. *International Journal of Electrical Power and Energy Systems, 122,* 23 – 45.

Kuroda, H. (2019). Demographic Changes and Challenges for Financial Sector. Remarks at the Paris EUROPLACE Financial Forum in Tokyo, November 28, 2019. Retrieved from Bank of Japan (Bank of Japan).

Maudos, J., & Fernandez de Guevara, J. (2007). The Cost of Market Power in Banking: Social Welfare Loss vs. Cost Inefficiency. *Journal of Banking & Finance, 31*(7), 2103-2125.

Njoku, C.O., Nwadike, E.C. & Azuama, G.U. (2020). Effect of Electronic Banking on the Economic Growth of Nigeria [2009-2018]. The International Journal of Business Management and Technology, 4(3), 196-219.

Obite, C. P., Bartholomew, D. C., Nwosu, U. I., Esiaba, G. E., & Kiwu, L. C. (2021). The Optimal Machine Learning Modeling of Brent Crude Oil Price. *Quarterly*

Journal of Econometrics Research, 7(1), 31-43. https://doi.org/10.18488/journal.88.2021.71.31.43

Odusina, A. O. (2014). Automated Teller Machine Usage and Customers' Satisfaction in Nigeria. *Global Journal of Management and Business Research: C Finance, 14*(4), 1 – 12.

Panga, Y. O., Moutinhoa, L. K., & Lib, Y. (2015). Cash flow prediction using a grey-box model. In Proceedings of the 21st International Conference on *Automation & Computing,* University of Strathclyde, Glasgow, UK. IEEE. Retrieved from https://ieeexplore.ieee.org/document/7313951.

Sjöberg, J. (2000). A Nonlinear Grey-Box Example Using a Stepwise System Identification Approach. *IFAC Proceedings, 33*(15), 941–946. https://doi.org/10.1016/s1474-6670(17)39 – 87.

Sohlberg, B. & Jacobsen, E. (2008). Grey box modeling – Branches and Experiences. In *IFAC Proceedings, 41*(2), 11415–11420.

Uzowulu, H. C., Anyanwu, F. A., & Amakor, I. C. (2024). Impact of cashless policy on the Nigerian economy (2009- 2022). *African Banking and Finance Review Journal, 9*(9), 40–54. Retrieved from https://www.abfrjournal.com/index.php/abfr/article/view/142

Zeng, J., & Qiao, W. (2013). Short-term solar power prediction using a support vector machine regression. *Renewable Energy, 52,* 118–127. https://doi.org/10.1016/j.renene.2012.10.009.

Zulu, J.J. & Banda, B.M. (2015). The impact of labour productivity on economic growth: the case of Mauritius and South Africa. *Southern African Journal of Policy and Development, 2*(1), 26-41.