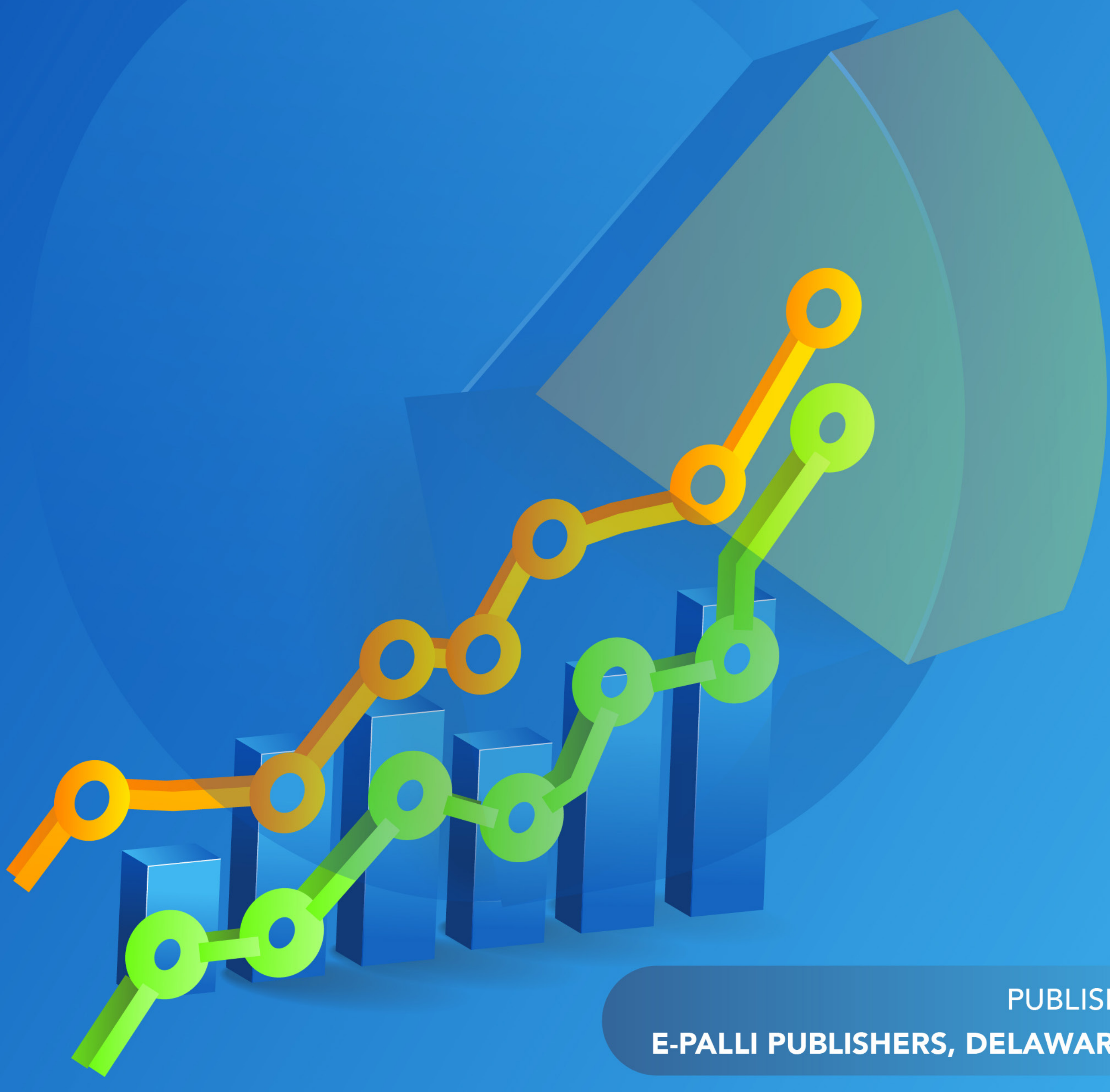
American Journal of Applied Statistics and Economics (AJASE)

ISSN: 2992-927X (ONLINE)

VOLUME 4 ISSUE 1 (2025)



PUBLISHED BY
E-PALLI PUBLISHERS, DELAWARE, USA

American Journal of Applied Statistics and Economics (AJASE)

Volume 4 Issue 1, Year 2025
ISSN: 2992-927X (Online)
DOI: https://doi.org/10.54536/ajase.v4i1.5751
https://journals.e-palli.com/home/index.php/ajase

Predictive Modeling of Ghana's Private Sector Pensions Asset under Management Contribution Using ARIMA Model

Chinton Emmanuel[1*], Donkoh Kojo Isaac[2], Acquah Oware Nana Emmanuel[3]

Article Information

Received: July 24, 2025

Accepted: August 25, 2025

Published: October 06, 2025

Keywords

ARIMA, Asset Under Management (AUM), Pensions Contribution in Ghana, Private Sector

ABSTRACT

This study applies an ARIMA (1,1,0) model to analyze the Private Sector Pension Assets Under Management (AUM) in Ghana. The model's parameters and performance metrics were evaluated using SARIMAX results and the Dickey-Fuller Test for stationarity. The SARIMAX model demonstrated a significant autoregressive term (ar. L1 = 0.9693) and acceptable performance metrics (MAE = 5.99, RMSE = 13.89, MAPE = 23.97%), indicating a strong influence of past values on current AUM. The diagnostic tests suggested that residuals were not autocorrelated and approximately normally distributed. The Dickey-Fuller Test further confirmed the stationarity of the time series, with a test statistic of -5.3314 and a p-value of 4.7116e-06, allowing us to reject the null hypothesis of a unit root. Overall, the ARIMA (1,1,0) model provides a reliable framework for forecasting and analyzing the Private Sector Pension AUM in Ghana, supported by robust statistical validation.

INTRODUCTION

Retirement planning plays a critical role in ensuring financial security during old age. Without adequate savings or income-generating assets, many individuals face severe financial challenges after leaving active employment (Diaw, 2017). To address this, most countries have adopted social security and pension systems that provide stable income for retirees and reduce old-age poverty.

In Ghana, the enactment of the National Pensions Act, 2008 (Act 766), marked a major reform of the pension system. The Act replaced the Social Security and National Insurance Law (PNDCL 247) and introduced a contributory three-tier pension scheme. These tiers comprise: (i) a mandatory basic national social security scheme managed by SSNIT, (ii) a mandatory occupational pension scheme managed by private trustees, and (iii) a voluntary provident and personal pension scheme. The Act also established the National Pensions Regulatory Authority (NPRA) to regulate and supervise pension administration. A key innovation of Act 766 was the extension of pension coverage to informal sector and self-employed workers, alongside those in formal employment (Abebrese, 2011).

The scheme requires a total monthly contribution of 18.5% of basic salary, with 13.5% allocated to Tier 1 and 5% to Tier 2. Tier 3 remains voluntary. The reform aimed to ensure income stability for retirees, harmonize pension provisions across the public and private sectors, and mobilize long-term funds for national development.

LITERATURE REVIEW

Theoretical Review

Life-Cycle Consumption Theory

Modigliani and Brumberg's (1954) life-cycle hypothesis provides the theoretical foundation for pension systems. It posits that individuals plan consumption and savings over their lifetime to smooth income across working and retirement years. Without adequate savings, retirees may face income insecurity. In Ghana, where extended family support systems are weakening, the theory underscores the need for deliberate retirement planning.

Positive Theory of Social Security

According to Sala-i-Martin (1996) and Tabellini (2000), public pensions improve economic efficiency by enabling older workers to retire, thereby creating employment opportunities for younger and more productive workers. Verbon (2012) further argues that pension systems act as retirement incentives where significant productivity gaps exist between older and younger generations.

Pooling Theory

Allen and Santomero (1998) highlight the efficiency of pension schemes in pooling risks, reducing transaction costs, and enhancing diversification. By mobilizing contributions, pension funds achieve economies of scale and improved investment outcomes (Matheson *et al.,* 2004; Bridgen & Meyer, 2008).

[1] Department of Statistics, University of Cape Coast, Ghana
[2] Financial Engineering, WorldQuant University, USA
[3] Department of Economics and Finance, Youngstown State University, USA
[*] Corresponding author's e-mail: emmanuelchinton7@gmail.com

Three-Tier Pension Scheme

Act 766 structures pensions into three tiers: the basic mandatory scheme (Tier 1), the mandatory occupational scheme (Tier 2), and the voluntary provident/personal pension scheme (Tier 3). Of the 18.5% total contribution, SSNIT retains 11% for retirement benefits and transfers 2.5% to the National Health Insurance Scheme. The remaining 5% is invested by Tier 2 trustees. Self-employed individuals may voluntarily participate in Tier 3, though participation levels remain low.

Empirical Review

Ghana's pension system has evolved since the colonial era, beginning with the Workmen's Compensation Ordinance of 1940 and the non-contributory Pension Ordinance of 1950 for civil servants (Darkwa, 2007; NPRA, 2010). Over time, reforms have sought to address sustainability, adequacy, and coverage gaps.

Kpessa (2011) observes that pensions in Africa play a crucial role in alleviating poverty among the elderly and supporting households under demographic pressure. Similarly, Agnew (2013) notes that pensions provide stable income for the aged, disabled, and unemployed. However, Fiiwe (2020) highlights shortcomings in benefit packages, particularly the absence of post-retirement healthcare, housing, and entrepreneurial support, which limit retirees' welfare.

International evidence shows similar trends. In the United States, private pension schemes date back to 1857, with American Express pioneering corporate pensions in 1878 (Bond, 2017). Pension benefits gained popularity during World War II as firms used them to retain workers amidst wage freezes (Pradmin, n.d.).

Informal Sector Participation

A major challenge in Ghana is extending pension coverage to the large informal sector. Although Act 766 permits voluntary participation through Tier 3, awareness and enrollment remain limited. A survey of self-employed workers revealed that over 70% were unaware of the scheme, while many who had knowledge of it contributed irregularly due to unstable incomes. This highlights the need for greater education, flexible contribution options, and innovative pension products tailored to informal sector workers.

Conceptual Framework of the Ghana Pension System

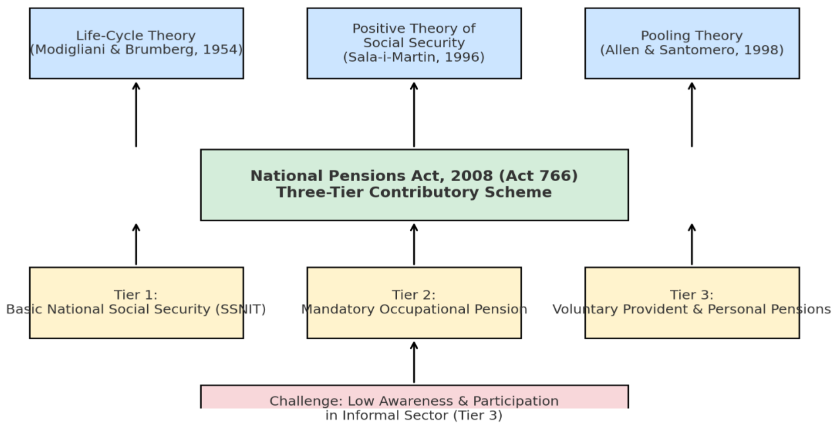


Figure 1: Conceptual Framework: Pension Theories and Ghana's Three-Tier Scheme

MATERIALS AND METHODS

Data Collection

The study utilized secondary data for the predictive modeling of the contributory pension assets under management of private sector in Ghana. For this study, we collected a time series of AUM of Private Sector pensions industry from the National Pensions and Regulatory Authority (NPRA) Annual reports from 2012 to 2023.

Statistical Analysis Tool

The study employed Time Series Statistical technique to analyze trends over time and forecast future pension AUM growth in Ghana's private sector and the analysis was done using Python Programming.

The study employed Autoregressive Integrated Moving Average (ARIMA) and SARIMAX to analyze the contributory trend over the period (2012-2023).

Sample Size

The sample size for the study comprised of 817 Pensions Trustees in Ghana

Autoregressive Integrated Moving Average (ARIMA)

The Auto-regressive integrated moving average (ARIMA) model is one of the most common prediction models, which is a time series analysis tool raised in the 1970s. It is a time series prediction model based on the fitting value of the past data sequence to extrapolate into future. It has 5 expressions: AR(P), MA(q), ARMA (p, q), ARIMA (p, d, q), ARIMA (p, d, q) × (P, D, Q)s

The Autoregressive Integrated Moving Average (ARIMA) model is a combination of the differenced autoregressive model with the moving average model. ARIMA model is said to be a unit-root non stationary because its AR polynomial has a unit-root and a conventional

Table 1: Asset Under Management of the Private Sector Pensions in Ghana (2012 -2022)

Years	Private Sector Pension AUM (GHS'Billion)
2023	46.5
2022	35.3
2021	28.0
2020	22.0
2019	17.3
2018	13.0
2017	9.8
2016	8.9
2015	8.8
2014	7.4
2013	4.8
2012	4.0

Source: National Pensions Authority (NPRA) Annual Report (2012- 2022).

approach for handling unit-root non-stationary is to use differencing (Tsay, 2010). If the differencing $W_t = Y_t - Y_{(t-1)} = (1 - B)\,Y_t$ or higher-order differencing $W_t = (1\text{-}B)^d\,Y_t$ of non- stationary time series then we call Y_t an ARIMA (p, d, q) process with order p of AR process, d the number of differences made for a series to become stationary and q is the order of MA process. It is expressed as:

$$Y'_t=I+\propto_1 Y'_{(t-1)}+\propto_2 Y'_{(t-2)}+....+\propto_p Y'_{(t-p)}+e_t+\theta_1 e_{(t-1)}+\theta_2 e_{(t-2)}+....+\theta_q e_{(t-q)} \quad (1.0)$$

$$\Phi_p(B)(1\text{-}B)^d Y_t=\theta_q(B)\propto_t \sim \text{ARIMA (p,d,q)} \quad (1.1)$$

Multiplicative Seasonal ARIMA (SARIMAX)

The seasonal ARIMA model incorporates both non-seasonal and seasonal factors in a multiplicative model: SARIMA (p, d, q) (P, D, Q) S. Box & Jenkins proposed the following model when dealing with a time series that contains seasonal fluctuations:

$$\Phi_p(B^S)\Phi_p(B)(1\text{-}B)^d(1\text{-}B^S)^D Y_t=\theta_q(B)\varepsilon_{(Q)}(B^S)\propto_t \quad (1.2)$$

Where Y_t is the observed value at time t, $\propto_t$ is the value at time t of white noise, d is order of differencing, is Φ_p (B) ordinary autoregressive component of order p and θ_q(B) and is the ordinary moving average component of order q, Sis number of seasons in a year and D is order of the seasonal differencing, Φ_p (B^S) and $\varepsilon_{(Q)}$(B^S) are the seasonal autoregressive and moving average difference of orders P and Q at lag s. According to Box & Jenkins (1976), the operator polynomials are:

$$\Phi_p(B)=(1\text{-}\varnothing_1 B\text{-}\ldots\varnothing_p B^p) \quad (1.3)$$

$$\theta_q(B)=(1+\varnothing_1 B\text{-}\ldots\theta_q B^p) \quad (1.4)$$

$$\Phi_p(B^S)=(1\text{-}\Phi B^S\text{-}\ldots\text{-}\Phi_p B^{sp}) \quad (1.5)$$

Box-Jenkins (ARIMA) Model

When performing a time series analysis using ARIMA models, three iterative steps must be used: diagnostic checking by examining residuals to assess the model's adequacy, parameter estimation by estimating the model's unknown parameters, and model identification by analyzing historical data.

Model Identification

Identification of the appropriate and suitable ARIMA model requires skills obtained by experience. Box & Jenkins postulates the following summary table on how to identify the model.

Table 2: Model identification (Box & Jenkins, 1976)

Model	ACF	PACF
ARIMA (p, d, 0)	Infinite. Tails off	Finite Cuts off after p lags
ARIMA (0, d, p)	Finite Cuts off after	Infinite. Tails off
ARIMA (p, d, q)	Infinite. Tails off	Infinite. Tails off

Model Identification

For ARIMA modeling, the autoregressive order (p) is usually determined by examining the partial autocorrelation function (PACF) of a stationary time series. If the PACF cuts off after a certain lag, the highest significant lag suggests the value of p. Conversely, if the PACF does not cut off, then p is often set to zero (Box & Jenkins, 1976). Similarly, the moving average order (q) is inferred from the autocorrelation function (ACF). A cutoff in the ACF after a few lags indicates the potential value of q, with the last significant lag serving as an estimate. In ARIMA (p, d, q) models, the autocorrelation patterns typically show exponential decay or damped sine-wave behavior after the first q–p lags.

Parameter Estimation

Once a tentative model structure is identified, parameter estimation follows. Box and Jenkins (1976) propose several approaches, including the method of moments, least squares, and maximum likelihood estimation (MLE). Given the non-linear nature of many ARIMA specifications, MLE is often preferred for its efficiency and robustness. When estimating residuals, backcasting may also be applied to obtain initial values for the error terms.

Diagnostic Checking

After estimation, diagnostic checks are essential to confirm model adequacy. Residuals should resemble white noise, meaning they are uncorrelated, normally distributed, and exhibit constant variance. A residual scatter plot should appear structureless, without systematic trends or patterns. Similarly, the residual autocorrelation function should not display significant spikes. Statistical tests, such as the Ljung-Box test or chi-square-based adequacy tests, are typically used to confirm that no significant autocorrelation remains in the residuals. Once these conditions are satisfied, the fitted ARIMA model can be considered adequate and used for forecasting.

RESULTS AND DISCUSSIONS

This section presents the outcome of the estimation of the model of this study. This begins with the forecast of private sector AUM using the ARIMA. For this study, the result presented in Chart 1 proves that private sector pensions AUM is projected to grow steeply by 2040.

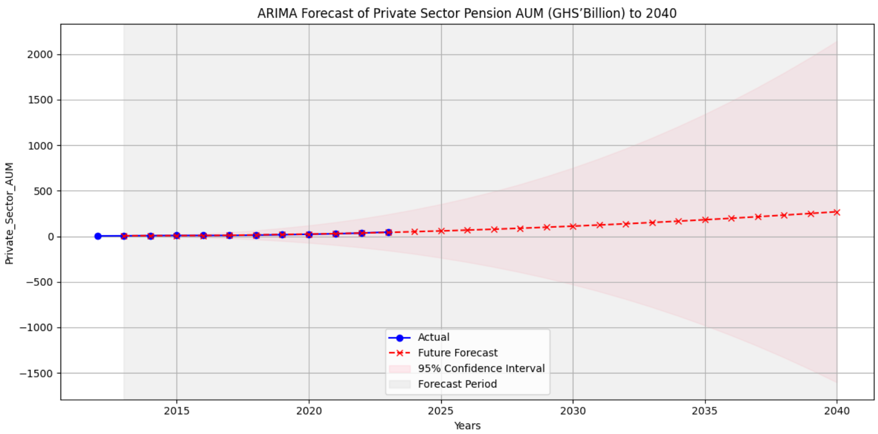


Figure 2: ARIMA Forecast of Private Sector Pension AUM (GHS' Billion) to 2040

Figure 3 presents the estimtaed results of the SARIMAX method for Ghana's private pensions AUM. The AR (AutoRegressive) term has a coefficient of 0.9693, which is significant (p-value 0.000). This indicates a strong influence of past values on the current value of the dependent variable. The variance of the error term is relatively high with a coefficient of 3.2070 and a marginally significant p-value (0.060), suggesting some level of uncertainty in the model. The diagnostic tests suggest that the residuals are not auto-correlated (Ljung-Box test) and are normally distributed (Jarque-Bera test). The heteroskedasticity test indicates no significant heteroskedasticity.

The performance metrics indicate that the model's predictions have a mean absolute error of 5.99, root mean squared error of 13.89, and mean absolute percentage error of 23.97%.

Overall, the SARIMAX model seems to fit the data well with significant AR term and acceptable performance metrics. However, the high variance of the error term suggests that there is some uncertainty in the model's predictions.

Figure 3 shows diagnostic plots for a statistical model. These plots are essential for diagnosing and validating the model's assumptions and fit.

```
                               SARIMAX Results
==============================================================================
Dep. Variable:      Private_Sector_AUM   No. Observations:                   12
Model:                  ARIMA(1, 1, 0)   Log Likelihood                 -23.422
Date:                 Fri, 27 Dec 2024   AIC                             50.844
Time:                         16:59:23   BIC                             51.639
Sample:                              0   HQIC                            50.342
                                  - 12
Covariance Type:                   opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1          0.9693      0.039     24.720      0.000       0.892       1.046
sigma2         3.2070      1.704      1.882      0.060      -0.133       6.547
===================================================================================
Ljung-Box (L1) (Q):                   0.61   Jarque-Bera (JB):                 0.27
Prob(Q):                              0.43   Prob(JB):                         0.87
Heteroskedasticity (H):               0.30   Skew:                            -0.32
Prob(H) (two-sided):                  0.27   Kurtosis:                         2.58
===================================================================================

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).

Performance Metrics:
Mean Absolute Error (MAE): 5.99
Root Mean Squared Error (RMSE): 13.89
Mean Absolute Percentage Error (MAPE): 23.97%
```

Figure 3: SARIMAX Results

Standardized Residuals for "P"

This plot displays the standardized residuals (differences between observed and predicted values) for a variable labeled "P" across different observations. The residuals fluctuate around the zero line, indicating how well the model's predictions match the actual data.

Histogram Plus Estimated Density

This plot combines a histogram of the residuals with

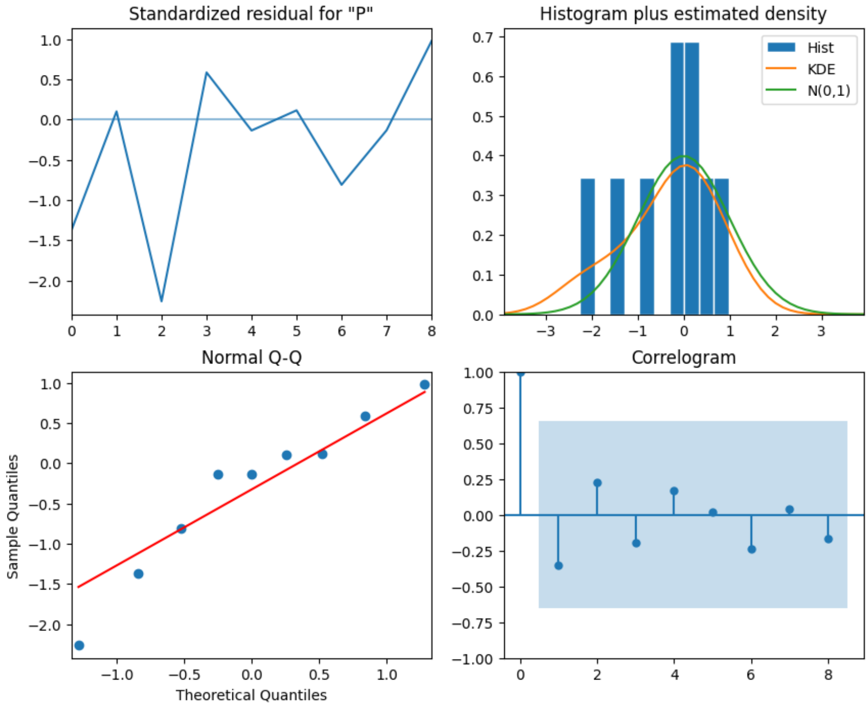


Figure 4: Diagnostic Plots for A Statistical Model

estimated density curves: -Histogram Bars: Show the frequency of residuals, Orange Line: Kernel Density Estimate (KDE) - a smoothed version of the histogram and Green Line: Represents the standard normal distribution (N(0,1)). The alignment of the orange and green lines with the histogram bars suggests whether the residuals follow a normal distribution.

Normal Q-Q Plot

A Quantile-Quantile (Q-Q) plot compares the sample quantiles of the residuals to the theoretical quantiles of a standard normal distribution:Red Line: Represents the expected line for normally distributed residuals and Points: Represent the actual residuals. The closer the points are to the red line, the more normally distributed the residuals are.

Correlogram

This plot shows the autocorrelation of the residuals at different lags: Points with Error Bars: Indicate the correlation values at various lags and the Shaded Area:Represents the confidence interval. Values within the shaded area suggest no significant autocorrelation, indicating the residuals are independent over time.

Table 3 shows the results of a Dickey-Fuller test, which is used to test for the presence of a unit root in a time series sample. The test statistic of -5.3314 is more negative than all the critical values at the 1%, 5%, and 10% significance levels. Combined with the very low p-value, this provides strong evidence to reject the null hypothesis. This suggests that the time series is stationary and does not have a unit root.

Chart 5 present the times series plots of the private sector over the last 12 years indicating that the AUM of private pensions has always been on the upward trajectory.

Table 3: Dickey-Fuller Test Result

Metric	**Value**
Test Statistic	-5.331440528065261
p-value	4.711563618849688e-06
# Lags Used	2.0
Number of Observations Used	9.0
Critical Value (1%)	-4.473135048010974
Critical Value (5%)	-3.28988060356653
Critical Value (10%)	-2.7723823456790124

Table 4 presents on the forcasted times series values of the private sector AUM from 2012 to 2040 taking into account a 95% confidence intervals for both the lower and upper bound.

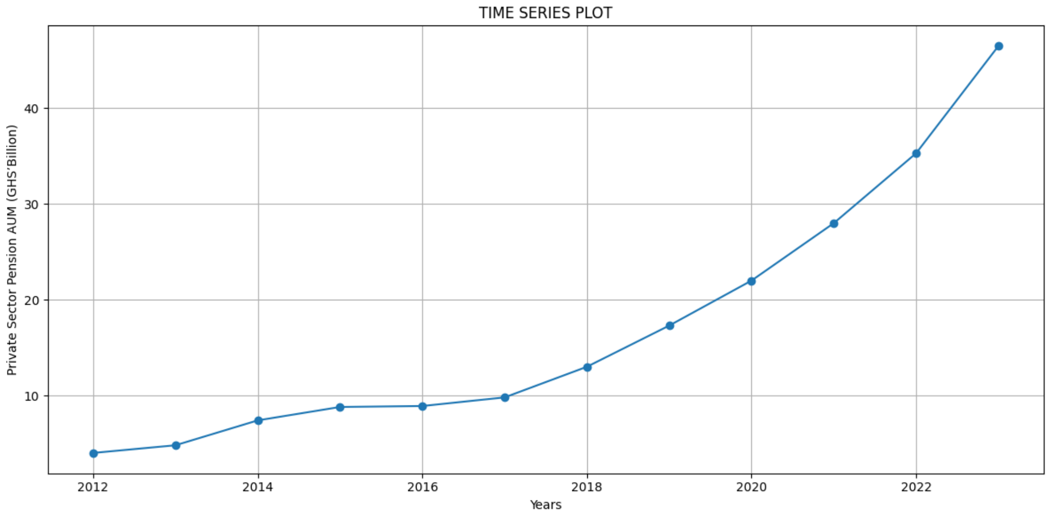


Figure 5: Time Series Plot

Table 4: Forcasted times series values of the private sector AUM from 2012 to 2040

Year	Forecasted Values (GHS'Billion)	Lower Bound (95% CI) (GHS'Billion)	Upper Bound (95% CI) (GHS'Billion)
2013	5.423536852	3.275484696	7.571589008
2014	6.720661869	-0.319161914	13.76048565
2015	7.247450909	-6.312625896	20.80752771
2016	8.354215719	-13.60046814	30.30889958
2017	11.15636763	-22.46023611	44.77297137
2018	15.24192579	-34.48246896	64.96632053
2019	19.61959785	-50.78297042	90.02216611
2020	24.0972536	-71.21520887	119.4097161
2021	29.28331686	-95.25346974	153.8201035
2022	35.6109406	-123.0317088	194.25359
2023	42.86059712	-155.3054794	241.0266737
2024	50.61826353	-192.6798141	293.9163411
2025	58.83731515	-235.2233702	352.8980005
2026	67.78828048	-282.8182577	418.3948187
2027	77.63626747	-335.6172463	490.8897813
2028	88.27024515	-394.0358105	570.5763008
2029	99.51904349	-458.4250597	657.4631466
2030	111.3777465	-528.9098723	751.6653653
2031	123.9652102	-605.5382018	853.4686221
2032	137.3429456	-688.4720041	963.1578953
2033	151.4567871	-777.979251	1080.892825
2034	166.2367155	-874.2954576	1206.768889
2035	181.6869161	-977.5598355	1340.933668
2036	197.8590336	-1087.881322	1483.599389
2037	214.7753175	-1205.417149	1634.967784
2038	232.4100458	-1330.365886	1795.185977
2039	250.7348986	-1462.908715	1964.378512
2040	269.7542712	-1603.185231	2142.693774

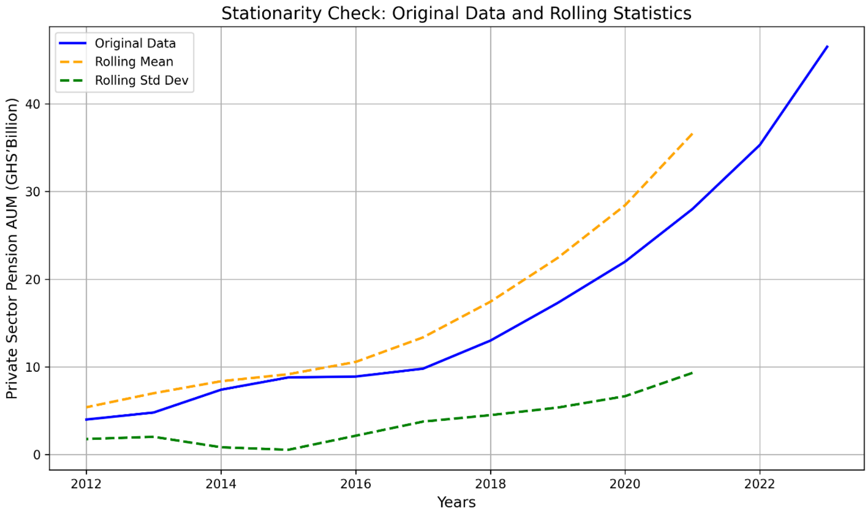


Figure 6: Stationarity Check: Original data and rolling statistics

CONCLUSION

The ARIMA (1,1,0) model provides a reasonably good fit for the Private Sector Pension AUM series. The statistically significant coefficient of the autoregressive term AR (1) indicates that current values are strongly influenced by their immediate past observations. Model diagnostics further show that the residuals are free from serious autocorrelation, as confirmed by the Ljung-Box test, and are approximately normally distributed according to the Jarque-Bera test. Nonetheless, signs of heteroskedasticity were detected, which suggests the need for additional adjustments to enhance the model's robustness.

In terms of accuracy, the model yields acceptable forecast error measures, including MAE, RMSE, and MAPE, making it suitable for short-term predictions. Moreover, the Augmented Dickey-Fuller test confirms the stationarity of the series, meaning its statistical properties such as mean and variance remain stable over time. This stationarity is particularly important, as it provides a strong foundation for reliable time series modeling and forecasting.

Finally, we suggest extensions to the model (ARIMA with GARCH, SARIMA with GARCH) to specifically tackle the heteroskedasticity problem for future research work

REFERENCES

Abebrese, J. (2011). Social Protection in Ghana: An Overview of existing programs and their prospects and challenges. *Friedrich Ebert Foundation.* (Pdf) Available at http://www.fesghana.org/uploads/PDF/FES_SocialProtectionGhana_2011_FINAL.pdf Accessed on November 28th, 2012

Agnew, J. (2013). Australia's Retirement System: Strengths, Weaknesses, and Reforms. *Issue in Brief,* (13). Boston College: Center for Retirement Research

Bond, T. (2017). *160 years of public pensions in the United States.* Retrieved from https://protectpensions.org/2017/04/29/public-pensions-early-history/

Box, GEP, Jenkins, G. (1970). *Time series analysis: forecasting and control.* 2nd Edition. San Francisco: Holden-Day. pp. 240.

Darkwa, O. (2007). Reforming the Ghanaian Social Security System; Prospects and Challenges. *Cross-Cultural Gerontology* (12).

Fiiwe, J. L. (2020). Analysis of Retirement Benefits in Nigeria: A Case Study of Selected Federal Establishment. *Equatorial Journal of Marketing and Insurance Policy, 4*(1), 1-15.

Kpessa, M. (2011). A Comparative Analysis of Pension Reforms and Challenges in Ghana and Nigeria. *International Social Security Review, 64*. 92-104

Kpessa, M. W. (2010). The Politics of Retirement Income Security Policy in Ghana: Historical Trajectories and Transformative Capabilities. *African Journal of political Science and International Relations,* 92-102

Modigliani, F. &Brumberg, R. H. (1954). *Utility analysis and the consumption function: an interpretation Of cross-section data,* (In Kenneth K. Kurihara, ed.). Post-Keynesian Economics, New Brunswick, NJ. Rutgers University Press. 388–436.

National Pensions Act of Ghana. (2008). Act 766.

Vroom, V. H. (1964). *Work and motivation.* San Francisco, CA: Jossey-Bass

World Bank (1994). *Averting the old age crises policies to protect: The old and Promote Growth.* Washington D.C: The World Bank.