World Academy of Science, Engineering and Technology
International Journal of Economics and Management Engineering
Vol:11, No:11, 2017

Application of Generalized Autoregressive Score Model to Stock Returns

Katleho Daniel Makatjane, Diteboho Lawrence Xaba, Ntebogang Dinah Moroke

***Abstract*—**The current study investigates the behaviour of time-varying parameters that are based on the score function of the predictive model density at time t. The mechanism to update the parameters over time is the scaled score of the likelihood function. The results revealed that there is high persistence of time-varying, as the location parameter is higher and the skewness parameter implied the departure of scale parameter from the normality with the unconditional parameter as 1.5. The results also revealed that there is a perseverance of the leptokurtic behaviour in stock returns which implies the returns are heavily tailed. Prior to model estimation, the White Neural Network test exposed that the stock price can be modelled by a GAS model. Finally, we proposed further researches specifically to model the existence of time-varying parameters with a more detailed model that encounters the heavy tail distribution of the series and computes the risk measure associated with the returns.

***Keywords*—**Generalized autoregressive score model, stock returns, time-varying.

I. Introduction

TIME-VARIATION in the parameters describing a stochastic time series process is pervasive in almost all applied scientific fields. Often, the model of interest is pigeonholed by parameters varying over time where literature has proposed innumerable possible specifications. Recently, [4]-[6] had noted that there is some difficulty in estimating certain models; particularly the stochastic models that are not accurate in yielding the shape of the conditional distribution of data. Generalised autoregressive conditional heteroscedasticity (GARCH) are the typical family and special cases of the generalised autoregressive score (GAS) models [1]. Though, we assume that demeaned return is the chauffeur of the time-varying within the conditional variance that is autonomous of the shape of the conditional distribution. To see this relationship, consider the case of observing a 10% return when the conditional mean is 0% and the volatility is 3%. By assuming a normal distribution, [1] emphasised that a 10% return entails a significantly increased volatility indicating fat-tail distribution with high likelihood of extreme values.

Economic and financial time series often exhibit intricate dynamic features. When the time series is analysed using a parametric dynamic model, it needs to be sufficiently flexible to describe its salient features. After the introduction of a GAS model by [4]-[6], we then take further the application of the model to the stock returns with the aim of investigating the behaviour of time-varying parameters that are based on the score function of the predictive model density at time t. The pros of using the conditional score as the driver are that, the estimation of the model parameters by maximum likelihood (ML) is frank. Furthermore, these models have proved to be more robust in modelling and predicting fat tail or skewed data such as that of the empirical finance. See for example [6], [7], [9]-[13]. The dynamics of correlations and volatilities in these models are driven by the score of the predictive conditional distribution. If the latter is fat-tailed, the score driven dynamics automatically correct the influential observations [3]. This empirical analysis is in four sections. Section II presents methods and procedures. Section III presents empirical analysis, and lastly, section IV presents the conclusion and recommendations.

Katleho Makatjane is with the North-West University, South Africa (e-mail: katleho.makatjane@outlook.com).

II. Methods and Procedures

This study assimilates the behaviour of time-varying parameters by utilising a GAS model. The results are presented in tables and graphs. Discussed in this section are the methods and procedures for preparing the data for empirical analyses. The results obtained here are meant to give guidance about the nature of the data and the type of models to estimate.

A. Nonlinearity

In the time series setting, there are methods to be executed prior model estimation. These methods include among others the linear test and or non-linear unit root tests. In testing the non-linear alternative hypothesis, [10] had employed a detailed test for linear regression analysis known as the Ramsey RESET test. The test was established by [14]. However, with the current study, we propose a white neural network test of [16] and the test is computed by first letting the pragmatic training set that is attained as a haphazard sequence be $Z^n = \{Z_t, t = 1, .., n\}$, $Z_t = (Y_t, X_t)$. Having Y_t as a scalar quantity, and X_t as the row vector for finite dimension, then, [16] rumoured Z_t to be neither identically but not independently from the past values but be distributed within $t = 1, \dots, n$. In this case, the conditional $E(Y_t|X_t)$ is the one governing the studying of the relationship between X_t and Y_t. To be more unpretentious, this can be extended to the regression function as:

$$g(X_t) = E(Y_t|X_t) \tag{1}$$

where, g is independent of t because of the identical distribution assumption. Here, the goal line is to investigate the competence of a given multilayer feedforward network as a representation of the unknown mapping function of g by considering a class of feedforward networks in which the network output scalar is determined by some given input x. At that moment, white demonstrated the network as follows:

$$o = \tilde{x}\gamma_0 + \sum_{j=1}^{q} \beta_j \delta(\tilde{x}\ \gamma_j) = f(x,\theta) \quad (2)$$

where, o is the network output and $\tilde{x} = (1, x)$ is a conformable column vector of the connection strengths from the input layer, $j = 0, \ldots, q$, with β_j being a scalar connection strength from the hidden unit j to the output unit $j = 0, \ldots, q$, while δ is a veil of secrecy function that follows either a logistic or hyperbolic tangent squasher. In this case, q symbolizes the number of hidden units with the connection strengths that are not constrained to zero. If the network can be meticulous in representing the indefinite function g, then θ^* is computed as a vector of the connection strengths such that $g(X_t) = f(X_t, \theta^*)$. Consequently, the null hypothesis is formulated as:

$$H_0: P[g(X_t) = f(X_t, \theta^*)] = 1 \text{ for some } \theta^* \quad (3)$$

This establishes a detailed statement of the null hypothesis of interest. Henceforth, the alternative is that the network is unable to exact the representation of g and it is formulated as:

$$H_a: P[g(X_t) = f(X_t, \theta^*)] < 1 \; for\ all\ \theta \quad (4)$$

A statistical test of the null hypothesis is a Lagrange multiplier test of [5] and the test is based on the linear regression:

$$\hat{\varepsilon}_t^2 = \eta_0 + \sum_{i=0}^{m} \eta_i \varepsilon_{t-i}^2 + e_t \quad (5)$$

where $\eta_0, \ldots, \eta_m$ are the parameter estimates, $\upsilon_t \sim i.i.d(\mu = 0, \sigma_\varepsilon^2 = 1)$ and the test statistic is the usual F-statistic:

$$F^* = \frac{\left(\frac{RSS}{TSS}\right)/_{m}}{1-\left(\frac{RSS}{TSS}\right)/_{n-m-1}} \sim F_{\alpha,(m,n-2m-1)} \quad (6)$$

where RSS and TSS are the regression sum of squares and total sum of squares of model (5), respectively. In general, the quantity $\frac{RSS}{TSS}$ is the mathematical computation of the coefficient of determination also known as R^2. The null hypothesis is rejected if the F value is greater than critical value of $F_{\alpha,(m,n-2m-1)}$ and concludes that the stock returns are non-linear in nature.

B. GAS Framework for Modelling Time-Varying Parameters

For both univariate and multivariate time series settings, the GAS framework is most striking because of its physiognomies to be able to outline time-varying as large. In reviewing the model, we follow [1], by firstly acquainting with the notation and presenting the GAS model when the parameter space is unrestricted. Furthermore, a mapping function is elaborated in showing how time-varying parameters are being modelled with some restricted parameter space.

According to [4], the GAS is specified by letting $Y_t \in \Re^N$ be an N-dimensional random vector at time t with the following subsequent conditional distribution:

$$Y_t \big| Y_{1;t-1} \sim p(Y_t; \theta_t) \quad (7)$$

$Y_{1;t-1} \equiv (Y_1', \ldots, Y_{t-1}')'$ holds the past values of Y_t up to time t-1 and $\theta_t \in \Theta \subseteq \Re^J$ is some time-varying parameters' vector which fully typifies p(.) and solitary be contingent on $Y_{1;t-1}$ and have some inert added parameters ζ for all t. The fruition in the time-varying parameter vector θ_t is the main feature of the GAS model and it is driven by the score of the conditional distribution defined in (7) with the following autoregressive component incorporated:

$$\theta_{t+1} \equiv \alpha + \phi \varrho_t + \varphi \theta_t \quad (8)$$

In this case, α, ϕ and φ are the coefficient matrices with the proper dimensions collected in ζ. A vector that is proportional to the score of (7) is denoted as ϱ_t and it is defined as follows:

$$\varrho_t \equiv \vartheta_t(\theta_t) \nabla_t(Y_t, \theta_t) \quad (9)$$

Note that $\vartheta_t = J * J$ is a positive definite scaling matrix that is known at time t and $\nabla_t(Y_t, \theta_t) \equiv \frac{\partial \ln p(Y_t;\theta_t)}{\partial \theta_t}$ is the score of (7) that is appraised at θ_t. The suggestion is to put the scaling matrix ϑ_t to a power of $\gamma > 0$ of the inverse from the information matrix of θ_t, to account for the variance of ∇_t as it is seen in [4] framework. Thus, [1] indicated that this simplifies to $\vartheta_t(\theta_t) \equiv \varphi_t(\theta_t)^{-\gamma}$ with:

$$\varphi_t(\theta_t) \equiv E_{t-1}[\nabla_t(Y_t, \theta_t) \nabla_t(Y_t, \theta_t)'] \quad (10)$$

Here, the expectation is taken with respect to the conditional distribution of $Y_t \big| Y_{1;t-1}$ and γ is been kept fixed at the subset values of $\gamma \subset \{0, 0.5, 1\}$. $\vartheta_t = I$ as $\gamma = 0$. Therefore, this implies no scaling [1]. I, in this case, is the identity metrics of the scaling parameter.

III. Empirical Analysis

A daily time series of Sanlam stock prices for the period of January 2002 to October 2015 is used for the effective analysis of investigating the behaviour of time varying parameters by utilising GAS and R 3.4.1 programming is utilized for the analysis. In accordance to [10], it is much more significant to assess the data behaviour patterns prior to data analysis. These perhaps help in identifying the associated data properties that enable an analyst to decide on the type of time series model to be used. Stock returns data are displayed in Fig. 1. The plot of stock returns in Fig. 1 uncovers that the series is non-stationary as it clearly shows some clustering volatility; hence, the series is to be modelled by time-varying model.

A. Exploratory Data Analysis

In this section, the preliminary data analyses are conducted with the purpose of assessing the behaviour of the data set as recommended by [11]. Nonetheless, we adopt the methods of [12] for data exploration where both authors used descriptive statistics, and the results are tabulated in Table I.

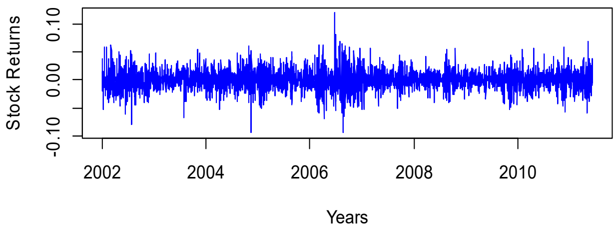

Fig. 1 Sanlam Stock Returns

TABLE I
EXPLORATORY DATA ANALYSIS

Statistic	Stock Returns
Mean	7.646
Std.dev	0.678
skewness	0.153
Kurtosis	2.104
JB	128.770 (0.00)
White neural network test	
Statistic	**Prob**
6.219	0.045

NB: The number in () presents the JB probability values

The descriptive statistics indicate that the return series do not have normal distribution character as revealed by Jarque-Bera (JB) test statistic. The JB test was established by [8]. This points toward that the time varying parameter model should be estimated using a non-normal distribution such as skewed student-t. The causes of the non-normality are the shocks imposed on the stock market activities, which causes stock price data to be volatile and vary with time. Consequently, the predictive performance of the GAS model would be better than the traditional time-driven models. The mean of the log returns is estimated at 7.646%, suggesting that daily, the average stock return is approximated at 7.5%. Additionally, in testing for the presence of time-varying parameters, the estimated White neural network test revealed that the stock returns are non-linear in nature implying that stock returns are time variant.

B. Generalised Autoregressive Score Model

The estimated coefficients of the marginal model are reported in Table II. Our results confirm the presence of those properties that are usually found in the financial econometric literature, specifically that focusing on volatility modelling, such as strong persistence of volatility and positive reaction of the conditional variance to negative innovations. The skewness parameter η_i of the skewed student distribution, as reported in row 7, is statistically significant and smaller than one justifying our choice of skewed innovations. However, the same results were found by [2] in their study of Switching-GAS copula models for systemic risk assessment. To confirm the presence of excess kurtosis and the departure from the normality, the parameter ν_i is found to be highly significant; hence, the conclusion is that the Sanlam stock price returns are leptokurtic. This is also evident from the visualization of the stock returns properties in Fig. 2. The quantile-quantile plot, and normal probability density function (PDF) with the normal histogram, significantly indicates the heavy tail properties. According to [15], stock market data significantly displays a leptokurtic behaviour, which is also known as heavy tailed behaviour, and hence, the need to model and predict the stock price returns behaviour with the model that qualifies this task. For Sanlam stock returns, the estimated value of ν is 47.313, implying a highly leptokurtic distribution. The conclusion here is that the findings are probably related to the 2007/2009 financial crisis that South Africa experienced.

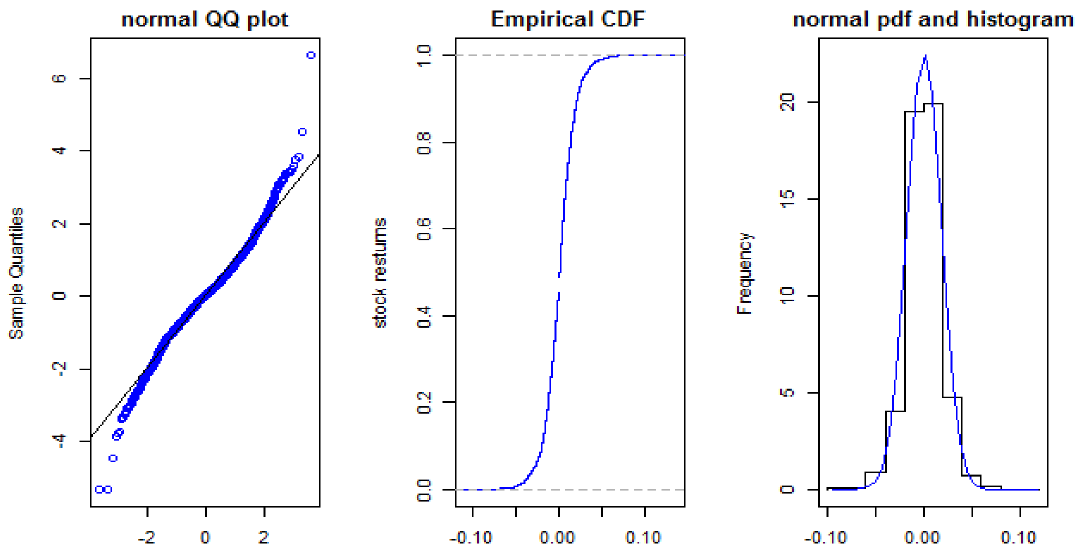

Fig. 2 Empirical properties of Stock returns

TABLE II
PARAMETER ESTIMATES

Parameter	Estimate	Std. Error	t value	prob
$\emptyset_{0,i}$	0.995	0.108	469.168	0.00
$\emptyset_{1i}$	0.029	0.0138	2.156	0.02
ϖ_i	16.486	0.049	33579290	0.00
ϑ_{1i}	12.275	0.0128	9581192	0.00
ϑ_{2i}	9.993	0.0125	79703660	0.00
ϑ_{3i}	0.219	0.0146	14.46457	0.00
η_i	0.975	0.055	17719.05	0.00
ν_i	47.313	0.0221	449.9282	0.00
	Unconditional Parameters			
	location	scale	skewness	shape
	1964.14	1050.51	1.5	50

IV. CONCLUSION

The current study aims to empirically investigate the behaviour of the time-varying parameter by estimating the GAS model to the South Africa Sanlam stock price returns. According to [1], the GAS model serves as an extension of the GARCH family models which assume that the conditional distribution does not vary over time. A vibrant advantage of the GAS model is that it exploits the full likelihood of information. Taking a local density score step as a driving mechanism, the time-varying parameters increase and produced a clear indication of a leptokurtic behaviour, in which the empirical properties revealed the same behaviour with the skewness of 1.5 in the unconditional parameters. The same behaviour of the leptokurtic was also realised in [15], [2]. Prior model estimation, the white neural test proved that the stock returns can be modelled with a time-varying parameter model. For further research, the $\mathrm{HEAVY\ GAS-tF}$ model of [13] should be utilised and go deeper in estimation of the correlation dynamics of the GAS model with heavy tails of the returns. Moreover, financial risk analysts and financial risk managers should, however, implement the model for risk computation and prediction of the heavy tails possessed in the financial time series.

REFERENCES

[1] Ardia, D., K. Boudt, and L. Catania, *Generalized Autoregressive Score Models in R: The GAS Package.* 2016.

[2] Bernardi, M. and L. Catania, *Switching-GAS copula models for systemic risk assessment.* arXiv preprint arXiv:1504.03733, 2015.

[3] Creal, D., S. J. Koopman, and A. Lucas, *A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations.* Journal of Business & Economic Statistics, 2011. 29(4): p. 552-563.

[4] Creal, D., S. J. Koopman, and A. Lucas, *Generalized autoregressive score models with applications.* Journal of Applied Econometrics, 2013. 28(5): p. 777-795.

[5] Engle, R. F., *Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation.* Econometrica: Journal of the Econometric Society, 1982: p. 987-1007.

[6] Harvey, A. C., *Dynamic models for volatility and heavy tails: with applications to financial and economic time series.* Vol. 52. 2013: Cambridge University Press.

[7] Janus, P., S. J. Koopman, and A. Lucas, *Long memory dynamics for multivariate dependence under heavy tails.* Journal of Empirical Finance, 2014. 29: p. 187-206.

[8] Jarque, C. M. and A. K. Bera, *Efficient tests for normality, homoscedasticity and serial independence of regression residuals.* Economics letters, 1980. 6(3): p. 255-259.

[9] Lucas, A., B. Schwaab, and X. Zhang, *Conditional euro area sovereign default risk.* Journal of Business & Economic Statistics, 2014. 32(2): p. 271-284.

[10] Makatjane, K., N. Moroke, and D. Xaba, *Threshold Cointegration and Nonlinear Causality test between Inflation Rate and Repo Rate.* Journal of Economics and Behavioral Studies, 2017. 9(3): p. 163-170.

[11] Makatjane, K. D. and N. D. Moroke, *Comparative Study of Holt-Winters Triple Exponential Smoothing and Seasonal Arima: Forecasting Short Term Seasonal Car Sales in South Africa.* 2016.

[12] Makatjane, K. D. and T. J. Makatjane, *Factors that Associated with the Academic performance of First year Students at the National University of Lesotho: Structural Equation Modelling Approach.* International Journal of Statistics and Applied Mathematics, 2017. 2(1): p. 42-49.

[13] Opschoor, A., et al., *New HEAVY models for fat-tailed realized covariances and returns.* Journal of Business & Economic Statistics, 2017: p. 1-15.

[14] Ramsey, J. B., *Tests for specification errors in classical linear least-squares regression analysis.* Journal of the Royal Statistical Society. Series B (Methodological), 1969: p. 350-371.

[15] Wentzel, D. and E. Mare, *Extreme value theory—An application to the South African equity market.* Investment Analysts Journal, 2007. 36(66): p. 73-77.

[16] White, H. *An additional hidden unit test for neglected nonlinearity in multilayer feedforward networks.* in *Proceedings of the international joint conference on neural networks.* 1989. Washington, DC.