*STATISTICS IN TRANSITION new series, September 2023*
*Vol. 24, No. 4, pp. 93–107,* https://doi.org/10.59170/stattrans-2023-053
Received – 06.11.2021; accepted – 21.11.2022

# Hyper-parametric Generalized Autoregressive Scores (GASs): an application to the price of United States cooking gas

**Rasaki Olawale Olanrewaju**[1], **Sodiq Adejare Olanrewaju**[2], **Omodolapo Waliyat Isamot**[3]

## Abstract

This paper presents the framework of the Generalized Autoregressive Score (GAS) model with a variety of symmetric conditional densities of different time-varying hyper-parameters. The distinctive trait and goal of the observation-driven GAS model is to use its score and information functions as the compeller of time-variation via hyper-parameters of conditional densities. 10 robust hyper-parametric conditional densities were used as random error drivers for the GAS model with an application to the price of the United States cooking gas in the period between 2005 and 2020. Out of the 10 robust hyper-parametric conditional noises for the GAS model, the Asymmetric Student–t with one tail decay parameter (AST1) outperformed other categories of its variants and other conditional densities subjected to the GAS model, achieving a minimum reduced-error performance of AIC=11943.277 and BIC=12014.525. The hyper-parametric model obtained a location score and scale score of -1.2634 and 0.6636, respectively, while its location information and scale information was 0.2691 and 0.0362, respectively. Furthermore, the GAS model via AST1 proved more efficient than the core volatile conditional heteroscedasticity model of the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) at GARCH (1,1) via a Gaussian distributed noise.

**Key words:** Asymmetric Student–t, Generalized Autoregressive Score, hyper-parameters, score, information.

## 1. Introduction

Describing and estimating time-varying variation in stochastic time series has been the process of aperture across all fields of applied statistics and most scientific

[1] Africa Business School (ABS), Mohammed VI Polytechnic University (UM6P), Rabat, Morocco. E-mail: olanrewaju_rasaq@yahoo.com. ORCID: https://orcid.org/0000-0002-2575-9254.
[2] Department of Statistics, University of Ibadan, Ibadan, Oyo State, Nigeria.
E-mail: sodiqadejare19@gmail.com. ORCID: . https://orcid.org/0009-0006-4494-2421.
[3] Department of Epidemiology and Medical Statistics, University of Ibadan, Ibadan, Oyo State, Nigeria. E- mail: omodolapo.isamot@gmail.com.

investigations. Time-varying variation is cognate in modelling parameter selection for strategizing and capturing behavioural dynamics of either multivariate or univariate stochastic time series process with different myriad of possible specifications (Cox, 1981; Creal *et al*., 2013). According to Harvey & Luati (2014), some time-varying parameters of some proposed time series models are not only difficult to estimate (especially the class of stochastic volatility models reviewed by Olanrewaju *et al.* (2020) & Shephard (2005)), but also at times fail to take into consideration the shape of the conditional distribution of the data. These time-varying models in time series are categorized in two classes: parameter-driven models and observation-driven models. In the latter, the time variations of the parameters are used by subjecting the stochastic parameters to be functions of lagged dependent variables as well as synchronous and lagged exogenous variables. A typical example of such a model is the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) by Engle (1982). In the former, parameters are stochastic processes with their associated source of error, such that given the past and concurrent information the parameters cannot be perfectly predictable. Examples of such models are Stochastic Volatility (SV) model by Shephard (2005) and stochastic intensity models by Koopman *et al*. (2008).

In order to strengthen the observation-driven based model, Creal *et al*. (2013) and Harvey (2013) proposed the score function of conditional density functions as the compeller of time-variation in the time series parameters to describe the data. This resulted into score-driven model called Generalized Autoregressive Score (GAS) model, otherwise known as Dynamic Conditional Score (DCS) model. Among the merits of GAS model over other observation-driven models are: it is based on complete density function rather than moments, its likelihood evaluation is free from ambiguity, its driven mechanism is based on score and information functions (Hessian). The model can be extended to long memory, asymmetric and other intricate dynamics. It is flexible enough to be used in all fields in which the use of time-varying parameter models is relevant. It can be subjected to real-value, integer-valued, strictly positive or (0, 1)-bounded observations provided the conditional density (either probability density function or probability mass function) for the score function and Hessian exists and is well-defined (Oh and Patton, 2016). It provides framework for modelling time variation in parametric models when computing the score of a parametric conditional observational density with respect to time-varying parameter. The practical relevance of GAS model includes default and credit risk modeling as affirmed by Lucas & Zhang (2016); stock volatility and correlation modeling as declared by Harvey & Sucarrat (2014); modeling time-varying dependence structures as established by Harvey & Thiele (2016); CDS spread modeling and questions relating to financial stability and systemic risk, modeling high frequency data as confirmed by Janus *et al*. (2014) and spatial econometrics as affirmed by Blasques *et al*. (2016).

The novelty of this article is to extend the driving mechanism of the score function and Hessian of the GAS model via its random noise to some probability density functions like Normal and its variants, Asymmetric Student–t with two tail decay parameters, Asymmetric Student–t with one tail decay parameter, Student-t, location-scale skew-normal distribution, Skew-t distribution, Asymmetric Laplace, Gamma, and exponential. The notion of the mentioned conditional densities to GAS model is to be able to improve its score function and Hessian robustly via each conditional density time-varying hyper-parameters like location, scale, skewness, and shapes indexes. The high frequency financial data to be subjected to the GAS model via the mentioned conditional densities is the price of United State cooking gas. The raw dataset of the price of the United State cooking gas from 2005-2020 will be used as extracted from U.S. Energy Information Administration (EIA).

## 2. Model Specification

In this section, the general class of observation-driven time-varying parameter model will be formulated. Thereafter, the Generalized Autoregressive Score (GAS) for the time-varying hyper-parameters driven by scale function of conditional likelihood will be formulated to drive the score-function and Hessian. According to Monache and Petrella (2014), time-varying parametric autoregressive model of order "$i$" can be defined as:

$$x_t = \phi_{0,t} + \phi_{1,t}x_{t-1} + \phi_{2,t}x_{t-2} + \cdots + \phi_{i,t}x_{t-i} + \omega_t \quad (1)$$

where the error term is $\omega_t \sim (0, \sigma_t^2) t = 1,2,\cdots,n$; $\phi_0,\cdots,\phi_i$ are the parameters of the autoregressive model; $x_{t-1},\cdots,x_{t-i}$ are the past series values (lags).

Olanrewaju & Folorunsho (2018) proposed an updating rule by defining the associated variation of the time-varying hyper-parameters in a vector form to be: $g_t = (\phi_t', \sigma_t^2) \ni \phi_t' = \left(\phi_{0,t}, \phi_{1,t}, \cdots, \phi_{i,t}\right)'$. This implies that equation (1) can be interpreted as the first order of a Markov process with

$$g_{t+1} = \eta + Kg_t + \xi_t, \xi_t \sim (0, \Sigma_t) \quad (2)$$

where $\eta$ is a vector of constants; K and $\Sigma$ are the matrices of hyper-parameters (updated location and scale parameters respectively), $g_t$ connotes the time-varying parameters. The Generalized Autoregressive Score (GAS) for the time-varying hyper-parameters driven by scale function of conditional likelihood of $g_t$ given the immediate past of "$t-1$", $g_{t-1} = (\phi_{t-1}', \sigma_{t-1}^2)$

$$g_{t+1/t} = \eta + Kg_{t/t+1} + Zc_t \quad (3)$$

where, $X_{t-1} = \{x_{t-1}, x_{t-2}, \cdots, x_1\}$, $\eta$and K are the same as defined above, where $Zc_t \sim (0, \sigma_t^2) t = 1,2, \cdots, n$ is the error term of the GAS time-varying hyper-parameters with driven mechanism called the score-function.

$$c_t = C_t \nabla_t \quad (4)$$

$$\ni \ \nabla_t = \frac{\partial[\log p(x_t/(Z_t;\theta_t)]}{\partial g_{t/t-1}}; \quad C_t = I_{t-1}^{-1} = \left[\frac{\partial[\log p(x_t/(Z_t;\theta_t)]}{\partial g_{t/t-1} g'_{t/t-1}}\right]^{-1}$$

with $I_{t-1}^{-1}$ being the Information matrix (Hessian), $Z_t = [G_t, X_{t-1}]$ and $G_t = \{g_{t/t-1}, g_{t-1/t-2}, \cdots, g_{1/0}\}$ defined for vector parameters of $\theta_t$; $p(x_t/Z_t;\theta)$ is probability of the past series values (lags) at time "t" given that the error $(Z_t)$ and vector parameters $(\theta_t)$ at time. Rewriting equation (1) in matrix form gives

$$x_t = A'\phi_{t/t-1} + \omega_t \ni \omega_t/X_{t-1} \sim (0, g_{t/t-1}), for t = 1, \cdots, n, \quad (5)$$

$$\sigma_{t/t-1}^2 = g_{t/t-1}, A' = [1, x_{t-1}, \cdots, x_{t-p}] \& \phi_{t/t-1} = [\phi_{0,t/t-1}, \phi_{1,t/t-1}, \cdots, \phi_{p,t/t-1}]'$$

$$\omega_t = x_t - A'\phi_{t/t-1},$$

$$\omega_t = x_t - A'\phi_{t/t-1}, \mu_t = A'\phi_{t/t-1}$$

The matrix form of equation (5) will be incorporated into:

**Student-t-Distribution** as

$$p(x_t;\theta_t) = \frac{\Gamma\left(\frac{v_t+1}{2}\right)}{\Gamma\left(\frac{v_t}{2}\right) g_t \sqrt{\pi v_t}} \left(1 + \frac{(2\phi'_{t/t-1}A + \omega'_t)}{v_t g_{t/t-1}^2}\right)^{-\frac{v_t+1}{2}} \quad -\infty < x_t < +\infty \quad (6)$$

$\omega_t/X_{t-1} \sim NID(0, g_{t/t-1}, v_t)$ for location parameter $\mu_t$, scale parameter $g_t$, $v_t$ degree of freedom $\theta_t = \{\mu_t, \phi, g_t, v_t\}'$. According to Jones and Faddy (2003), **Asymmetric Student–t with two tail decay parameters** (that is the Student t-distribution, which is both heavy tailed and skew). Then, the density function of this new distribution is

$$p(x_t;a,b) = C_{a,b}^{-1} \left\{1 + \frac{x_t}{(a+b+x_t^2)^{\frac{1}{2}}}\right\}^{a+\frac{1}{2}} \left\{1 - \frac{x_t}{(a+b+x_t^2)^{\frac{1}{2}}}\right\}^{b+\frac{1}{2}} \quad (7)$$

where,

$C_{a,b} = 2^{a+b-1} B(a,b)(a+b)^{\frac{1}{2}}$, where $B(a,b)$ denotes the beta function. When a = b, $p(x_t;b,a)$ reduces to the Student- t-distribution on (2a) degrees of freedom (**Asymmetric Student–t with one tail decay parameter**). When a<b or a>b, $p(x_t;b,a)$ is negatively or positively skewed respectively. In fact, $p(x_t;b,a) = p(-x_t;b,a)$. Note that "a" and "b" are positive real numbers and need not to be integer or half-integer.

**Location-Scale Skew-Normal distribution**

According to Owen (2008), a random variable X is said to be a location-scale skew-normal distribution, with location at $\mu$, scale at δ and shape parameter $\alpha$, and denoted $X \sim \theta = SN(\mu, \delta^2, \alpha)$ if its probability density function (pdf) is given by

$$p(x_t\,; \theta_t) = \frac{2}{\delta}\phi\left(\frac{x_t-\mu}{\delta}\right)\Phi\left(\alpha\frac{x_t-\mu}{\delta}\right), x_t \in \mathbb{R}(\alpha, \mu \in \mathbb{R}, \delta \in \mathbb{R}^+),$$

Then, $$p(x_t\,; \theta_t) = \frac{2}{\delta}\phi\left(\frac{x_t-A'\phi_{t/t-1}}{\delta}\right)\Phi\left(\alpha\frac{x_t-A'\phi_{t/t-1}}{\delta}\right) \quad (8)$$

$$\ni \omega_t/X_{t-1} \sim N\big(0, g_{t/t-1}\big)$$

**Normal Distribution**

$$p(x_t\,; \theta_t) = \frac{1}{\sqrt{2\pi g_{t/t-1}^2}} \frac{\left(x_t-A'\phi_{t/t-1}\right)'\left(x_t-A'\phi_{t/t-1}\right)}{2g_{t/t-1}} - \infty < x_t < +\infty \quad (9)$$

$$\ni \omega_t/X_{t-1} \sim N\big(0, g_{t/t-1}\big)$$

Its inverse, that is **Inverse Normal distribution** is

$$p(x_t\,; \theta_t) = \sqrt{\frac{\lambda}{\sqrt{2\pi x^3}}}\,exp\left[\frac{\lambda\left(x_t-A'\phi_{t/t-1}\right)'\left(x_t-A'\phi_{t/t-1}\right)}{2(A'\phi_{t/t-1})^{2x}}\right] \quad (10)$$

$\lambda$ is the shape parameter. The inverse normal distribution always works on sided tail.

**Skew-t Distribution**

To accommodate asymmetry and long tailed data, Hansen (1994) introduced the so-called skewed-t-distribution while maintaining the property of a zero mean and variance equal to one. Skew-t-distribution is derived by introducing a universalization of the Student-t distribution as follows:

$$p(x_t\,; \lambda, r) = b\frac{\Gamma\left(\frac{r+1}{2}\right)}{\Gamma\left(\frac{r}{2}\right)\sqrt{\pi(r-2)}}\left(1+\frac{\zeta^2}{r-2}\right)^{-\frac{r+1}{2}} \quad (11)$$

where,

$$\zeta = \begin{cases}(bx_t+a)/(1-\lambda) if\, x_t < -a/b\\(bx_t+a)/(1+\lambda) if\, x_t \geq -a/b\end{cases}$$

such that the constant terms “a” and “b” are defined as: $a = 4\lambda c\frac{r-2}{r-1}$; $b = 1 + 3\lambda^2 - a^2$. In this distribution, $2 < r < \infty$ denotes the degrees of freedom parameter and $-1 < \lambda < 1$ is the asymmetry parameter.

**Asymmetric Laplace**

A random variable has an Asymmetric Laplace (μ, λ, κ) Distribution (ALD) if its probability density function is

$$p(x_t\,;\mu\,,\lambda,\kappa)=\left(\frac{\lambda}{\kappa+\frac{1}{\kappa}}\right)e^{-(x_t-\mu)\lambda S\kappa^S}$$

So, $$p(x_t\,;\phi_{t/t-1},\lambda,\kappa)=\left(\frac{\lambda}{\kappa+\frac{1}{\kappa}}\right)e^{-(x_t-A'\phi_{t/t-1})\lambda S\kappa^S} \quad (12)$$

where $S=sign(x_t-\mu)$

$\mu$ is a location parameter, $\lambda>0$is a scale parameter, and $\kappa$ is an asymmetry parameter. When $\kappa=1$, $(x_t-\mu)S\kappa^S$simplifies to $|x_t-\mu|$ and the distribution simplifies to the Laplace distribution.

**Gamma Distribution**

A random variable X is said to be a Gamma distribution if:

$$p(x_t)=\left(\frac{x_t}{\beta}\right)^{\alpha-1}\times\frac{e^{\left(-\frac{x_t}{\beta}\right)}}{\beta\Gamma(\alpha)}x_t\in(0,\infty) \quad (13)$$

where $\Gamma(\alpha)=\int_0^\infty e^t t^{\alpha-1}\partial t$

with scale parameter$\beta>0$ and shape parameter$\alpha>0$.

**Exponential distribution**

A random variable X is said to be an exponential distribution (λ) if its probability density function is

$$p(x_t\,;\lambda)=\lambda e^{-\lambda x_t}x_t\in 0,\infty)$$
$$p(x_t\,;\lambda)=A'\phi_{t/t-1}e^{-A'\phi_{t/t-1}x_t} \quad (14)$$

The autoregressive score and information functions, hyper-parameters and autoregressive coefficients for the distributions specified from equation (6) to equation (14) can be estimated via the specifications made in equation (3), (4) and (5) using Maximum Likelihood (ML) or Reweighted Least Square Algorithm. See Creal *et al.* (2013), Harvey (2014), Olanrewaju & Folorunsho (2018).

## 3. Numerical Analysis

This section discusses the analyzes and results of the time-varying and time series hyper-parametric Generalized Autoregressive Scores (GASs) of the aforementioned conditional densities. The data to be subjected to the GASs with the random noise densities will be the averge monthly price of cooking gas in the United State from 1:2005 to 12:2020. The raw dataset of the price of the United State cooking gas will be used as extracted from U.S. Energy Information Administration (EIA). The monthly unit of the price is in US Dollar ($).

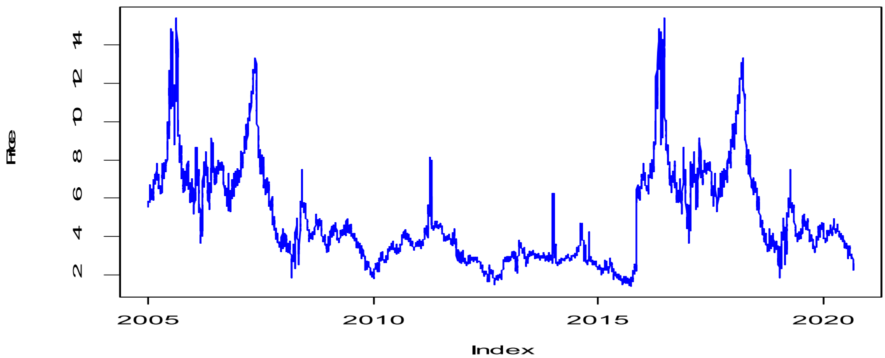


**Figure 1:** Time Plot of the Price of the Cooking Gas

From Figure 1, it is glaring that the monthly price of cooking gas in ($) was firstly pegged at around 6 ($) before skyrocketing to over 14($) towards ending of 2005 until 2006. It maintained an oscillating price between 12($) and 2($) from 2006 to 2015. The price also skyrocketed again mid-2015 to over 16($), it pendulum between 14($) and 10($) until around 2017 before a continuous drastic to 2($) was experienced. In general, from 2005 to 2020 the price of the cooking experienced a shocky zig-zag fluctuation.

**Table 1:** Coefficients of Skewness and Kurtosis

| | | | |
|---|---|---|---|
| D'Agostino Skewness test | Skew. = 1.564 | z = 29.540 | P-value < 2.2e-16 |
| Anscombe-Glynn kurtosis test | Kurt. = 5.5841 | z = 15.6419 | P-value < 2.2e-16 |
| Bonett-Seier test for Geary kurtosis | tau = 1.7997 | z = 10.6083 | p-value < 2.2e-16 |

Under the hypothesis of normality, that is under the null hypothesis that the price of the cooking gas dataset is not skewed, which is the data should be symmetry (i.e. skewness should be equal to zero). However, since the D'Agostino Skewness coefficient is 29.540 with its P-value < 2.2e-16<0.05, there is sufficient evidence that the price of the cooking gas dataset is skewed with indication that the dataset is not normally distributed (this connotes that we fail to accept the null hypothesis). In a similar vein, since the Anscombe-Glynn kurtosis coefficient of 5.5841 is far greater than three, this suggested that the price of the US cooking gas is affected by heaviness in the tail of normal distribution. In collaboration, since Geary's kurtosis coefficient of

1.7997 ≠ sqrt(2/pi) (0.7979), tailedness of the normal distribution of the price of the cooking gas data is no doubt affected. Consequently, there is a need for hyper-parameters in the conditional densities to modify the lacuna.

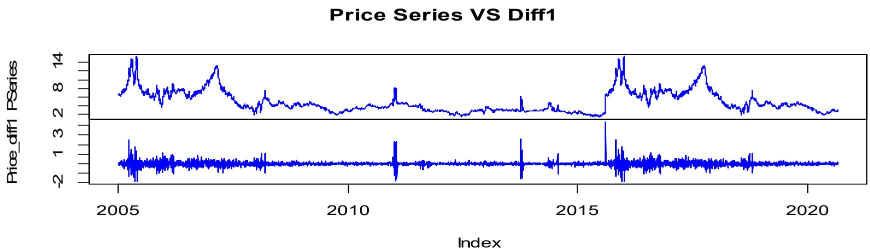


**Figure 2:** Time Plot versus the First Differencing Plot of the Price of the Cooking Gas

The upper visual time series plot in Figure 2 above is the raw plot of the price of the cooking gas in the US from 2005 to 2020, but was not stationary due to visual characterization of up and down shocks.

**Table 2:** Test of Stationarity and ARCH Effects for the Price of the US Cooking Gas

| Estimates | ADF Test Statistic | Lag | P-value | LM Statistic | LM P-value |
|---|---|---|---|---|---|
| Price Series | -1.553 | 12 | 0.674 | 32.84 | 0.005 |
| First Differencing | -50.715 | 12 | 0.01 | 26.678 | 0.0002 |

We tested the stationarity of the price of the US cooking gas via the Augmented Dickey-Fuller Test (ADF). We hypothesized both the price of the cooking series and its first differencing that their Null hypotheses display a unit root, that is both series are nonstationary. The number of lag used for testing is 12. The Test Statistic for the former was -1.553, while the latter gave -50.715. Since the p-value for the latter (first differencing) is 0.01 and the only one less than 0.05. We concluded that there is enough evidence to reject the Null hypothesis, meaning that the first differencing of the price series is the only one that is stationary. We also tested for Autoregressive Conditional Heteroskedasticity (ARCH) in order to ascertain if conditional variance on the information exists at a given point in time for both price of the cooking series and its first differencing. The formulated Null hypothesis for both series was there are no ARCH effects. Since the p-values for both the latter and former are less than 5% level of significance, the Null hypotheses are rejected and it is concluded that both series possessed ARCH effects. The first differencing series of the price of the cooking gas was used to model time-varying hyper-parameters for the GAS model because of its stationarity.

**Table 3:** Model Adequacy of the Density GAS w.r.t to the Price of United State Cooking Gas.

| Specification | ddist | Pdist | qdist | Location-Score | Scale-Score | Skewness-Score | Shape1-Score | Shape-2 Score | Informatin-Location | Information -Scale | Information -Skewness | Information -Shape | Information -Shape2 | AIC | BIC |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Normal** | -3.2807 | 0.0000 | 5.1845 | -1.0278 | 0.3313 | ----- | ----- | ----- | 0.4275 | 0.9941 | ----- | ----- | ----- | 58242.958 | 58247.334 |
| **Inv. Normal** | -3.6068 | 0.0026 | 1.1101 | -0.8511 | 0.2783 | ----- | ----- | ----- | 0.1676 | 0.0140 | ----- | ----- | ----- | 12958.755 | 12994.380 |
| **Inv. Sqrt. Normal** | -3.6068 | 0.0026 | 1.1101 | -0.8511 | 0.2783 | ----- | ----- | ----- | 0.1676 | 0.0140 | ----- | ----- | ----- | 12958.755 | 12994.380 |
| **Skewed Normal** | -6.2708 | 0.0050 | 2.7476 | -2.3015 | 2.4255 | 0.0000 | ----- | ----- | 0.4460 | 0.0995 | 0.0000 | ----- | ----- | 12304.230 | 12357.666 |
| **Student-t** | -3.9999 | 0.0004 | 1.5743 | -1.2635 | 0.6636 | 0.0000 | ----- | ----- | 0.2691 | 0.0362 | 0.0000 | | | 12786.397 | 12839.833 |
| **Student-t Inv.** | -3.9999 | 0.0004 | 1.5743 | -1.2635 | 0.6636 | 0.0000 | ----- | ----- | 0.2691 | 0.0362 | 0.0000 | ----- | ----- | 12786.397 | 12839.833 |
| **Student-t Inv.Sqrt.** | -3.9999 | 0.0004 | 1.5743 | -1.2635 | 0.6636 | 0.0000 | ----- | ----- | 0.2691 | 0.0362 | 0.0000 | ---- | ----- | 12786.397 | 12839.833 |
| **SkewStudent-t** | -3.9999 | 0.0004 | 1.5743 | -1.2635 | 0.6636 | 0.0000 | 0.0000 | ----- | 0.2691 | 0.0362 | 0.0000 | 0.0000 | ----- | 12198.097 | 12269.346 |
| **AST** | -3.9999 | 0.0004 | 1.5743 | -1.2635 | 0.6636 | 0.0000 | 0.0000 | 0.0000 | 0.2691 | 0.0362 | 0.0000 | 0.0000 | 0.0000 | 11983.633 | 12072.693 |
| **AST Inv.** | -3.9999 | 0.0004 | 1.5743 | -1.2634 | 0.6636 | 0.0000 | 0.0000 | 0.0000 | 0.2691 | 0.0362 | 0.0000 | 0.0000 | 0.0000 | 11983.633 | 12072.693 |
| **AST Inv. Sqrt.** | -3.9999 | 0.0004 | 1.5743 | -1.2635 | 0.6636 | 0.0000 | 0.0000 | 0.0000 | 0.2691 | 0.0362 | 0.0000 | 0.0000 | 0.0000 | 11983.633 | 12072.693 |
| **AST1 (Identitiy)** | -3.9999 | 0.0004 | 1.5743 | -1.2634 | 0.6636 | ----- | ----- | ----- | 0.2691 | 0.0362 | ----- | ----- | ----- | **11943.277*** ** | **12014.525*** ** |
| **ALD(Identity)** | -2.8941 | 0.0001 | 0.7969 | -1.8781 | 1.4196 | ----- | ----- | ----- | 0.6879 | 0.2366 | ----- | ----- | ----- | 11992.746 | 12046.182 |
| **Gamma (Identity)** | -11.626 | 0.0192 | 3.4824 | -5.0779 | 12.392 | ----- | ----- | ----- | 1.0002 | 0.5002 | ----- | ----- | ----- | 12153.334 | 12188.958 |
| **Exponential** | -1.7321 | 0.0000 | 0.0100 | 5.0779 | ----- | ----- | ----- | ----- | 26.2950 | ----- | ----- | ----- | ----- | 14760.256 | 14778.068 |
| **Negative - Binomial** | 0.0000 | 0.0000 | 0.0000 | 4.6542 | -0.1386 | ----- | ----- | ----- | 41.4354 | 1.0000 | ----- | ----- | ----- | 44052.93 | 44056.91 |
| **Skellam** | -1.6243 | 0.1538 | 0.0000 | -0.1655 | -0.1129 | ----- | ----- | ----- | 1.0000 | 1.0000 | ----- | ----- | ----- | 12510.243 | 12545.867 |

**Table 4:** Model Adequacy of GARCH (1,1) Model w.r.t to the Price of United State Cooking Gas.

| Specification | Estimate | Std. Error | t-value | Pr(>\|t\|) | Information Criteria | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | AIC | BIC | Shibata | Hannan-Quinn |
| **Omega** | 0.4642 | 0.2905 | 1.5978 | 0.0101 | 22235.08 | 22237.99 | 22235.94 | 22236.034 |
| **Alpha1** | 0.9801 | 0.0684 | 14.3364 | 0.00000 | | | | |
| **Beta1** | 0.0000 | 0.0630 | 0.0001 | 0.9999 | | | | |

Description of each conditional density with respect to GAS or Dynamic Conditioal Correlation (DCC) was explicitly tabled in Table 1 and Table 2 (Table 1A in appendix). DCC is one of the most famous models for multivariate volatility. It uses multivariate GAS to model and analyze volatilities when the framework is based on score-driven time series for time-varying parameters. The model summary includes for each density of GAS includes their long-term value of the time-varying hyper-parameters, their estimated score and Hessian values, their model performance and concerned estimated autoregressive coefficients. Among the ten(10) hyper-parametric conditional noises that were subjected to the GAS model via the application of the price of the cooking gas, Asymmetric Student–t with one tail decay parameter (AST1) outperformed other category of its variants as well as other conditional densities for the GAS/DCC model with the minimum reduced-error performance of AIC=11943.277 and BIC=12014.525. The model hyper-parametric scores for the location-score and scale-score are -1.2634 and 0.6636 respectively. Its location-information and scale- information are 0.2691 and 0.0362 respectively. The concerned estimated coefficients of $kappa_1$, $kappa_2$, $kappa_3$ and $kappa_4$ are the elements of vector $\eta$ i.e. $\eta_{\mu}$, $\eta_{\phi}$, $\eta_{g_t}$,$\eta_{v_t}$which are 0.1444, 0.7434, -10195 and -0.8836 respectively. Analogously, $a_1$, $a_2$, $a_3$, $a_4$ are estimates of $a_{\mu}$, $a_{\phi}$, $a_{g_t}$,$b_{v_t}$ with 0.0000, 0.0000, 0.0000, and 0.0000 respectively, similarly to that of $b_1$, $b_2$, $b_3$, b with 0.9468, 0.5166, 0.1566 and 0.5993 respectively.

In comparison of the GAS or DCC model with the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model, it was affirmed that model GARCH(1,1) was the optimal lag (that is both Autoregressive (AR) and Moving Average (MA) at lag 1 each) for the volatiled price of the United State cooking gas studied over the period of time. It is to be noted that the GARCH model was subjected to different distributional error noises, like Student-t, Gaussian, Skew-Normal, etc., but Gaussian noise gave a robust generalization. The estimated model criteria of AIC, BIC, Shibata and Hannan-Quinn of 22235.08, 22237.99, 22235.94, and 22236.034 respectively for the model performance of the GARCH (1,1) model were far below the model performance of the GAS or DCC model via the random noise of the Asymmetric

Student–t with one tail decay parameter (AST1). The robustness of the GAS model via the Asymmetric Student–t with one tail decay parameter (AST1) might be via the location and scale scores of the noise.

## 4. Conclusions

This article introduced the possible conditional densities for the Generalized Autoregressive Score (GAS) model with embedded time-varying hyper-parameters. The score and Hessian functions (via location, scale, skewness, and shapes parameters) are of paramount interest due to their capability to curtail the lacuna of heaviness in the tail of normal distribution and possibility of skewed observations. Due to the flexibility of the GAS model to several statistical distributions, an empirical application to financial data of the price of the United State cooking gas was subjected to the GAS model with ten (10) different conditional densities. Each of the conditional density subjected to the GAS model via the application of the price of cooking gas from 2005 to 2020 was driven by the mechanism of time-varying score and Hessian functions of their embedded hyper-parameters. Asymmetric Student–t with one tail decay parameter (AST1) outperformed other category of its variants as well as other re-parameterized distributions used. In addition, the GAS model via Asymmetric Student–t with one tail decay parameter (AST1) random noise outshined the core volatile conditional heteroscedasticity of Generalized Autoregressive Conditional Heteroscedasticity (GARCH) with Gaussian distributed noise. For further studies, the conditional densities of the GAS model might be subjected to a driven mechanism of family of distributions with strictly positive values or integer values.

## Acknowledgement

The authors personally extend gratitude to the US Energy Information Administration (EIA).

## References

Blasques, F., Koopman, S. J., Łasak, K., Lucas, A., (2016). In-Sample Confidence Bands and Out-of-Sample Forecast Bands for Time-Varying Parameters in Observation-Driven Models. *International Journal of Forecasting*, Vol. 32(3), pp. 875–887, doi: 10.1016/j.ijforecast.2015. 11.018.

Cox, D. R., (1981). Statistical analysis of time series: some recent developments. *Scandinavian Journal of Statistics*, Vol. 8(2), pp. 93–115.

Creal, D., Koopman, S.J., Lucas, A., (2013). Generalized Autoregressive Score Models with Applications. *Journal of Applied Econometric*, Vol. 28(5), pp. 777–795. https://doi.org/10.1002/jae.1279.

Janus, P., Koopman, S. J., Lucas, A., (2014). Long Memory Dynamics for Multivariate Dependence under Heavy Tails. *Journal of Empirical Finance*, Vol. 29, pp. 187–206, doi: 10.1016/j.jempfin, 2014.09.07.

Engle, R. F., (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica,* Vol. 50(4), pp. 987–1007.

Hansen, B. E., (1994). Autoregressive conditional density estimation. *International Economic Review*, Vol. 35 (3), pp. 705–729.

Harvey, A. C, Thiele, S. (2016). Testing Against Changing Correlation. *Journal of Empirical Finance*, Vol. 38(B), pp. 575–589. doi:10.1016/j.jempfin.2015.09.003.

Harvey, A. C., Sucarrat, G., (2014). EGARCH Models with Fat Tails, Skewness, and Leverage. *Computational Statistics & Data Analysis*, Vol.76, pp. 320–338, doi: 10.1016/j.csda.2013.09. 022.

Harvey, A. C., Luati, A., (2014). Filtering with Heavy Tails. *Journal of the American Statistical Association*, Vol. 109(507), pp. 1112–1122, doi: 10.1080/ 01621459.2014.887011.

Harvey, A. C., (2013). *Dynamic Models for Volatility and Heavy Tails: With Applications to Financial and Economic Time Series*. Cambridge University Press.

Koopman, S. J., Lucas, A., Monteiro, A., (2008). The multi-state latent factor intensity model for credit rating transitions. *Journal of Econometrics,* Vol. 142(1), pp. 399–424.

Owen, C. B., (2008). *Parameter Estimation for the Beta Distribution*. Brigham Young University – Provo.

Jones, M. C., Faddy, M. J., (2003). A skew extension of the student t-distribution, with applications. *Journal of Royal Statistics Society B*,. Vol 65, Part 1, pp. 159–174.

Lucas, A., Zhang, X., (2016). Score-Driven Exponentially Weighted Moving Averages and Value-at-Risk Forecasting. *International Journal of Forecasting*, Vol. 32(2), pp. 293–302, doi: 10.1016/ j.ijforecast.2015.09.003.

Oh, D. H., Patton, A. J., (2016). Time-Varying Systemic Risk: Evidence from a Dynamic Copula Model of CDS Spreads. *Journal of Business* & *Economic Statistics,* Vol. 36(2), pp. 181–195, doi: 10.1080/07350015.2016.1177535.

Olanrewaju, R. O., Ojo J. F., Adekola, L. O., (2020). Bayesian latent autoregressive stochastic volatility: an application of naira to eleven exchangeable currencies rates. *Open Journal of Mathematical Sciences*, Vol.4 (1), pp. 386–396, doi: 10.30538/oms2020.0128.

Olanrewaju, R. O., Folorunsho, S. A., (2018). Generalized autoregressive score (GAS) functions under Gaussian and Student-t distributions. *International Journal of Statistics and Applied Mathematics*, Vol. 3(5), Part A, pp. 56–61.

Monache, D. D., Petrella, I., (2014). *Adaptive Models and Heavy Tails*. School of Economics and Finance, Working Paper No. 720, ISSN 1473-0278.

Shephard, N., (2005). *Stochastic Volatility: Selected Readings*. Oxford University Press, Oxford.

# APPENDICES

## Appendix 1.

**Table 1A:** Coefficients of the Hyper-Parameterization of the Generalized Autoregressive Scores (GASs)

| $a_3$ | $a_4$ | $a_5$ | $b_1$ | $b_2$ | $b_3$ | $b_4$ | $b_5$ | Location | Scale | Skewness | Shape1 | Shape2 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ----- | ----- | ----- | 0.9789 (0.0009) | 0.8686 (0.0042) | ----- | ----- | ----- | 5.4877 | 0.0334 | ----- | ----- | ----- |
| ----- | ----- | ----- | 0.5000 (4.4854) | 0.4999 (4.4967) | ----- | ----- | ----- | 5.1279 | 5.9665 | ----- | ----- | ----- |
| ----- | ----- | ----- | 0.5000 (4.4854) | 0.4999 (4.4967) | ---- | ----- | ----- | 5.1279 | 5.9665 | ----- | ----- | ----- |
| 0.000001 (0.0000) | ----- | ----- | 0.8807 (0.0000) | 0.5166 (0.0000) | 0.5166 (0.000) | ----- | ----- | 5.2107 | 2.2423 | 1.5000 | ----- | ----- |
| 0.000001 (0.0000) | ----- | ----- | 0.8807 (2.2661) | 0.5662 (0.0000) | 0.56620 (3.7787) | ----- | ----- | 4.7451 | 3.7160 | 5.0892 | ----- | ----- |
| 0.000001 (0.0000) | ----- | ----- | 0.8807 (2.2661) | 0.5662 (3.7788) | 0.5662 (3.7788) | ----- | ----- | 4.7451 | 3.7160 | ----- | 5.0892 Realized=1.85 | ----- |
| 0.000001 (0.0000) | | ----- | 0.8807 (2.2661) | 0.5662 (0.0000) | 0.5662 (3.7788) | ----- | ----- | 4.7451 | 3.7160 | ----- | 5.0892 | ----- |
| 0.000001 (0.0000) | 0.000001 (0.0000) | ----- | 0.7648 (0.0000) | 0.5166 (2.0346) | 0.5165 (0.3261) | 0.5166 (1.0512) | | 4.9908 | 2.3230 | 1.4999 | 5.7047 | |
| 0.000001 (0.0000) | 0.000001 (0.0000) | 0.0000 (0.0000) | 0.8999 (0.0000) | 0.9138 (0.5077) | 0.9138 (1.3458) | 0.9138 (0.0000) | 0.9138 (0.0000) | 2.7768 | 4.3048 | 0.1259 | 4.0343 | 4.0001 Realized=1.85 |
| 0.000001 (0.0000) | 0.000001 (0.0000) | 0.000001 (0.0000) | 0.8999 (0.0000) | 0.9138 (0.5077) | 0.9138 (1.3457) | 0.9138 (2.5291) | 0.9138 (3.4201) | 2.7768 | 4.3047 | 0.1258 | 4.0342 | 4.0000 |
| 0.000001 (0.0000) | 0.000001 (0.0000) | 0.000001 (0.0000) | 0.8999 (0.0000) | 0.9138 (0.5077) | 0.9138 (1.3457) | 0.9138 (1.1827) | 0.9138 (3.7102) | 2.7768 | 4.3047 | 0.1258 | 4.0342 | 4.0000 |
| **0.000001 (0.0000)** | **0.000001 (0.0000)** | ----- | **0.9468 (0.0129)** | **0.5166 (0.0008)** | **0.5166 (0.0050)** | **0.5993 (0.01431)** | ----- | **2.7203** | **4.6539** | **0.1161** | **8.5669** | ----- |
| 0.000001 (0.0000) | ----- | ----- | 0.8476 (0.0009) | 0.5331 (5.3925) | 0.5496 (0.0000) | ----- | ----- | 2.7800 | 1.4536 | 0.3758 | ----- | ----- |
| ----- | ----- | ----- | 3.3109 (0.0000) | 3.3109 (0.0000) | ----- | ----- | ----- | ----- | 5.1271 | 0.9998 | ----- | ----- |
| ----- | ----- | ----- | 0.90000 (0.1478) | ----- | ----- | ----- | ----- | 0.1950 | ----- | ----- | ----- | ----- |
| ----- | ----- | ----- | 0.90000 (0.0000) | 0.90000 (0.0000) | ----- | ----- | ----- | 0.8593 | 31.3185 | ----- | ----- | ----- |
| | | | 0.5829 (0.0002) | 0.5131 (0.0002) | | | | 5.1288 | 5.1333 | | | |

## Appendix 2.

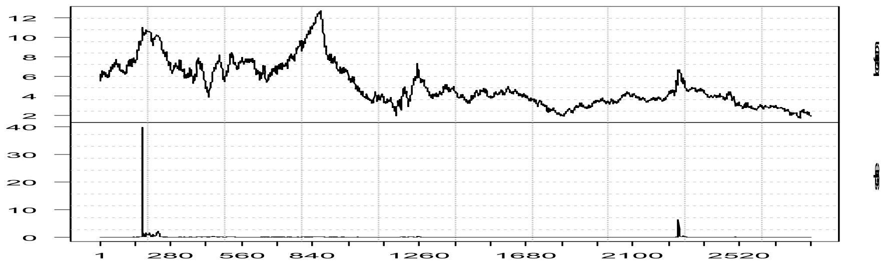


**Figure 1A:** Graphical Plot of the AST1 Location and Scale Parameters.

## Appendix 3.

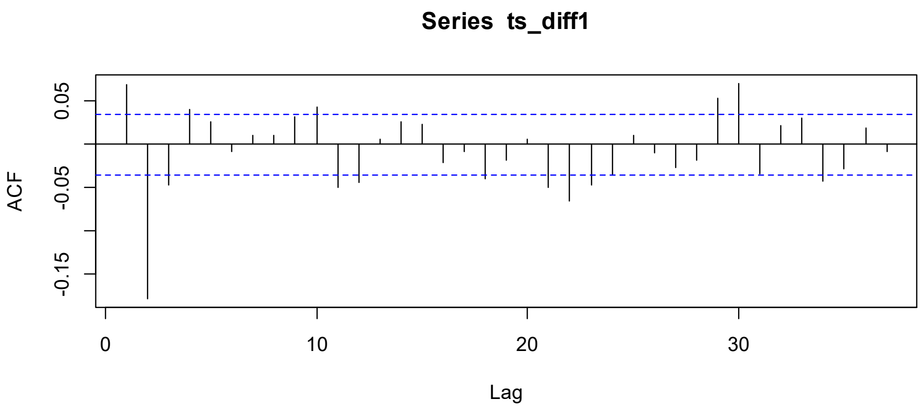


**Figure A2:** ACF Graph of the First Differencing of the Price of the US Cooking GAS