ASSESSING THE EFFICIENCY OF CLASSICAL AND BAYESIAN APPROACHES IN ADDRESSING HETEROSCEDASTICITY UNDER KNOWN FUNCTIONAL FORMS.

[1]Oluwaseun Ayobami Adesina, [2]Taiwo Abideen Lasisi and [3]Temitayo Stephen Fadare
[1,2&3]Department of Statistics, Ladoke Akintola University of Technology, Ogbomoso, Nigeria

(e-mail:[1]oaadesina26@lautech.edu.ng; [2]talasisi26@lautech.ced.ng and [3]fadaretemitayo2@gmail.com;)

Corresponding Author[1,*]
Tel: +2348033971773 Email: *oaadesina26@lautech.edu.ng*

Abstract

This study provides a robust comparison of the traditional and Hierarchical Bayesian approaches for addressing heteroscedasticity, evaluated under known functional forms where the variance of errors is modeled as a function of exogenous variables. Using simulated data generated through Gibbs Sampling in a Monte Carlo framework, the study examines the performance of hierarchical Bayesian (HB), ordinary least squares (OLS), and generalized least squares (GLS) approaches across different sample sizes and replications. The findings indicate that the HB demonstrates superior efficiency in addressing heteroscedasticity compared to the traditional approaches, consistently outperforming them across various scenarios. These results underscore the advantage of the HB approach in modeling relationships involving predictor variables and a dependent variable exhibiting heteroscedasticity, offering a robust alternative for researchers and practitioners.

Keywords: Efficiency, Heteroscedasticity, Classical approach, Hierarchical Bayesian approach, Monte Carlo.

ISSN NUMBER: 1116-249X

1. INTRODUCTION

In the linear regression framework, the ordinary least squares (OLS) estimator is prominently considered under several assumptions; among which include the assumption of homoscedasticity (constant variances). It is generally known that ordinary least squares (OLS) deliver accurate and unbiased estimates of the parameters when the linear regression model's underlying assumptions are true. If the error term has non-constant variance, then the best linear unbiased estimator (BLUE) is the generalized least squares (GLS) estimator The GLS estimator is an extension of the ordinary least squares (OLS) estimator, which is used when the errors have a scalar variance-covariance matrix (homoscedastic and uncorrelated). The GLS estimator is more efficient than OLS in the presence of heteroscedasticity as it takes into account the structure of the errors which is also called the weighted least squares (WLS) estimator. Essentially, least squares assumes that the variance of the error term is constant and independent or serially uncorrelated. Specifically, the efficiency of the least squares estimators comes to bear whenever the assumption of constant error variance (homoscedasticity) is met (Wooldridge, 2010).

A notable limitation of Generalized Least Squares (GLS) estimators is that they can perform poorly in finite samples if the conditional variance model is mis-specified or estimated with significant error (Angrist and Pischke, 2008). In particular, the weighted estimators may even underperform their unweighted counterparts. This is because accurately modeling the variance function can be challenging, which may render Feasible GLS (FGLS) efforts ineffective. However, not all researchers agree with this assessment.

Leamer (2010) contends that researchers should be working to model the heteroscedasticity in order to determine whether sensible reweighting affects the estimates. In the context of a random sample for which only heteroscedasticity is a concern, Romano and Wolf (2017) shows that GLS

ISSN NUMBER: 1116-249X

can improve the efficiency of the estimates of the relevant parameters even when the scedastic function of the errors is mis-specified, though it is assumed that the correct covariates are used. Even when the true form of the heteroscedasticity is unknown, HR standard errors can be used to base valid inference; e.g. DiCiccio *et al* (2019).

In recent times, the application of Bayesian principles in econometrics has witnessed tremendous growth (Geweke, 2010). The principle is based on a degree-of-belief interpretation of probability contrary to the relative-frequency interpretation of the classical methods (Koop, 2003). The Bayesian principle assumes that coefficients and covariance matrix of the normal linear regression model (NLRM) have prior distributions (Li, 2018).

Bayesian procedures usually behave well in small samples. Thus, the Bayesian normal linear regression with nonparametric heteroscedasticity can also be an attractive alternative to classical semi-parametrically efficient estimators from Carroll (1982) and Robinson (1987).

Lancaster (2003) and Poirier (2011) do not assume linearity of the regression function and treat the linear projection coefficients as the parameters of interest. Rubin, (1981) uses Bayesian bootstrap to justify from the Bayesian perspective the use of the ordinary least square estimator with a heteroscedasticity robust covariance matrix. Pelenis (2014) demonstrates posterior consistency in a semiparametric model with a parametric specification for the regression function and a nonparametric specification for the conditional distribution of the regression error term.

Several different approaches to inference in a regression model have been proposed in the Bayesian framework. In a standard textbook linear regression model, normality of the error terms is assumed. More recent literature relaxed the normality assumption by using mixtures of normal or Student t-distributions. However, if the shape of the error distribution depends on covariates then

ISSN NUMBER: 1116-249X

the posterior may not concentrate around the data generating values of the linear coefficients (Müller, 2013).

Salois and Balcombe (2014) uses potential endogeneity of participation in Supplemental Nutrition Assistance Program (SNAP) as a potential problem in investigating its causal influence on obesity using instrumental variable (IV) approaches and due to the presence of heteroscedasticity in the errors, the approach for dealing with heteroscedastic errors in Geweke (1993) is extended to the Bayesian instrumental variable estimator outlined in Rossi *et al*. (2005).

Oseni *et al* (2019) derived Bayesian estimators of the NLRM in the presence of functional forms of heteroscedasticity. The Bayesian principle assumes that coefficients and covariance matrix of the NLRM have prior distributions. This approach is very attractive to applied econometricians because it combines out-of-sample information with observed data. Estimation of a NLRM using the Bayesian approach in the presence of heteroscedasticity is a relatively new area being explored in the econometric literature. Variance was treated as a linear function and as an exponential function of exogenous variables. The estimators are found to be unbiased and consistent and the precision values tend to zero. The estimates obtained from the estimators approximately 95% draws fall within each of the corresponding credible interval.

Hierarchical Bayesian estimation is a multi-level analysis which brings about flexibility in parameter estimation. The hierarchical Bayesian technique is premised on assuming that hierarchical prior distributions are independently drawn from the same distribution with unknown parameters.

Hierarchical Bayesian (HB) is of greater advantage in improving estimates for groups with limited data by shrinking individual estimates toward the group mean (shrinkage effect). HB is helpful in handling complex, nested structures and models the heterogeneity between subgroups more

ISSN NUMBER: 1116-249X

naturally than flat models. HB also allows for uncertainty quantification at multiple levels (subgroup and global levels).

Akinlade *et al* (2021) uses Hierarchical Bayesian Estimation (HBE) of unobserved individual heterogeneity of dynamic panel models to improve on a static panel model even for a panel with small, moderate, and large N. Three experiments for the individual (N) and time (T) were considered: (10, 15), (20, 20), and (100, 15). Theoretical findings are accompanied by extensive Markov Chain Monte Carlo experiments, which show that the estimator performs well and handled the complicated pattern exhibited by the data.

This paper therefore investigates the performance comparison of the Ordinary Least Squares (OLS) and Generalized Least Square (GLS) of the classical approaches, and hierarchical structures of the Bayesian approach of correcting for heteroscedasticity when the functional form is known (Quadratic) using random normal simulated data.

Results obtained are of importance to researchers and practitioners in modeling complex relationships between the predictor variables and the dependent variable with the known functional form of heteroscedasticity.

2. METHODOLOGY

2.1 Ordinary Least Squares (OLS)

The specification of the linear regression model is given by

$$y = X\beta + U \tag{1}$$

where

ISSN NUMBER: 1116-249X

y is the dependent variable defined as $y = \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ y_n \end{bmatrix}$

X is the independent variable defined as $X = \begin{Bmatrix} x_{11} & x_{12} \cdots & x_{1k} \\ x_{21} & x_{22} \cdots & x_{2k} \\ x_{31} & x_{32} \cdots & x_{3k} \end{Bmatrix}$

β is the parameter estimate defined as $\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ . \\ . \\ \beta_k \end{bmatrix}$

U is the error term defined as $U = \begin{bmatrix} u_1 \\ u_2 \\ . \\ . \\ u_n \end{bmatrix}$

The least square estimator of the linear regression model seeks to minimize the residual sum of squares in the model

$$SSE = \sum_{i=1}^{n}\left(y_i - X_i\hat{\beta}\right)^2 \quad (2)$$

The estimated vector $\hat{\beta}$ “that minimizes the objective function β is obtained by taking the derivative of equation with respect to β, setting it equal to zero, and solving for $\hat{\beta}$."

$$\hat{\beta} = (X^T X)^{-1} X^T y \quad (3)$$

And the estimated value of s^2 is computed by

$$s^2 = \frac{\left(y - X\hat{\beta}\right)^{-1}\left(y - X\hat{\beta}\right)}{n-k} \quad (4)$$

where v = n - k interpreted as the degrees of freedom

ISSN NUMBER: 1116-249X

2,2 Generalized Least Squares

Generalized Least Squares Estimation (GLSE) is a statistical technique used to estimate the parameters of a regression model in the presence of non-constant variance (heteroscedasticity). Unlike ordinary least squares (OLS), which assumes that the variance of the errors is constant, GLSE accounts for varying error variances across different observations. This is particularly important because heteroscedasticity can lead to inefficient estimates and biased inference results. In GLSE, weights are assigned to each observation based on the estimated error variances, allowing for more reliable parameter estimates.

The linear regression model with three (3) explanatory variables when weighted becomes:

$$\boldsymbol{w_t} Y_t = w_t\beta_1 + \beta_2(w_t X_{12}) + \beta_3(w_t X_{13}) + w_t\mu_t \tag{5}$$

where

$w_t = \frac{1}{\sigma_t^2}$, $W^{-1} = \sigma^2(\varepsilon)$

The variance covariance matrix is

$$\sigma^2(\varepsilon) = \begin{pmatrix} \sigma_1^2 & 0 \cdots & 0 \\ 0 & \sigma_2^2 \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & \sigma_3^2 \end{pmatrix}$$

The weighted estimator is then given as:

$$\hat{\beta} = (X'W\,X)^{-1}X'WY \tag{6}$$

2.3 Bayesian Approach

The Markov Chain Monte Carlo (MCMC) simulation method is utilized through the Gibbs Sampler Algorithm, configured with an independent normal-gamma prior. The hierarchical prior is also normal -gamma, derived as a conjugate prior from the posterior distribution. This posterior

ISSN NUMBER: 1116-249X

is obtained by combining the likelihood function, incorporating a quadratic functional form of heteroscedasticity, with the prior.

2.3.1 The Likelihood Function

The likelihood function for this model $y = X\beta + U$ using multivariate Normal density when the variance differs across observations is given as

$$P\left(y \mid \beta, h, \Omega\right)\} = \frac{h^{\frac{N}{2}}}{(2\pi)^{\frac{N}{2}}} \left|\Omega\right|^{\frac{1}{2}} \{\exp[-\frac{h}{2}(Y - X\beta)'\Omega^{-1}(Y - X\beta)]\} \tag{7}$$

Rewriting (7) in multiple form gives:

$$P\left(y / \beta, h, \Omega\right) = \frac{h^{\frac{N}{2}}}{(2\pi)^{\frac{N}{2}}} \{\exp[-\frac{h}{2}((n-k)s^2 + (\hat{\beta} - \beta)' X' \Omega^{-1} X(\hat{\beta} - \beta))]\} \tag{8}$$

where

$$\hat{\beta} = (X^{'}\Omega X)^{-1} X^{'}\Omega^{-1} Y \tag{9}$$

$$s^2 = \frac{(Y - X\beta)^{'}\Omega(Y - X\beta)}{n-k} \tag{10}$$

$$s^2 = \frac{(Y - X\beta)^{'}\Omega(Y - X\beta)}{N-k} \tag{10}$$

Thus,

$$(N-k)s^2 = (Y - X\beta)'\Omega(Y - X\beta) \tag{11}$$

Setting $v = N - k$, which is interpreted as the degrees of freedom in the above to have

$$P\left(y / \beta, h, \Omega\right) = \frac{h^{\frac{N}{2}}}{(2\pi)^{\frac{N}{2}}} \{\exp[-\frac{h}{2}(vs^2 + (\hat{\beta} - \beta)' X' \Omega^{-1} X(\hat{\beta} - \beta))]\} \tag{12}$$

If (12) is partitioned by using $N = v + k,$ the likelihood function becomes

ISSN NUMBER: 1116-249X

$$P\left(y \mid \beta, h, \Omega\right) = \frac{1}{(2\pi)^{\frac{N}{2}}} \{h^{\frac{k}{2}} \exp[-\frac{h}{2}(\hat{\beta} - \beta(\Omega))' X' \Omega^{-1} X(\hat{\beta} - \beta(\Omega))] \left\{h^{\frac{v}{2}} \exp\left[\frac{-hv}{2s^{-2}(\Omega)}\right]\right\}. \quad (13)$$

The quantity $\{h^{\frac{k}{2}} \exp[-\frac{h}{2}(\hat{\beta} - \beta(\Omega))' X' \Omega^{-1} X(\beta - \beta(\Omega))]\}$ in (13) resembles the kernel of the multivariate Gaussian density while $\left\{h^{\frac{v}{2}} \exp\left[\frac{-hv}{2s^{-2}(\Omega)}\right]\right\}$ also looks like the kernel of the gamma density. These results simply suggest a normal-gamma prior for the likelihood function (Koop, 2003).

2.3.2 The Prior

This is the information at hand about a particular study before seeing the data, we denote the independent prior by $P\left(\beta, h\right)$

In the independent random variables its follows that,

$P\left(\beta, h\right) = P\left(\beta\right) \cdot P\left(h\right)$ with $P\left(\beta\right)$ being Normal and $P\left(h\right)$ being Gamma:

The likelihood in (11) suggests that Normal-Gamma prior are appropriate for the parameters β and h in this study.

Prior for β is of the form:

$$P(\beta) = \frac{h^{\frac{k}{2}}}{(2\pi)^{\frac{k}{2}} |\Omega_0|^{\frac{1}{2}}} \left\{exp\left[-\frac{1}{2}\left(\hat{\beta} - \beta_0\right)(\Omega_0)^{-1}\left(\hat{\beta} - \beta_0\right)\right]\right\} \quad (14)$$

and prior for h is of the form

$$p(h) = \frac{1}{\Gamma\left(\frac{v_0}{2}\right)\left(\frac{2s_0^{-2}}{v_0}\right)^{\frac{v_0}{2}}} \left\{h^{\frac{v_0-2}{2}} exp\left[\frac{hv_0}{2s_0^{-2}}\right]\right\} \quad (15)$$

ISSN NUMBER: 1116-249X

where, β_0 and $\frac{1}{\Gamma\left(\frac{v_0}{2}\right)\left(\frac{2s_0^{-2}}{v_0}\right)^{\frac{v_0}{2}}}$ in (14) and (15) are the priors for β and integrating constant respectively.

So that the joint prior for β and h then becomes

$$p(\beta \mid h) = \frac{h^{\frac{v_0-2}{2}-1}}{(2\pi)^{\frac{k}{2}}|\Omega_0|^{\frac{1}{2}}\Gamma\left(\frac{v_0}{2}\right)\left(\frac{2s_0^{-2}}{v_0}\right)^{\frac{v_0}{2}}}\left\{exp\left[-\frac{1}{2}\left(\hat{\beta}-\beta_0\right)(\Omega_0)^{-1}\left(\hat{\beta}-\beta_0\right)+\frac{v_0}{s_0^{-2}}\right]\right\} \tag{16}$$

The compact form of (15) can be expressed as:

$$p(\beta \mid h) = f_{NG}(\beta, h \mid \beta_0, s_0^{-2}, v_0) \tag{17}$$

Finally, the specification for non-informative uniform prior, is $p(\Omega_0) \propto 1$

2.3.3 The Posterior distribution

The posterior distribution can be obtain by combining the likelihood function in (13) and the prior distributions in (16)

Then, from the joint density $p(\beta, h \mid y)$ is given by

$$p(\beta, h \mid y) = p(y \mid \beta, h)\, p(\beta, h) \tag{18}$$

which becomes

$$p(\beta, h, \Omega \mid y) = \frac{1}{(2\pi)^{\frac{N}{2}}}\{h^{\frac{k}{2}}\exp[-\frac{h}{2}(\hat{\beta}-\beta(\Omega))'X'\Omega^{-1}X(\hat{\beta}-\beta(\Omega))]\left\{h^{\frac{v}{2}}\exp\left[\frac{-hv}{2s^{-2}(\Omega)}\right]\right\}.$$

$$\times \frac{h^{\frac{v_0-2}{2}-1}}{(2\pi)^{\frac{k}{2}}|\Omega_0|^{\frac{1}{2}}\Gamma\left(\frac{v_0}{2}\right)\left(\frac{2s_0^{-2}}{v_0}\right)^{\frac{v_0}{2}}}\left\{exp\left[-\frac{1}{2}\left(\hat{\beta}-\beta_0\right)(\Omega_0)^{-1}\left(\hat{\beta}-\beta_0\right)+\frac{v_0}{s_0^{-2}}\right]\right\} \tag{19}$$

From the joint posterior distributions in (18), the following conditional densities were obtained.

(i) The conditional posterior density of β:

$$p(\beta \mid h, \Omega, y) = N(\beta_n, \Omega_n) \tag{20}$$

ISSN NUMBER: 1116-249X

where

$$\beta_n = \Omega_n\left[\Omega_0^{-1}\beta_0 + hX'\Omega_0^{-1}X\hat{\beta}_{GLS}\right] \quad (21)$$

$$\Omega_n = [\Omega_0^{-1} + hX'\Omega_0^{-1}X]^{-1} \quad (22)$$

(ii) The conditional posterior density of h is;

$$p(h, \beta, \Omega, y) = G[s_0^{-2} . v_0] \quad (23)$$

where

$$s_0^{-2} = \frac{v_0}{(y - X\beta)'\Omega^{-1}(y - X\beta) + v_0 s_0^2} \quad (24)$$

and

$$v_n = N + v_0 \quad (25)$$

2.3.4 Hierarchical Structures

The combination of the likelihood in (13) and the prior in (16) gives a Normal-Gamma distribution for the posterior which is the conjugate prior that suggests the distribution of the hierarchical prior and hierarchical posterior to be normal-gamma in the form of the hyper-parameters given below

2.3.5 Hierarchical Prior

$$\mu_\alpha \sim N(\underline{\mu}_\alpha, \underline{\sigma}_\alpha^2)$$

and

$$V_\alpha^{-1} \sim G(\underline{V}_\alpha^{-1}, \underline{v}_\alpha)$$

2.3.6 Hierarchical Posterior

$$\hat{\beta} \,/\, y, h, \alpha, \mu_\alpha, V_\alpha \sim N(\hat{\beta}, \bar{V}_\beta)$$

and

$$h \,/\, y, \hat{\beta}, \alpha, \mu_\alpha, V_\alpha \sim G(S^{-2}, \bar{v})$$

ISSN NUMBER: 1116-249X

where

$$\bar{V}_\beta = (\underline{V}_\beta^{-1} + \text{h} \sum_{i=1}^{N} \bar{X}_i' \ \bar{X}_\iota)^{-1} \quad (26)$$

$$\hat{\beta} = \bar{V}(\underline{V}_\beta^{-1} + \text{h} \sum_{i=1}^{N} \bar{X}_i' \, [y_i - \alpha_i l_T]) \quad (27)$$

$$\bar{v} = \text{TN} + \underline{v} \quad (28)$$

The scenarios of the three error variance structures are sample sizes, number of replications and the MCMC algorithm. The sample size which refers to the number of observations or data points available at different levels of the hierarchy. In this framework, data is structured in multiple layers (e.g., individual-level data nested within groups), and each level contributes to the estimation process. A small sample size at any level can influence the uncertainty of parameter estimates, but the hierarchical model can borrow strength from other levels or groups, sharing information across the hierarchy. This allows the model to make better estimates, even with limited data at one level, by leveraging commonalities across the dataset.

The quadratic variance function is given as:

$$\underline{V}(\varepsilon) = \mu + \theta_1 X_1 + \theta_2 X_2 + \theta_3 X_3 + \gamma_1 X_1^2 + \gamma_2 X_2^2 + \gamma_3 X_3^2 + \lambda_1 X_1 X_2 + \lambda_2 X_1 X_3 + \lambda_3 X_2 X_3. \quad (29)$$

where the initial values that are set for the parameters above are given below:

θ_1= 0.5, θ_2= 0.3, θ_3= 0.4, γ_1= 0.2, γ_2 = 0.1, γ_3= 0.15 λ_1 = 0.10, λ_2= 0.05, λ_3 = 0.08.

The model parameters was estimated using Markov Chain Monte Carlo (MCMC) methods with gibbs sampler because of its helpful Bayesian modelling platform The model performance is assessed using predictive potential, standard deviation which is a function of the variance and the relative efficiency and also cross-validation was used.

ISSN NUMBER: 1116-249X

Empirical Illustrations

The data used for this study were simulated from the random normal distribution with mean zero and variance one $X_{ij} \sim N[0,1]$, i=1, 2,…,N and j =1,2,…,N, using Gibbs Sampler of Monte Carlo simulation technique. The initial values were assumed for the regression coefficients, such that β_0=1,β_1=2, $\beta_2 =$ -1.5, β_3= 0.5 for the scenarios considered; and the simulation was conducted for three (3) different sample sizes n=250, n=500, and n=1000 and two replications r=100 and r=250.

3. Results and Discussions

The full summary of the simulated results of a known form of heteroscedasticity with different sample sizes and replications are presented below

ISSN NUMBER: 1116-249X

Table 1: Simulation Results for Quadratic Heteroscedasticity Functional Form with Sample Size [$n = 250$] and Replication [$r = 100$]

Parameters	Initial values	Estimates	Standard Deviation	Relative Efficiency	
				GLS versus OLS	Hierarchical Bayesian versus GLS
			OLS		
β_0	1.0000	2.8565	1.0600	-	-
β_1	2.0000	3.8084	1.0623	-	-
β_2	-1.5000	0.1138	1.0632	-	-
β_3	0.5000	0.8476	1.0598	-	-
			GLS		
β_0	1.0000	1.0701	0.6405	0.3658	-
β_1	2.0000	2.0341	0.6346	0.3577	-
β_2	-1.5000	-1.4697	0 6357	0.3582	-
β_3	0.5000	0.5353	0.6336	0.3581	-
			HIERARCHICAL BAYESIAN		
β_0	1.0000	1.0717	0.0949	-	0.0217
β_1	2.0000	2.0290	0.0300	-	0.0023
β_2	-1.5000	-1.4761	0.0300	-	0.0023
β_3	0.5000	0.0271	0.0300	-	0.0023

ISSN NUMBER: 1116-249X

Table 2: Simulation Results for Quadratic Heteroscedasticity Functional Form with Sample Size [$n = 250$] and Replication [$r = 250$]

Parameters	Initial values	Estimates	Standard Deviation	Relative Efficiency: GLS versus OLS	Relative Efficiency: Hierarchical Bayesian versus GLS
OLS					
β_0	1.0000	2.8492	1.0587	-	-
β_1	2.0000	3.8071	1.0608	-	-
β_2	-1.5000	0.1118	1.0604	-	-
β_3	0.5000	0.8481	1.0586	-	-
GLS					
β_0	1.0000	1.0711	0.6407	0.3675	-
β_1	2.0000	2.0349	0.6352	0.3599	-
β_2	-1.5000	-1.4680	0.6351	0.3600	-
β_3	0.5000	0.5361	0.6342	0.3602	-
HIERARCHICAL BAYESIAN					
β_0	1.0000	1.0720	0.0943	-	0.0216
β_1	2.0000	2.0282	0.0300	-	0.0023
β_2	-1.5000	-1.4745	0.0300	-	0.0023
β_3	0.5000	0.5285	0.0300	-	0.0023

ISSN NUMBER: 1116-249X

Table 3: Simulation Results for Quadratic Heteroscedasticity Functional Form with Sample Size [$n = 500$] and Replication [$r = 100$]

Parameters	Initial values	Estimates	Standard Deviation	Relative Efficiency	
				GLS versus OLS	Hierarchical Bayesian versus GLS
OLS					
β_0	1.0000	1.8370	1.8570	-	-
β_1	2.0000	3.8152	1.8663	-	-
β_2	-1.5000	0.1031	1.8637	-	-
β_3	0.5000	0.8623	1.8621	-	-
GLS					
β_0	1.0000	1.0680	0.6903	0.3732	-
β_1	2.0000	2.0307	0.6814	0.3665	-
β_2	-1.5000	-1.4696	0.6792	0.3661	-
β_3	0.5000	0.8312	0.6792	0.3662	-
HIERARCHICAL BAYESIAN					
β_0	1.0000	1.0683	0.0076	-	0.0110
β_1	2.0000	2.0252	0.0008	-	0.0012
β_2	-1.5000	-1.4741	0.0008	-	0.0012
β_3	0.5000	0.5241	0.0008	-	0.0012

ISSN NUMBER: 1116-249X

Table 4: Simulation Results for Quadratic Heteroscedasticity Functional Form with Sample Size [$n = 500$] and Replication [$r = 250$]

Parameters	Initial values	Estimates	Standard Deviation	Relative Efficiency	
				GLS versus OLS	Hierarchical Bayesian versus GLS
			OLS		
β_0	1.0000	2.8464	1.8703	-	-
β_1	2.0000	3.8125	1.8746	-	-
β_2	-1.5000	0.1092	1.8791	-	-
β_3	0.5000	0.8524	1.8765	-	-
			GLS		
β_0	1.0000	1.0650	0.6891	0.3698	-
β_1	2.0000	2.0304	0.6775	0.3629	-
β_2	-1.5000	-1.4699	0.6789	0.3626	-
β_3	0.5000	0.5292	0.6778	0.3626	-
			HIERARCHICAL BAYESIAN		
β_0	1.0000	1.0651	0.0075	-	0.0108
β_1	2.0000	2.0228	0.0008	-	0.0012
β_2	-1.5000	-1.4750	0.0008	-	0.0012
β_3	0.5000	0.5245	0.0008	-	0.0012

ISSN NUMBER: 1116-249X

Table 5: Simulation Results for Quadratic Heteroscedasticity Functional Form with Sample Size [$n = 1000$] and Replication [$r = 100$]

Parameters	Initial values	Estimates	Standard Deviation	Relative Efficiency: GLS versus OLS	Relative Efficiency: Hierarchical Bayesian versus GLS
			OLS		
β_0	1.0000	2.8469	2.7936	-	-
β_1	2.0000	3.8086	2.7951	-	-
β_2	-1.5000	0.1037	2.7979	-	-
β_3	0.5000	0.8515	2.7881	-	-
			GLS		
β_0	1.0000	1.0657	1.0276	0.3689	-
β_1	2.0000	2.0225	1.0101	0.3623	-
β_2	-1.5000	-1.4718	1.0093	0.3619	-
β_3	0.5000	0.5288	1.0075	0.3625	-
			HIERARCHICAL BAYESIAN		
β_0	1.0000	1.0670	0.0056	-	0.0054
β_1	2.0000	2.0212	0.0006	-	0.0006
β_2	-1.5000	-1.4802	0.0006	-	0.0006
β_3	0.5000	0.5215	0.0006	-	0.0006

ISSN NUMBER: 1116-249X

Table 6: Simulation Results for Quadratic Heteroscedasticity Functional Form with Sample Size [$n = 1000$] and Replication [$r = 250$]

Parameters	Initial values	Estimates	Standard Deviation	Relative Efficiency	
				GLS versus OLS	Hierarchical Bayesian versus GLS
			OLS		
β_0	1.0000	2.8477	2.7935	-	-
β_1	2.0000	3.8051	2.7943	-	-
β_2	-1.5000	0.1094	2.8001	-	-
β_3	0.5000	0.8542	2.7988	-	-
			GLS		
β_0	1.0000	1.0621	1.0263	0.3684	-
β_1	2.0000	2.0239	1.0076	0.3616	-
β_2	-1.5000	-1.4730	1.0086	0.3614	-
β_3	0.5000	0.5078	1.0093	0.3617	-
			HIERARCHICAL BAYESIAN		
β_0	1.0000	1.0635	0.0056	-	0.0054
β_1	2.0000	2.0204	0.0006	-	0.0006
β_2	-1.5000	-1.4799	0.0006	-	0.0006
β_3	0.5000	0.5216	0.0006	-	0.0006

ISSN NUMBER: 1116-249X

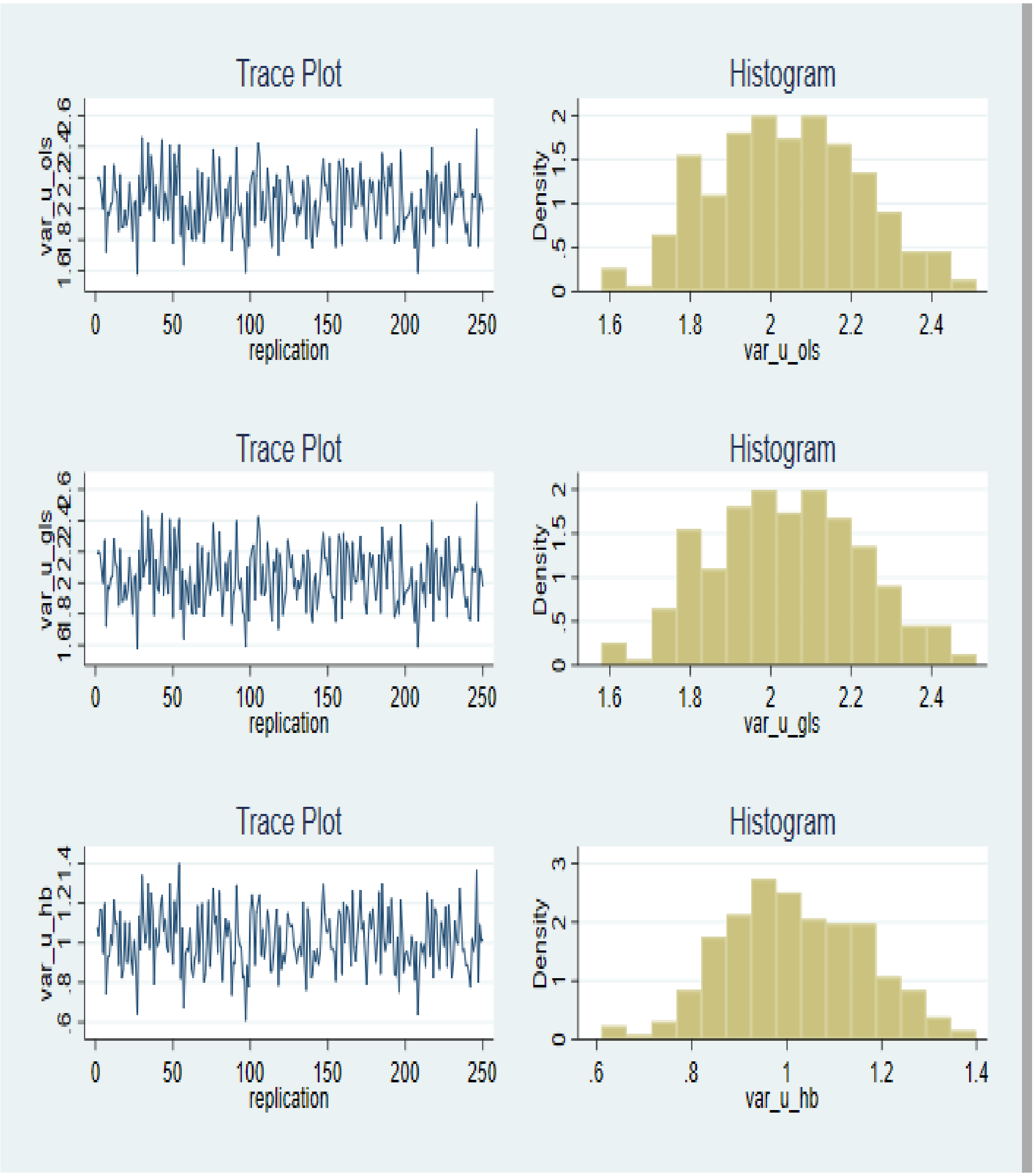

Figure 1: Simulated Dependent Variable for Quadratic Heteroscedastic Structures with replication, r = 250 and Sample Size, $n = 250$

ISSN NUMBER: 1116-249X

The tables 1, 2, 3, 4, 5, and 6 above shows the parameter, the actual value for the parameter, the standard deviation and the relative efficiency comparing the variances of a pair of contending estimators (GLS versus OLS, and HB versus GLS) under the three (3) specified estimators which are OLS, GLS, and HB.

From table 1, in considering the parameters starting from β_1 for OLS, GLS, and HB, the estimate of HB (2.0258) is observed to be closest to the initial value compared to that of GLS (2.0349) and OLS (3.8071), the HB standard deviation (0.0300) appears to be the smallest out of the three specified estimators [HB (0.0300), GLS (0.6352), and OLS (1.0608)] which indicates the stability of the estimate, then the relative efficiency of HB versus GLS (0.0023) is smaller compare to that of GLS versus OLS (0.3599) which is an indication that HB variance is the smaller to GLS variance and also GLS variance is smaller to OLS variance . Despite that the two relative efficiencies are both lesser than one, HB is said to be the most efficient estimator in correcting for heteroscedasticity compared to GLS and OLS due to its smallest variance**.**

For $\beta_{2,}$ out of the three (3) specified estimators, the estimate of HB (-1.4745) is the closest to the initial value, its standard deviation (0.0300) is the smallest and the relative efficiency for HB to GLS (0.0023) is smaller compared to that of GLS to OLS (0.3600). The HB estimates, standard deviations and relative efficiencies for β_3, appears to be most predictive, (i.e the closest estimates to the actual value), most stable, and most efficient (smallest variance).

Across the parameters ($\beta_{1,}$ $\beta_{2,}$ β_3) of the specified estimators, the HB estimates gives the most predictive estimate (i.e its estimate closer to the initial value), its standard deviation gives stability to the parameters, and its variances is the smallest which makes it more efficient compared to that of GLS and OLS

ISSN NUMBER: 1116-249X

The observations of the behavior of the parameter estimates, standard deviations and relative efficiencies from tables 2, 3, 4, 5, and 6 follows the same pattern as in table 1. Additionally, from the result, it is obvious that as the sample size increases, the variance of the HB decreases which makes it more efficient in handling the problem of heteroscedasticity.

Figure 1 reveals the range of values of the error variance of each specified estimators on a trace plot and histogram with replication r=250 and sample size n=250 It is observed that the range of values of the HB (0.4 – 1.4) is smaller than that of the GLS and OLS (1.6 - 2.6) which is an indication that the HB outperforms the other estimators in terms of its efficiency in correcting for heteroscedasticity.

4. Conclusion

The study demonstrates the efficiency of the hierarchical Bayesian method to model relationships of one or more predictor variables and a dependent variable that exhibits specified functional forms of heteroscedasticity (quadratic). The HB approach yields more accurate and precise parameter estimates compared to the traditional OLS and the GLS methods. This feat has important implications for researchers and practitioners who need to model complex relationships with varying error structures, since HB has been shown to perform best regardless of the functional form that was considered.

REFERENCES

Akinlade, Y. O., Adesina, O. A., Afolabi, N. O., & Ojo, O. O. (2021). Hierarchical Bayesian estimation of random co-efficient dynamic panel data model. *The International Journal of Science & Technology*.

Angrist, J. D., and J.-S. Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.

Carroll, R. J. (1982). Adapting for heteroscedasticity in linear models. *The Annals of Statistics, 10(*4), 1224–1233.

DiCiccio, C. J., J. P. Romano, and M. Wolf. 2019. "Improving Weighted Least Squares Inference." *Econometrics and Statistics* 10: 96−119.

ISSN NUMBER: 1116-249X

.Geweke, J. (1993). Bayesian treatment of the independent Student-t linear model. Journal of Applied Econometrics, 8(2), 169-179

Geweke, J. (2010). Bayesian econometrics. In D. J. Henderson & J. L. Glasby (Eds.), The Oxford handbook of econometrics (Vol. 2, pp. 712-733). Oxford University Press.

Koop, G. (2003). Bayesian econometrics. Wiley.

Koop, G., & Korobilis, D. (2010b). Bayesian multivariate time series methods for empirical macroeconomics. *Foundations and Trends in Econometrics, 3*(4), 267–358.

Lancaster, T. (2003). *An introduction to modern Bayesian econometrics.* Blackwell Publishing.

Leamer, E. E. 2010. "Tantalus on the Road to Asymptopia." *The Journal of Economic Perspectives* 24 (2): 31−46..

Li, Y. (2018). Bayesian linear regression. In Bayesian methods in finance (pp. 15-40). CRC Press.

Müller, P. (2013). Bayesian nonparametric data analysis. Springer.

Oseni, et al. (2019). Bayesian estimators of the normal linear regression model in the presence of functional forms of heteroscedasticity. [Journal title not provided]

Pelenis, J. (2014). Posterior consistency in Bayesian quantile regression with Heteroscedasticity *Journal of Econometrics, 178* (3), 494–507.

Poirier, D. J. (2011). *The growth of Bayesian methods in statistics and economics since1970.* Wiley Online Library.

Robinson, P. M. (1987). Asymptotically efficient estimation in the presence of heteroscedasticity of unknown form. *Econometrica, 55* (4), 875–891.

Romano, J. P., and M. Wolf. 2017. "Resurrecting Weighted Least Squares." *Journal of Econometrics* 197 (1): 1−19.

Rossi, P. E., Allenby, G. M., & McCulloch, R. (2005). Bayesian statistics and marketing. Wiley.

Rubin, D. B. (1981). The Bayesian bootstrap. *The Annals of Statistics, 9* (1), 130–134.

Salois, M. J., & Balcombe, K. G. (2014). Do food stamps cause obesity? A generalized Bayesian instrumental variable approach in the presence of heteroscedasticity. University of Reading.

Wooldridge, J. M. (2010). Econometric analysis of cross section and panel data. MIT Press.

ISSN NUMBER: 1116-249X