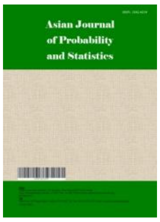

Asian Journal of Probability and Statistics

15(2): 1-17, 2021; Article no.AJPAS.74094
ISSN: 2582-0230

Robustness Test of Selected Estimators of Linear Regression with Autocorrelated Error Term: A Monte-Carlo Simulation Study

Rauf Ibrahim Rauf[1*], Okoli Juliana Ifeyinwa[1] and Haruna Umar Yahaya[1]

[1]Department of Statistics, Faculty of Science, University of Abuja, Abuja, Nigeria.

Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/AJPAS/2021/v15i230348
Editor(s):
(1) Dr. Jaya Bishwal, University of North Carolina, USA.
Reviewers:
(1) Cliff Richard Kikawa, Kabale University, Uganda.
(2) Maad M. Mijwil, Baghdad College of Economics Sciences University, Iraq.
Complete Peer review History: https://www.sdiarticle4.com/review-history/74094

Original Research Article

Received 28 July 2021
Accepted 04 October 2021
Published 12 October 2021

Abstract

Assumptions in the classical linear regression model include that of lack of autocorrelation of the error terms and the zero covariance between the explanatory variable and the error terms. This study is channeled towards the estimation of the parameters of the linear models for both time series and cross-sectional data when the above two assumptions are violated. The study used the Monte-Carlo simulation method to investigate the performance of six estimators: ordinary least square (OLS), Prais-Winsten (PW), Cochrane-Orcutt (CC), Maximum Likelihood (MLE), Restricted Maximum- Likelihood (RMLE) and the Weighted Least Square (WLS) in estimating the parameters of a single linear model in which the explanatory variable is also correlated with the autoregressive error terms. Using the models' finite properties(mean square error) to measure the estimators' performance, the results shows that OLS should be preferred when autocorrelation level is relatively mild ($\rho = 0.3$) and the PW, CC, RMLE, and MLE estimator will perform better with the presence of any level of AR (1) disturbance between 0.4 to 0.8 level, while WLS shows better performance at 0.9 level of autocorrelation and above. The study thus recommended the application of the various estimators considered to real-life data to affirm the results of this simulation study.

Keywords: Prediction; estimators; linear regression model; simulation; autocorrelation.

**Corresponding author: Email: rauf.ibrahim@uniabuja.edu.ng;*

1 Introduction

An AR(p) model is an autoregressive model where specific lagged values of e_t are used as predictor variables. Lags are where results from one time period affect following periods. The value for "p" is called the ***order*.** For example, an AR (1) would be a "first order autoregressive process". The outcome variable in a first order AR process at some point in time *t* is related only to time periods that are one period apart (i.e. the value of the variable at t – 1). A second or third order AR process would be related to data two or three periods apart [1,2,3].

1.2 The AR(p) model is defined by the equation

First we define the standard OLS estimation to be of the form:

$$y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_k X_{ik} + e_t$$

where i = 1, 2,…t.

Thus,

$$y_1 = \beta_0 + \beta_1 X_{11} + \beta_2 X_{12} + \ldots + \beta_k X_{1k} + e_1$$
$$y_2 = \beta_0 + \beta_1 X_{21} + \beta_2 X_{22} + \ldots + \beta_k X_{2k} + e_2$$
$$\vdots$$
$$y_n = \beta_0 + \beta_1 X_{t1} + \beta_2 X_{t2} + \ldots + \beta_k X_{tk} + e_t$$

In vector form

$$\underset{t \times 1}{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \cdot \\ \cdot \\ \cdot \\ y_t \end{bmatrix}} = \underset{t \times (k+1)}{\begin{bmatrix} 1 & X_{11} & X_{12} & \ldots & X_{1k} \\ 1 & X_{21} & X_{22} & \ldots & X_{2k} \\ 1 & X_{31} & X_{32} & \ldots & X_{3k} \\ & & \cdot & & \\ & & \cdot & & \\ & & \cdot & & \\ 1 & X_{t1} & X_{t2} & \ldots & X_{tk} \end{bmatrix}} \underset{(k+1) \times 1}{\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \cdot \\ \beta_k \end{bmatrix}} + \underset{t \times 1}{\begin{bmatrix} e_0 \\ e_1 \\ e_2 \\ \cdot \\ \cdot \\ \cdot \\ e_t \end{bmatrix}}$$

The general form is:

$$y = X\beta + e_t$$

where *y* is an (t × 1) vector of observations of the dependent variable, *X* matrix is an *n* × (*k*+1) full rank matrix of explanatory variables, β is a ((*k*+1) ×1 vector of unknown parameters to be estimated, e_t is (t × 1) vector of random error. The parameter β in a linear regression model are commonly estimated using the Ordinary Least Squares Estimator (OLSE). The OLSE of β is given as:

$$\hat{\beta}_{OLS} = (X'X)^{-1} X'y$$

The estimator is generally preferred if there is no violation in any of the assumptions of the linear regression model [4]. However, if the model violates the autocorrelation assumption, then the error term e_t will take the form:

$$e_t = \varphi_0 + \varphi_1 e_{t-1} + \varphi_2 e_{t-2} + \ldots + \varphi_p e_{t-p} + \delta_t$$

Where:

- e_{t-1}, e_{t-2}… e_{t-p} are the past series values (lags) of the error term e_t,

- φ_{0-p} is the coeeficient of the past series (lags) of the error term e_t,
- δ_t is white noise (i.e. randomness),
- and δ_t is defined by the following equation:

$$\delta_t = \left(1 - \sum_{t=1}^{p} \emptyset_i\right)$$

where $\emptyset_i$ is the process mean

To estimate the parameters of linear models and also to enable inferences to be made about these estimators, certain underlying assumptions are made. The absence of autocorrelation of the error terms and that X is a matrix with non-stochastic elements and has rank k < t, hence $\emptyset_i$ and e_t are independent for all *i* and *t* [5,6] is expected.

This study is channeled towards the estimation of the parameters of the linear models when the above assumption is violated using simulated data set. This would help researchers and practitioners in the choice of estimator in empirical work when the regressor and the error terms are not well behaved. It would also allow correct inferences in linear models plagued by autocorrelated disturbances, which are also significantly correlated with the exponential trended explanatory variable [7]. Some researchers have worked on the methods for detecting the presence of autocorrelation and alternative estimators to estimate the parameters in the linear regression model with autocorrelation error [8,9,4]. These include [10-21].

However, thus paper evaluates the estimation ability of six methods of estimation of AR (1) order one process. We compared the ordinary least square (OLS), Prais-Winsten (PW), Cochrane-Orcutt (CC), Maximum Likelihood (MLE), Restricted Maximun- Likelihood (RMLE) and the Weighted Least Square (W+LS).

1.3 Aim and Objectives

The aim of the study is to compare the various method for estimating model with AR (1) disturbance. Particularly the specific objectives are:

i. Determine all the criteria for comparing the estimators for model with autoregressive of order one disturbance.
ii. Compare the performance of selected AR of order one process estimator with various level of AR(i) disturbance.
iii. Determine the best AR (1) estimator in terms of efficiency (minimum standard error) which could be Bias, Mean Absolute Error and Mean Square Error.

2 Material and Methods

2.1 Structures of disturbance term

The following structures are popular in autocorrelation:

- Autoregressive (AR) process.
- Moving average (MA) process.
- Joint Autoregression moving average (ARMA) process.

This study is based on Autoregressive process only so we will limit our discussion to autoregressive process only [22,23].

2.2 The selected AR (1) estimators for comparison

In the next section the description for the different estimation methods used in this paper as presented.

2.2.1 Cochrane-orcutt procedure

This procedure utilizes *P* matrix defined while estimating β_0 and β_1 is known. It has following steps:

i. Apply OLS to $y_t = \beta_0 + \beta_1 X_t + u_t$ and obtain residual vector e.
ii. Estimate $\rho \; by \; r = \frac{\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=1}^{n} e_t^2}$ Note that r is a consistent estimator of ρ.
iii. Replace ρ by r is

$$y_t - \rho y_{t-1} = \beta_0(1 - \rho) + \beta(x_t - \rho x_{t-1}) + \varepsilon_t$$

and apply OLS to the transformed model.

$$y_t - r y_{t-1} = \beta_0^* + (x_t - r x_{t-1}) + disturbance\ term$$

And obtained estimators of β_0^* and β as $\hat{\beta}_0^*$ and $\hat{\beta}$ respectively

This is Cochrane-Orcutt procedure. Since two successive applications of OLS are involved, so it is also called as **two-step procedure.**

This application can be repeated in the procedure as follows:

i. Put $\hat{\beta}_0^*$ and $\hat{\beta}$ in original model ii. Calculate the residual sum of squares.
iii. Calculate $\rho \; by \; r = \frac{\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=2}^{n} e_t^2}$ and substitute it in the model.

$$y_t - \rho y_{t-1} = \beta_0(1 - \rho) + \beta(x_t - \rho x_{t-1}) + \varepsilon_t$$

and again obtain the transformed model.

iv. Apply OLS to this model and calculate the regression coefficients.

This procedure is repeated until convergence is achieved, i.e., iterate the process till the two successive estimates are nearly same so that stability of estimator is achieved. This is an iterative procedure and is numerically convergent procedure. Such estimates are asymptotically efficient and there is a loss of one observation [24].

2.2.2 Prais-Winsten procedure

This is also an iterative procedure based on two step transformations.

i. Calculate $\rho \; by \; \hat{\rho} = \frac{\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=3}^{n} e_t^2}$ where $e_t'^s$ are residuals based on OLSE.
ii. Replace $\rho \; by \; \hat{\rho}$ in the model as in Cochrane-Orcutt procedure

$$\left(\sqrt{1-\hat{\rho}^2}\right) y_1 = \left(\sqrt{1-\hat{\rho}^2}\right) \beta_0 + \left(\sqrt{1-\hat{\rho}^2}\right) x_t + \left(\sqrt{1-\hat{\rho}^2}\right) u_t$$

$$y_t - \hat{\rho} y_{t-1} = (1-\hat{\rho})\beta_0 + \beta(x_t - \hat{\rho} x_{t-1}) + (u_t - \hat{\rho} u_{t-1}), t = 1,2,\cdots,n$$

iii. Use OLS for estimating the parameters.

The estimators obtained with this procedure are asymptotically as efficient as best linear unbiased estimators. There is no loss of any observation.

2.2.3 Maximum likelihood procedure

Assume that $y \sim N\ (X\beta,\ \sigma_\varepsilon{}^2\varphi)$. the likelihood function for $\beta,\ \rho\ and\ \sigma_\varepsilon{}^2$ is

$$L = \frac{1}{(2\pi)^{n/2}|\varphi|^{n/2}} exp\left[-\frac{1}{2\sigma_\varepsilon^2}(y-X\beta)'\varphi^{-1}(y-X\beta)\right]$$

Ignoring the constant and using

$$|\varphi| = \frac{1}{1-\rho^2}$$

The log-likelihood is

$$lnL = lnL(\beta, \rho\ and\ \sigma_\varepsilon^2) = \frac{-n}{2} ln\sigma_\varepsilon^2 + \frac{1}{2}\ln(1-\rho^2) - \frac{1}{2\sigma_\varepsilon^2}(y-X\beta)'\varphi^{-1}(y-X\beta)$$

The maximum likelihood estimators of $\beta, \rho\ and\ \sigma_\varepsilon^2$ can be obtained by solving the normal equations

$$\frac{\partial lnL}{\partial \beta} = 0, \qquad \frac{\partial lnL}{\partial \rho} = 0, \qquad \frac{\partial lnL}{\partial \sigma_\varepsilon^2} = 0$$

Here normal equations turn out to be nonlinear in parameters and cannot be easily solved.
One solution is to

- first derive the maximum likelihood estimator of σ_ε^2
- Substitute it back into the likelihood function and obtain the likelihood function as the function of $\beta\ and\ \rho$
- Maximize this likelihood function with respect to $\beta\ and\ \rho$

Thus,

$$\frac{\partial lnL}{\partial \sigma_\varepsilon^2} = 0 \Rightarrow \frac{-n}{2\sigma_\varepsilon^2} + \frac{1}{2\sigma_\varepsilon^2}(y-X\beta)'^{\varphi^{-1}}(y-X\beta) = 0$$

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{n}(y-X\beta)'^{\varphi^{-1}}(y-X\beta)$$

Is the estimator of $\sigma_\varepsilon{}^2$.

Substituting $\hat{\sigma}_\varepsilon^2$ in place of σ_ε^2 is the log-likelihood function yields

$$lnL^* = lnL^*(\beta,\rho) = \frac{-n}{2} ln\left[-\frac{1}{n}(y-X\beta)'\varphi^{-1}(y-X\beta)\right] + \frac{1}{2}\ln(1-\rho^2) - \frac{n}{2}$$

$$= \frac{-n}{2}\left[ln\{(y-X\beta)'\varphi^{-1}(y-X\beta)\} - \frac{1}{2}\ln(1-\rho^2)\right] + k$$

$$k - \frac{n}{2} ln\left[\frac{(y-X\beta)'\varphi^{-1}(y-X\beta)}{(1-\rho^2)^{1/n}}\right]$$

$$where\ k = \frac{n}{2}\ln n - \frac{n}{2}$$

Maximization of ln L^* is equivalent to minimizing the function

$$\left[\frac{(y - X\beta)'\varphi^{-1}(y - X\beta)}{(1-\rho^2)^{1/n}}\right]$$

Using optimization techniques of non-linear regression, this function can be minimized and estimates of β and ρ can be obtained.

If n is large and $|\rho|$ is not too close to one, then the term $(1 - \rho^2)^{1/n}$ is negligible and the estimates of β will be same as obtained by nonlinear least squares estimation.

2.2.4 Restricted maximum likelihood method

This method is sometimes called the reduced maximum likelihood or residual maximum likelihood method. The likelihood of a sample is the prior probability of obtaining the data in your sample. Assume that

$$Y = (Y_1^T, Y_2^T, \cdots, Y_n^T)^T$$

Follows the k dimensional AR (p) process, given by

$$Y_t = \mu + \tilde{Y}_t$$

$$\tilde{Y}_t = \sum_{i=1}^{p} A_i \tilde{Y}_{t-1} + e_t$$

Where e_t is an independent (0,1) series and the roots of

$$\left| I_k z^p - \sum_{i=1}^{p} A_i z^{p-1} \right| = 0$$

Are at most one in absolute value. The initial value of $\tilde{Y}_t$ are assumed to $\tilde{Y}_t = 0 \; for \; t =$ 0, −1, ⋯, $-p + 1$. Letting l_k = Var(Y), the log restricted likelihood for Y is

$$RL = \frac{-1}{2} log|\Sigma| - \frac{1}{2} log|X^T \Sigma X| - \frac{1}{2} Q$$

Where

$$Q = YT_k - 1Y - YT_k - 1(XT_k - 1X) - 1XT_k - 1Y$$

And X is the nk x k matrix (I_k, ⋯, I_k)

For simplicity, we illustrate the weighted least squares approximate restricted maximum likelihood estimator through the univariate AR (1) model, where

$$Y_t = \alpha \tilde{Y}_t + e_t$$

The restricted maximum likelihood estimator $\hat{\alpha}_{REML}$ is the minimizer of

$$RL\ (\alpha, \sigma_\varepsilon^2) = -\frac{(n-1)}{2} log\sigma_\varepsilon^2 + logw(\alpha) - \frac{1}{2} Q$$

Where

$(\alpha) = \{1 + (n - 1)(1 - \alpha)^2\}^{-1}$

2.2.5 Ordinary least squares

The ordinary least squares (OLS) for an AR (1) model is:

$$\widehat{\emptyset}_{ols} = \frac{\sum_{t-1}^{T-1}(y_t - \bar{y})(y_{t+1} - \bar{y})}{\sum_{t-1}^{T-1}(y_t - \bar{y})^2}$$

The asymptotic standard error for $\widehat{\emptyset}_{\text{ols}}$ is:

$$SE_{ols} = \sqrt{\frac{T-(T-1)\emptyset^2-1}{T-Ty^2{}_T}}$$

The OLS estimation is capable of handling non-stationary data under certain restrictions. This means that it is possible to obtain a non-stationary estimate (i.e., $|\widehat{\emptyset}_{\text{ols}}|>1|\phi\text{^ols}|>1$). To identify possible different behaviors, we distinguish two types of OLS analysis results: OLS-A will refer to the complete results, where OLS-S will refer to the results where the non-stationary results are left out [22,23].

Autoregressive (AR) Process

The structure of disturbance term in autoregressive process (AR) is assumed as

$$u_t = \emptyset_1 u_{t-1} + \emptyset_2 u_{t-2} + \cdots + \emptyset_p u_{t-p} + \varepsilon_t$$

That is, the current disturbance term depends on the q lagged disturbances and $\emptyset_1, \emptyset_2, \cdots, \emptyset_p$ are the parameters (coefficients) associated with $u_{t-1}, u_{t-2}, \cdots, u_{t-p}$ respectively. An additional disturbance term is introduced in ε_t which is assumed to satisfy the following conditions:

$(\varepsilon_t) = 0$

$(\varepsilon t{-}s) = \{\sigma 0_\varepsilon{}^2\ if\ if\ s\ s \neq= 00$

This process is termed as AR (p) process. In practice, the AR (1) process is more popular.

2.3.1 Estimation under the first order autoregressive process

Consider a simple linear regression model

$$y_t = \beta_0 + \beta_1 X_t + \varepsilon_t \; t = 1,2,\cdots,n$$

Assume the $u_s{}'s$ follow a first order autoregressive scheme defined as

$$u_t = \rho u_{t-1} + \varepsilon_t$$

Where

$$|\rho| < 1, E(\varepsilon_t) = 0$$

$$E(\varepsilon_t \varepsilon_{t-s}) = \begin{cases} \sigma_\varepsilon^2 \; if \; s = 0 \\ 0 \; if \; s \neq 0 \end{cases}$$

for all t =1, 2..., n where ρ is the first order autocorrelation between u_t and u_{t-1} t =1, 2..., n. Now

$$u_t = \rho u_{t-1} + \varepsilon_t$$
$$= \rho(u_{t-2} + \varepsilon_t) + \varepsilon_t$$
$$=\vdots$$
$$= \varepsilon_t + \rho\varepsilon_{t-1} + \rho^2\varepsilon_{t-2} + \cdots$$
$$\sum_{r=0}^{\infty} \rho^r \varepsilon_{t-r}$$
$$E(u_t) = 0$$
$$E(u_t^2) = E(\varepsilon_t^2) + \rho^2 E(\varepsilon_{t-1}^2) + \rho^4 E(\varepsilon_{t-1}^2) + \cdots$$
$$= (1 + \rho^2 + \rho^4 + \cdots)\sigma_\varepsilon^2$$
$$E(u_t^2) = \sigma_u^2 = \frac{\sigma_\varepsilon^2}{1-\rho^2} for\ all\ t$$
$$E(u_t u_{t-1}) = E[(\varepsilon_t + \rho\varepsilon_{t-1} + \rho^2\varepsilon_{t-2} + \cdots) * (\varepsilon_{t-1} + \rho\varepsilon_{t-2} + \rho^2\varepsilon_{t-3} + \cdots)]$$
$$= E[\{\varepsilon_t + \rho(\varepsilon_t + \rho\varepsilon_{t-1} + \cdots)\} * \{\varepsilon_t + \rho(\varepsilon_{t-2} + \rho\varepsilon_{t-3} + \cdots)\}]$$
$$= \rho E[(\varepsilon_t + \rho\varepsilon_{t-1} + \cdots)^2] = \rho\sigma_u^2$$

Similarly, $E(u_t u_{t-2}) = \rho^2\sigma_u^2$.

In general, $E(u_t u_{t-s}) = \rho^s\sigma_u^2$.

3 Procedure for the Monte Carlo Simulation Experiment

To compare the various estimators for the autocorrelation (ϕ), we simulate according to an AR(1) model. In the generation of the data we vary the length of the time series T and the autocorrelation ϕ. For T we use nineS different sizes, namely 10,20,30,40, 50,100,250,500 and 1000 For ϕ, we use an autocorrelation of 0.0, ±0.1 ±0.3 ±0.4 ±0.5 ±0.8 ±0.9 and ±0.99. Earlier studies show that there is a difference between the bias for the negative and positive ϕ for several estimators, including r 1 and the C-statistic [12], (Solanas et al. 2010). This indicates that a thorough test is required to include both positive and negative autocorrelations. Finally, the number of replications must be set. However, a pilot study showed that the maximum standard deviation of the mean $\widehat{\emptyset}$ over 2000–5000 replications was 0.0007, when T = 10 and ρ=0.7, for all estimators. Therefore, we use N = 2000 replications per condition. Considering a fully crossed experimental design, this yields 13× 9 × 2000 = 234,000 simulated data sets.

Across all conditions, μ is set to zero and σ2 to one, which can be done without loss of generality. This results in a standard normal distribution for yt given ϕ.

3.1 Analyses procedure

For the simulations and analyses we use the program ‘R’ (R Core Team 2019). The Estimator was computed directly with the basic functions available. For the Yule–Walker, OLS and MLE methods we use the command ‘ar’ from the software package ‘stats’.

4 Results and Discussion

This section gives the result of the analysis based on the proposed methods. The ordinary least square (OLS), Prais-Winsten (PW), Cochrane-Orcutt (CC), Maximum Likelihood (MLE), Restricted Maximun- Likelihood (RMLE) and the Weighted Least Square (WLS). The analysis is based on Monte Carlo simulation study and application to real life with details described in the previous chapter. The estimators deviations from the true value is used to with the predictive ability of the estimators to compare their performance in the presence of AR(1) disturbance. The performance of these estimators towards the autocorrelated structure or AR(1) disturbance was justified by considering their error measurement, Bias, mean square error (MSE) and mean absolute error (MAE).

4.1 Simulation study results

4.1.1 Comparison Based on BIAS, MSE and MAE of the Parameters

Table 1. Comparison among estimators when ρ = 0.0 (AR-disturbance present at 0.0)

Mean Square Error (MSE)						
SAMPLE	**OLS**	**PW**	**CC**	**MLE**	**RMLE**	**WLS**
10	1.677556439	1.829163065	2.539859486	1.802230649	1.585652855	1.415168571
20	0.494560939	0.566457888	0.588459044	0.569505163	0.555301771	3.056439212
30	1.742255165	1.830558218	1.838034381	1.828889002	1.825156999	8.183528328
40	11.41665309	12.02446339	12.40201021	12.03199849	11.97156466	8158.092501
50	5.473026322	5.828757816	5.964987538	5.828384995	5.816128397	297.5615122
100	3.376579875	3.414000651	3.460433867	3.41400669	3.410434891	30.89522509
250	0.518051205	0.519907542	0.52308319	0.519903164	0.519994043	0.69079669
500	0.135289529	0.135434841	0.13612885	0.135437562	0.135450361	0.142934951
1000	0.03359212	0.033615699	0.033680409	0.033615829	0.033617169	0.041381115

Table 2. Comparison among estimators when ρ = 0.10 (AR-disturbance present at 0.10)

Mean Square Error (MSE)						
SAMPLE	**OLS**	**PW**	**CC**	**MLE**	**RMLE**	**WLS**
10	0.462315739	0.478483458	0.418142379	0.4790216	0.47757641	0.463951493
20	3.159950473	3.421204811	3.593516828	3.422884958	3.290534938	50.45198821
30	0.741496085	0.779386738	0.808085996	0.776178486	0.769438922	4.235307818
40	11.08924831	11.43755971	11.93610356	11.45269536	11.39616205	137.4550253
50	15.14334068	15.68743787	16.19074988	15.68821432	15.64868408	193.753046
100	3.644542377	3.646340615	3.68792863	3.646604484	3.641554484	8.409397655
250	0.567372485	0.561628161	0.564595037	0.561632496	0.561562572	0.623883457
500	0.143229511	0.142042615	0.142305562	0.142046053	0.142057439	0.157277545
1000	0.036287272	0.035904117	0.035971877	0.035903954	0.035904298	0.042978231

Table 3. Comparison among estimators when ρ = -0.10 (AR-disturbance present at -0.10)

Mean Square Error (MSE)						
SAMPLE	**OLS**	**PW**	**CC**	**MLE**	**RMLE**	**WLS**
10	0.258389418	0.349034633	0.349502268	0.343482842	0.34832576	0.2633241
20	2.913600132	3.230113438	3.467847342	3.244529307	3.18393118	7.66851369
30	0.726430661	0.745376528	0.770134343	0.741083986	0.740568764	1.845775022
40	10.28127075	10.72849003	11.17650426	10.74036628	10.71991632	5148.136844
50	14.06591229	14.42331129	14.86645718	14.41729431	14.4150196	230.9845854
100	3.339214275	3.375751677	3.410906683	3.376785626	3.372670598	19.8952734
250	0.519381878	0.514054377	0.516581114	0.514007485	0.513980642	0.626532794
500	0.131038114	0.128490883	0.12871002	0.128489304	0.128505273	0.149336984
1000	0.032873913	0.03231277	0.032365042	0.03231307	0.03231432	0.040702466

Table 4. Comparison among estimators when ρ = 0.30 (AR-disturbance present at 0.30)

	Mean Square Error (MSE)					
SAMPLE	**OLS**	**PW**	**CC**	**MLE**	**RMLE**	**WLS**
10	2.041778792	2.225949917	3.043870284	2.212873968	2.058200467	1.938352593
20	7.982790454	8.044362429	9.328612318	8.061568522	7.962559912	32.35866816
30	4.335231373	3.947134196	4.224114621	3.934717875	3.921577835	69.89287969
40	0.106079042	0.106359855	0.097743996	0.106202037	0.105466352	0.132088781
50	0.200140947	0.166508761	0.186175401	0.166550684	0.165437042	0.688089708
100	4.587826784	4.122500096	4.164997454	4.122882799	4.12032928	12637.54277
250	0.727810692	0.64746263	0.648769648	0.647479199	0.647388094	0.798788069
500	0.172274123	0.153310026	0.153842293	0.153309423	0.153306735	0.179290748
1000	0.045442006	0.040883036	0.040899504	0.040882731	0.040882488	0.050783469

Table 5. Comparison among estimators when ρ = -0.30 (AR-disturbance present at -0.30)

	Mean Square Error (MSE)					
SAMPLE	**OLS**	**PW**	**CC**	**MLE**	**RMLE**	**WLS**
10	1.602507546	1.26102012	1.259510017	1.294280269	1.33084165	1.577842422
20	6.022518184	6.632577267	7.042412438	6.506753723	6.336043496	15.25446723
30	3.785250278	3.334929585	3.467960367	3.330842152	3.359660619	100.4551538
40	0.102985419	0.079158488	0.075988067	0.079335558	0.080559638	0.257269497
50	0.143589292	0.140327761	0.145705389	0.140363407	0.139113222	0.345503277
100	3.378911404	2.988342875	3.011085204	2.987467299	2.986572381	8.04378765
250	0.530654217	0.467848684	0.467733586	0.467837895	0.467743906	0.642459775
500	0.126793257	0.108964936	0.108839174	0.108964831	0.108966845	0.151172183
1000	0.032815129	0.028132671	0.028145681	0.028133403	0.028133908	0.044364189

Table 6. Comparison among estimators when ρ = 0.50 (AR-disturbance present at 0.50)

	Mean Square Error (MSE)					
SAMPLE	**OLS**	**PW**	**CC**	**MLE**	**RMLE**	**WLS**
10	6.088655792	4.910332683	9.346574679	4.500885901	5.315081067	5.137357085
20	0.362032346	0.239085798	0.376184652	0.251233321	0.255886256	0.9362035
30	3.685031778	2.984971919	3.210477343	2.992793468	2.979198808	9.395262384
40	40.9819629	33.44534932	35.26855533	33.52223614	33.70344067	267801.5212
50	8.463105482	6.760905155	6.94539354	6.914344499	7.541400193	143.4221623
100	6.667181788	5.414045619	5.552208509	5.416114151	5.415977911	28.48404332
250	1.059214625	0.857567933	0.869805479	0.85769179	0.857710762	0.948768603
500	0.264300873	0.210390137	0.21118345	0.210385262	0.210384229	0.237670706
1000	0.064446736	0.051951842	0.052065036	0.051952697	0.051953457	0.066922614

Table 7. Comparison among estimators when ρ = 0.80 (AR-disturbance present at 0.80)

	Mean Square Error (MSE)					
SAMPLE	**OLS**	**PW**	**CC**	**MLE**	**RMLE**	**WLS**
10	0.642137301	0.634332376	0.628571464	0.658109753	0.723236123	1.080884084
20	1.088308534	0.61242658	2.10932706	0.333037518	0.290868885	0.335017258
30	3.551666925	3.559135643	3.773053495	3.440295456	3.951679721	2.844658131
40	1.661510764	1.336792916	1.255191691	1.328076238	1.363428958	1.201322753
50	51.96614494	39.60880832	48.80707849	39.93475882	41.35312601	237.0424919
100	26.73562329	21.16686959	22.60145347	21.33198801	21.48582487	40.16296644
250	4.348565682	3.500147067	3.620431435	3.727741351	6.561116295	2.119983957
500	1.084455459	0.859619053	0.878303252	0.85969345	0.85978076	0.54900904
1000	0.280510916	0.224825352	0.227028661	0.224838722	0.224836591	0.163092895

Table 8. Comparison among estimators when ρ = 0.90 (AR-disturbance present at 0.90)

	Mean Square Error (MSE)					
SAMPLE	**OLS**	**PW**	**CC**	**MLE**	**RMLE**	**WLS**
10	1.684155932	0.88496137	10.72849003	11.17650426	10.74036628	10.71991632
20	7.534035456	7.218775537	7.301436667	7.283278202	7.150336223	2.935884825
30	1.312845166	3.034413954	5.067729304	4.914366001	4.061003117	0.105569111
40	40.3975682	32.1648494	57.4594885	32.60808561	34.63794658	43.61441184
50	5.97209945	4.455509263	7.265226057	4.468124848	4.407906998	8.881103702
100	66.42616422	54.66694646	67.92106598	53.3321713	55.3841561	46.92584652
250	14.93244	12.7879846	13.96771344	12.79672588	12.80904169	4.112949593
500	3.794334793	3.261157225	3.445274371	3.261189426	3.261289974	0.985940996
1000	0.925407474	0.812914399	0.824430775	0.812994713	0.813046033	0.274086904

Table 9. Comparison among estimators when ρ = 0.99 (AR-disturbance present at 0.99)

	Mean Square Error (MSE)					
SAMPLE	**OLS**	**PW**	**CC**	**MLE**	**RMLE**	**WLS**
10	39.0236102	44.57706083	37.47012345	36.33916451	35.81777338	8.125857224
20	12.27700069	8.928107439	8.018848488	8.561764836	8.475860607	3.938409337
30	77.21111884	75.0462983	99.78911919	72.02676346	72.97828565	19.26653639
40	48.35996234	588.8090677	54.44198461	33.06072197	29.48920424	15.30803365
50	114.7620092	107.4621138	103.0759894	94.19029371	93.38343874	38.91425177
100	37.91026034	36.47390218	30.04655628	35.1406444	35.77366318	2.476228869
250	6.451623796	4.796415844	5.067729304	4.914366001	4.061003117	0.105569111
500	1.838888221	0.810574411	0.654540666	0.857791122	2.55448943	0.046821323
1000	76.04714833	70.97343513	85.41929284	70.41062272	201.231671	2.137437148

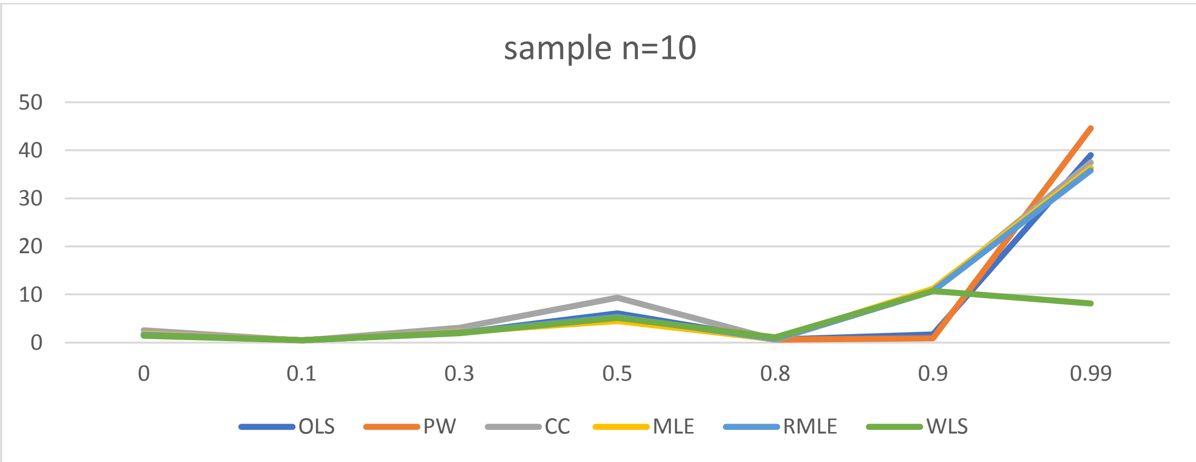

Fig. 1. Plots of the MSE on the various estimators with sample sizes of 10 across different AR disturbance (0,0.10, 0.30,0.50,0.80,0.90,0.99)

After a critical observations on the data and the line graph of the MSE of the respective estimators at the introduction of AR(1) of different levels i.e ρ =0, 0.1,0.3,0.5,0.8,0.9, 0.99 levels in the simulation program, the following are deduction about the Mean Square Errors (MSE) of the Estimators for sample sizes 10 (n=10). It was observed that PW has the largest value of MSE and WLS has the lowest when AR (1) is at level 0.99. We then conclude that estimation by the WLS is the best in the presence of AR(1) of level 0.99 when the sample size is 10.

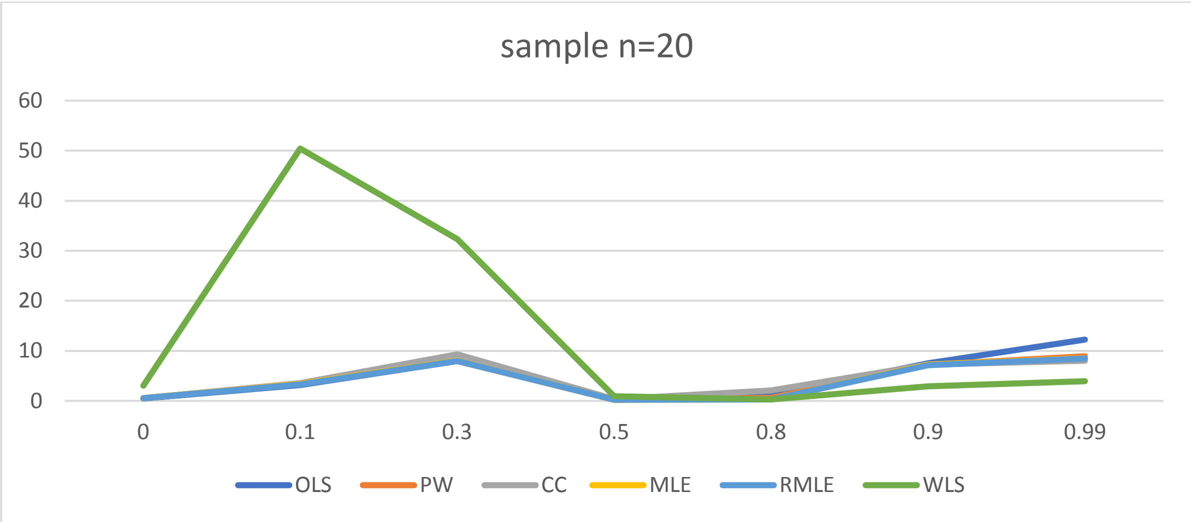

Fig. 2. Plots of the MSE on the various estimators with sample sizes of 20 across different AR disturbance (0,0.10, 0.30,0.50,0.80,0.90,0.99)

After a critical observations on the data and the line graph of the MSE of the respective estimators at the introduction of AR(1) of different levels i.e ρ =0, 0.1,0.3,0.5,0.8,0.9, 0.99 levels in the simulation program, the following are deduction about the Mean Square Errors (MSE) of the Estimators for sample sizes 20 (n=20). It was observed that OLS has the largest value of MSE and WLS has the lowest when AR (1) is at level 0.99. We then conclude that estimation by the WLS is the best in the presence of AR(1) of level 0.99 when the sample size is 20. It is also observed that when AR(1) is at 0 to 0.5, WLS was mostly affected among other estimators.

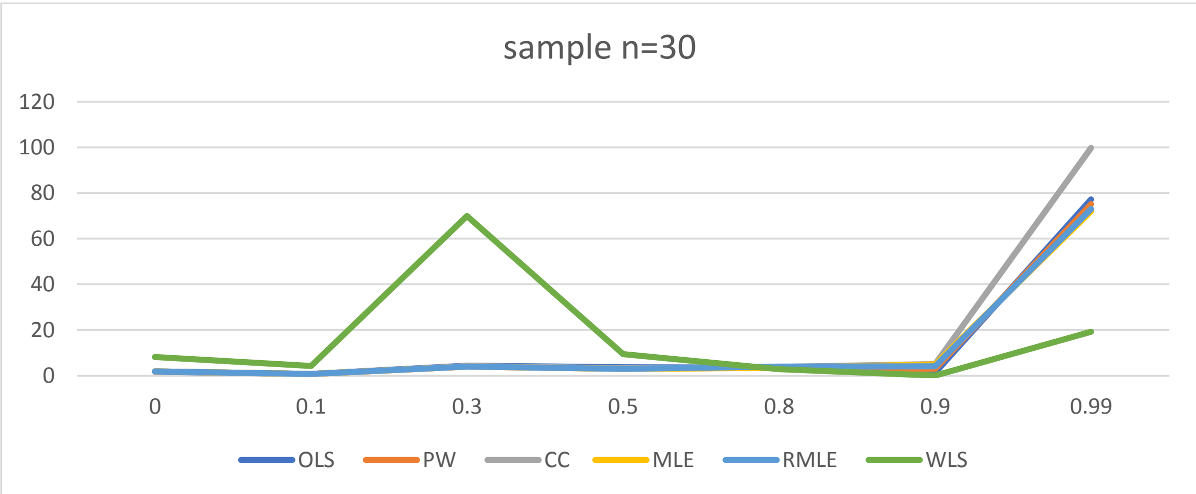

Fig. 3. Plots of the MSE on the various estimators with sample sizes of 30 across different AR disturbance (0,0.10, 0.30,0.50,0.80,0.90,0.99)

After a critical observations on the data and the line graph of the MSE of the respective estimators at the introduction of AR(1) of different levels i.e ρ =0, 0.1,0.3,0.5,0.8,0.9, 0.99 levels in the simulation program, the following are deduction about the Mean Square Errors (MSE) of the Estimators for sample sizes 30 (n=30). It was observed that CC has the largest value of MSE and WLS has the lowest when AR (1) is at level 0.99. We then conclude that estimation by the WLS is the best in the presence of AR(1) of level 0.99 when the sample size is 30. It is also observed that when AR(1) is at 0 to 0.5, WLS was mostly affected among other estimators.

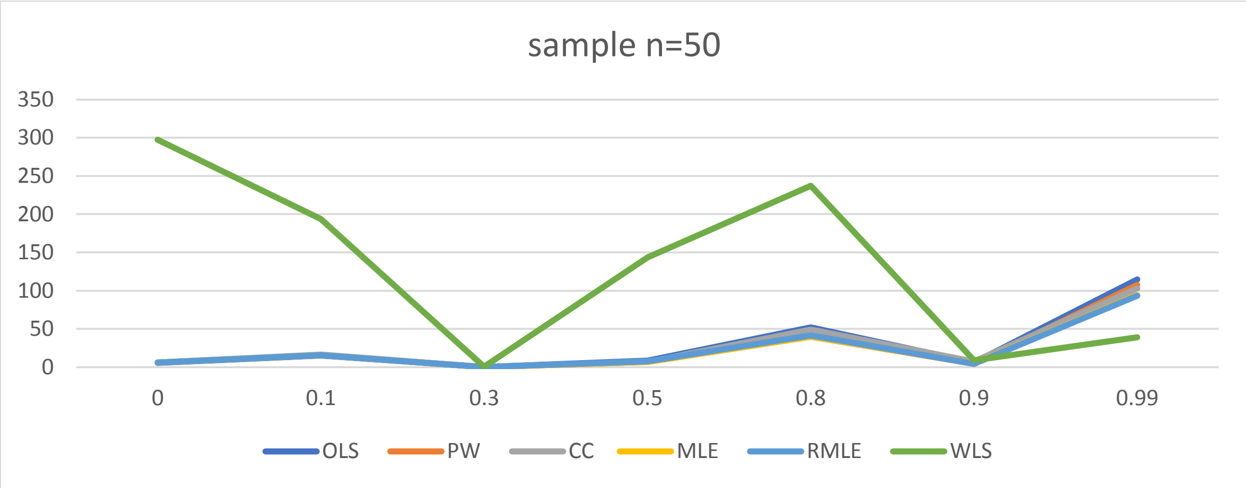

Fig. 4. Plots of the MSE on the various estimators with sample sizes of 50 across different AR disturbance (0,0.10, 0.30,0.50,0.80,0.90,0.99)

After a critical observations on the data and the line graph of the MSE of the respective estimators at the introduction of AR(1) of different levels i.e ρ =0, 0.1,0.3,0.5,0.8,0.9, 0.99 levels in the simulation program, the following are deduction about the Mean Square Errors (MSE) of the Estimators for sample sizes 50 (n=50). It was observed that OLS has the largest value of MSE and WLS has the lowest when AR (1) is at level 0.99. We then conclude that estimation by the WLS is the best in the presence of AR(1) of level 0.99 when the sample size is 50. It is also observed that when AR(1) is at 0 to 0.9, WLS was mostly affected (highest value) among other estimators.

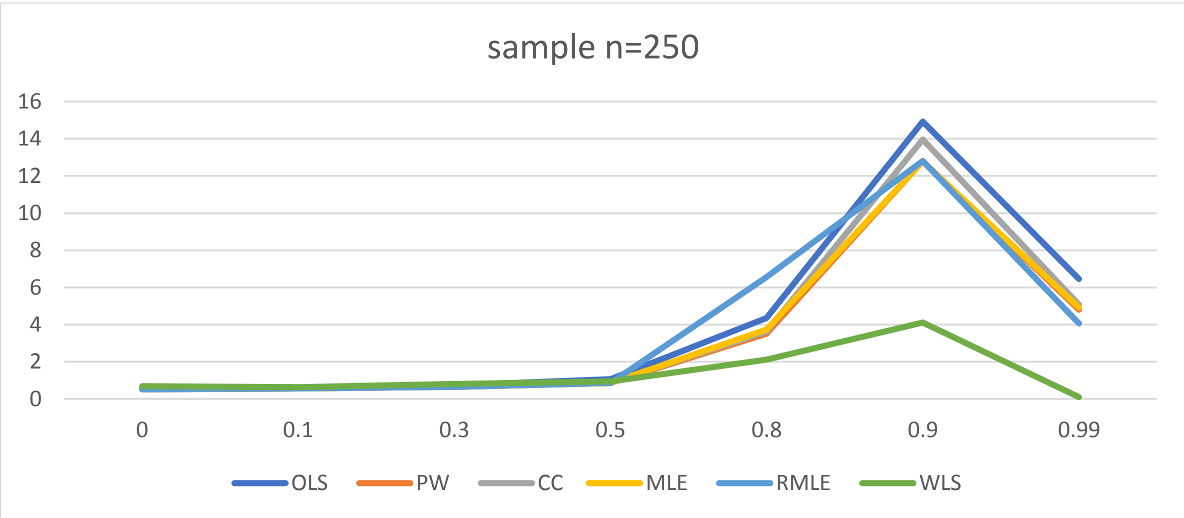

Fig. 5. Plots of the MSE on the various estimators with sample sizes of 250 across different AR disturbance (0,0.10, 0.30,0.50,0.80,0.90,0.99)

After a critical observations on the data and the line graph of the MSE of the respective estimators at the introduction of AR(1) of different levels i.e ρ =0, 0.1,0.3,0.5,0.8,0.9, 0.99 levels in the simulation program, the following are deduction about the Mean Square Errors (MSE) of the Estimators for sample sizes 250 (n=250). It was observed that OLS has the largest value of MSE and WLS has the lowest when AR (1) is at level 0.99. We then conclude that estimation by the WLS is the best in the presence of AR(1) of level 0.99 when the sample size is 250. It is also observed that when AR(1) is at 0.5, 0.8,0.9,0.99, WLS the best estimator to use when the sample size is 250.

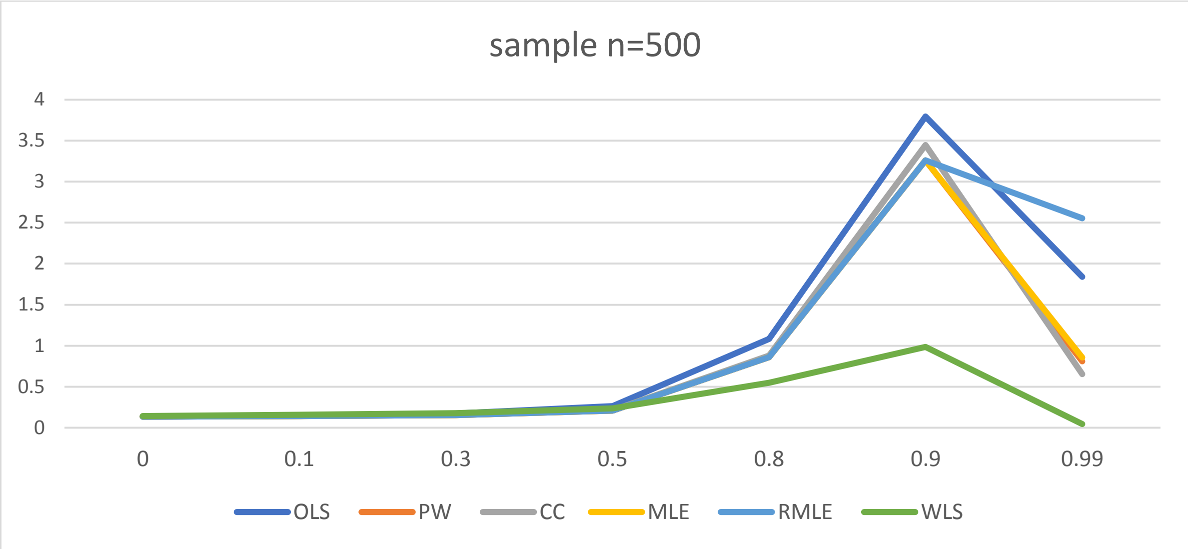

Fig. 6. Plots of the MSE on the various estimators with sample sizes of 500 across different AR disturbance (0,0.10, 0.30,0.50,0.80,0.90,0.99)

After a critical observations on the data and the line graph of the MSE of the respective estimators at the introduction of AR(1) of different levels i.e ρ =0, 0.1,0.3,0.5,0.8,0.9, 0.99 levels in the simulation program, the following are deduction about the Mean Square Errors (MSE) of the Estimators for sample sizes 500 (n=500). It was observed that RMLE has the largest value of MSE and WLS has the lowest when AR (1) is at level 0.99. We then conclude that estimation by the WLS is the best in the presence of AR(1) of level 0.99 when the sample size is 500. It is also observed that when AR(1) is at 0.5, 0.8,0.9,0.99, WLS the best estimator to use when the sample size is 500.

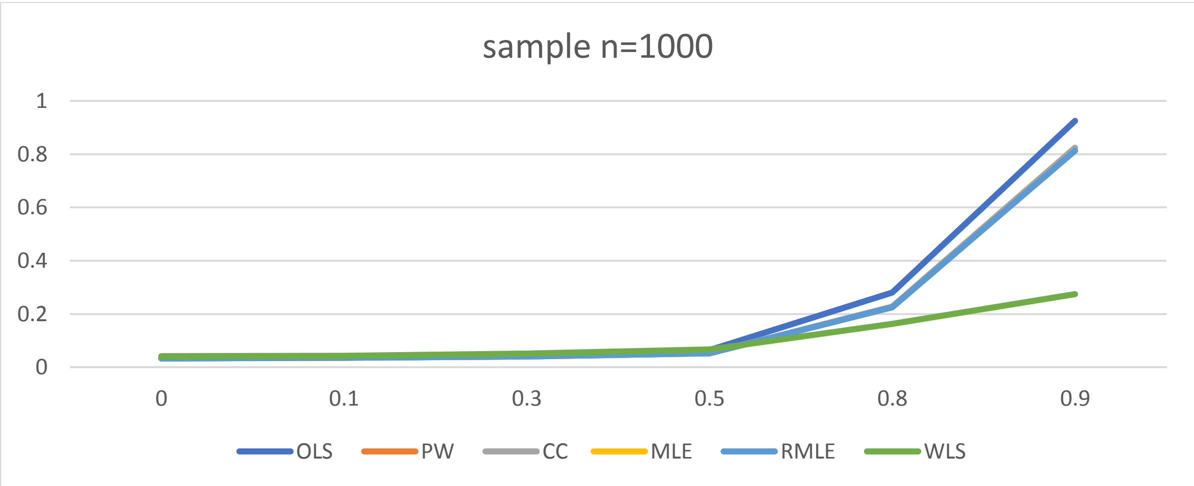

Fig. 7. Plots of the MSE on the various estimators with sample sizes of 1000 across different AR disturbance (0,0.10, 0.30,0.50,0.80,0.90,0.99)

After a critical observations on the data and the line graph of the MSE of the respective estimators at the introduction of AR(1) of different levels i.e ρ =0, 0.1,0.3,0.5,0.8,0.9, 0.99 levels in the simulation program, the following are deduction about the Mean Square Errors (MSE) of the Estimators for sample sizes 1000 (n=1000). It was observed that OLS has the largest value of MSE and WLS has the lowest when AR (1) is at level 0.99. We then conclude that estimation by the WLS is the best in the presence of AR(1) of level 0.99 when the sample size is 1000. It is also observed that when AR(1) is at 0.5, 0.8,0.9,0.99, WLS the best estimator to use when the sample size is 1000.

5 Conclusion

The results shows the following:

- That the robust methods are as efficient as the OLS if the basic assumptions are satisfied. In addition, small presence of AR1 disturbance in the data did not substantially impair robust methods like, CC, PW, MLE and RMLE. However, the robust estimators are PW, CC, RMLE and MLE in the presence of AR1 disturbance of almost 0.8 level. While, WLS shows better performance when the sample is large and the autocorrelation levels are at 0.9 and 0.99. Especially when the sample size increase and becomes relatively large. The robustness of the estimators was also seen clearly when the sample size is 250,500 and 1000 and also the level of AR1 is higher than 0.5 level. The pattern of performance of the estimators in terms of MSE was relatively similar as the level of the AR1 disturbance increase. Thus, strongly suggesting that the estimators are more effective when the AR1 disturbance are highly obvious.
- The findings of this study also strongly show that the use of the OLS when the data set is contaminated with AR1 disturbance will only lead to a misleading result, thus estimators like PW, CC, RMLE and MLE may be recommended in such situation. In extreme cases, the WLS can also be useful.
- However, in every situation of test the PW, CC, RMLE and MLE Estimator did very well. This was as a result of the fact that, these methods were developed as modifications to the OLSE and other robust methods, therefore, they are able to resist the influences of AR1 disturbance that limit the performances of OLSE and WLS methods.

6 Recommendations

In respect of our findings, the following recommendations are given on the use and application of robust methods in linear model analysis.

- We recommend the use of robust methods because of the effects of masking and swamping. Robust methods help to uncover observations which may be outliers but are behaving as usual observations, or the observations which are not outliers but because of other data points they appear as one.
- The central limit theorem is based on large sample theory and it is not in all situations that this law holds, therefore it is advisable to use robust methods, since they do not impose strict distributional assumptions on the datasets.
- Again, robust methods can be used concurrently with the Ordinary Least Squares method of Estimation as diagnostic tools.
- Finally, many statisticians do not use robust methods because, they believe these methods are computationally complex with less information on how they are used. However, we recommend the use of these methods because, there are statistical packages which now have functions for the application of robust methods.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Dow MM, Burton ML, White DR. Network autocorrelation: A simulation study of a foundational problem in regression and survey research. Social Networks. 1982;4(2):169–200.

[2] Epps TW, Epps ML. The robustness of some standard tests for autocorrelation and heteroskedasticity when both problems are present. Econometrica: Journal of the Econometric Society. 1977;745–753.

[3] Evans M. Robustness of size of tests of autocorrelation and heteroscedasticity to nonnormality. Journal of Econometrics. 1992;51(1–2):7–24.

[4] Ayinde Kayode, Olaomi JO. Robustness of Some Estimators of Linear Model with Autocorrelated Error Terms When Stochastic Regressors are Normally Distributed. Journal of Modern Applied Statistical Methods. 2008;7(1):20.

[5] Godfrey LG. Discriminating between autocorrelation and misspecification in regression analysis: An alternative test strategy. The Review of Economics and Statistics. 1987;128–134.

[6] Henshaw Jr RC. Testing single-equation least squares regression models for autocorrelated disturbances. Econometrica: Journal of the Econometric Society. 1966;646–660.

[7] Hidalgo J, Robinson PM. Adapting to unknown disturbance autocorrelation in regression with long memory. Econometrica. 2002;70(4):1545–1581.

[8] Ayinde K, Lukman AF, Rauf RI, Alabi OO, Okon CE, Ayinde OE. Modeling Nigerian Covid-19 cases: A comparative analysis of models and estimators. Chaos, Solitons and Fractals. 2020;138. Available:https://doi.org/10.1016/j.chaos.2020.109911

[9] Ayinde K, Lukman AF, Arowolo OT. Combined parameters estimation methods of linear regression model with multicollinearity and autocorrelation. Journal of Asian Scientific Research. 2015;5(5):243–250.

[10] Bumb BL, Kelejian HH. Autocorrelated and Heteroskedastic Disturbances in Linear Regression Analysis: A Monte Carlo Study. Sankhyā: The Indian Journal of Statistics, Series B. 1983;257–270.

[11] Calabrese R, Elkink JA. Estimators of binary spatial autoregressive models: A Monte Carlo study. Journal of Regional Science. 2014;54(4):664–687.

[12] DeCarlo LT, Tryon WW. Estimating and testing autocorrelation with small samples: A comparison of the C-statistic to a modified estimator. Behaviour Research and Therapy. 1993;31(8):781–788.

[13] Dubin R. Robustness of spatial autocorrelation specifications: some Monte Carlo evidence. Journal of Regional Science. 2003;43(2):221–248.

[14] Edgerton D, Shukur G. Testing autocorrelation in a system perspective testing autocorrelation. Econometric Reviews. 1999;18(4):343–386.

[15] Gosling BJ, Puterman ML. Ridge estimation in regression problems with autocorrelated errors: A monte carlo study. Communications in Statistics-Simulation and Computation. 1985;14(3):577–613.

[16] Harris P. A simulation study on specifying a regression model for spatial data: Choosing between autocorrelation and heterogeneity effects. Geographical Analysis. 2019;51(2):151–181.

[17] Hildreth CG, Lu JY. A Monte Carlo study of the regression model with autocorrelated disturbances. Rand Corp Santa Monica CA; 1969.

[18] Magee L, Ullah A, Srivastava VK. Efficiency of estimators in the regression model with first-order autoregressive errors. Routledge; 2018.

[19] Park RE, Mitchell BM. Estimating the autocorrelated error model with trended data. Journal of Econometrics. 1980;13(2):185–201.

[20] Racine MD. The simultaneous occurrence of heteroscedasticity and autocorrelation: a Monte Carlo study. Theses (Dept. of Economics)/Simon Fraser University; 1983.

[21] Zubair MA, Adenomon MO. Comparison of estimators efficiency for linear regressions with joint presence of autocorrelation and multicollinearity. Science World Journal. 2021;16(2):103–109.

[22] West KD. Another heteroskedasticity-and autocorrelation-consistent covariance matrix estimator. Journal of Econometrics. 1997;76(1–2):171–191.

[23] Wilford LL, Taylor D. The power of four tests of autocorrelation in the linear regression model. Journal of Econometrics. 1975;3(1):1–21.

[24] Thornton DL. A note on the efficiency of the Cochrane-Orcutt estimator of the AR (1) regression model. Journal of Econometrics. 1987;36(3):369–376.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
https://www.sdiarticle4.com/review-history/74094