ESTIMATING THE PARAMETERS OF LINEAR REGRESSION MODEL IN THE PRESENCE OF AUTOCORRELATED ERROR TERMS (USING BOOTSTRAP APPROACH).

Femi J. Ayoola

Department of Statistics, University of Ibadan, Nigeria.

Corresponding Emails: (fj.ayoola@.ui.edu.ng ayoolafemi@.yahoo.com)

Abstract

Autocorrelation of error terms in the linear econometric model remain a major features of most time series data. The variety of scenario in which time series observations can be caused by autocorrelated disturbances are so many that in spite of numerous analytical and empirical contribution already made on this subject, the available diagnostics leave many questions yet to be answered. This effort is channeled towards the estimation of the parameters of the linear regression models when the above two assumptions are violated.

In this work, we used the Bootstrap method to investigate the performance of five different estimation methods were considered in comparing the performance of one specification of explanatory variable in a Bootstrapping experiment with 50 replications and the sample sizes of 20 and 60 with autocorrelation levels of 0.4. 0.8 and 0.9.

The simulation results, were investigated under the finite sampling properties of Bias, Variance and Root Mean Squared Error, show that all estimators are adversely affected as autocorrelation coefficient(ρ) is close to unity. In this regard, the estimators rank as follows in ascending order of performance: HILU, CORC, ML, MLGRID and. OLS

This result helps in the choice of estimator in empirical work when the regressor and the error terms are not well behaved.

KEYWORDS: Bootstrap Methods, Estimators, Autocorrelated Error Terms, Simulation, Replication

Introduction

In the classical statistical linear model

$$Y=X\beta + U \qquad (1)$$

Where

Y = N x 1 vector; X = (N x k) matrix of rank k β=(k x 1) vector of parameters; U = (N x 1) vector of disturbance terms

Ordinary Least Square (OLS) method estimates the parameters and also to enable inferences to be made about these estimators, certain underlying assumptions are made. Two of them are the absence of autocorrelation of the error terms and that X is a matrix with nonstochastic elements and has rank k < N, hence U_i and X_j are independent for all i and j.

This research is to investigate the behavoural pattern of autocorrelation estimators in the estimation of the parameters of the linear models when the usual two assumptions are not hold.

Consider the model:

$$Y_i = X\beta + U_i;\ U_t = \rho U_{t-1} + \varepsilon_i \qquad t = 1 \ldots\ldots N$$

$$p < 1, E(\epsilon) = 0, E(\epsilon'\varepsilon) = \sigma_\varepsilon^2 I, E(U) = 0\ and\ E(U'U) = \sigma^2\Omega \qquad (2)$$

If we multiply the model (2) by some TxT nonsinglular transformation matrix P to obtain

$$PY = PX\beta + PU \qquad (3)$$

The variance matrix for the disturbance in Eq 3 is

$$E(PUUP') = \sigma^2 P\Omega P'\ since\ E(PU) = 0$$

Since we can specify P such that:

$$P\Omega P' = 1$$

Then the resulting OLS estimates of the transformed variables PY and PX in Eq. 3 have all the optimal properties of OLS and could be validly subjected to the usual inference procedures. Applying OLS to Eq. 3 results in minimizing the quadratic form.

$$U'\Omega^{-1}U = (y - X\beta)'\Omega^{-1}(Y - X\beta)$$

With optimal solutions as

$$\frac{\delta}{\delta\beta}\left(U'\Omega^{-1}U\right) = \left(X'\Omega^{-1}X\right)\beta - X'\Omega^{-1}Y = 0 \qquad (4)$$

Which gives

$$\hat{\beta}_{(GLS)} = \left(X'\Omega^{-1}X\right)^{-1} X'\Omega^{-1}Y \qquad (5)$$

With the variance-covariance matrix given by var(β) = $\sigma^2\left(X'\Omega^{-1}X\right)^{-1}$ .This estimator is known as $\hat{\beta}$ (GLS) the Aiken or Generalized Least Squares (GLS) estimator. If we assume normality for the error terms the Us, the likelihood function is given by:

$$1\left(\beta, \frac{\sigma^2}{Y}\right) = (2\pi\sigma^2)^{-T/2} - I\Omega I^{-1/2} exp\left\{\frac{(Y - X\beta)'\Omega^{-1}(Y - X\beta)}{2\sigma^2}\right\} \qquad (6)$$

Where IΩI is the determinant of Ω. Optimizing this likelihood function with respect to β means maximizing the weighted sum of squares to obtain:

$$\hat{\beta}_{(GLS)} = \left(X'\Omega^{-1}X\right)^{-1} X'\Omega^{-1}Y \qquad (7)$$

In obtaining $\hat{\beta}_{(OLS)}$ *and* $\hat{\beta}_{(GLS)}$, we assume Ωis known. When Ω is not known, we resort to estimating Ω by $\widehat{\Omega}$ in which case, we obtain an Estimated Generalized Lest Squares (EGLS) or Estimated Generalised Maximum Likelihood (EGLM) estimator and therefore;

$$\hat{\beta}_{(GLS)} = \left(X'\Omega^{-1}X\right)^{-1} X'\Omega^{-1}Y \qquad (8)$$

For this model, in Eq. 5, the TxT covariance matrix of the error vector is

$$E(UU') = \sigma_U^2 V = \sigma_U^2 \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{T-1} \\ \rho & 1 & \rho & \dots & \rho^{T-2} \\ \rho^2 & \rho & 1 & \dots . & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \rho^{T-1} & \dots & \dots & \dots & 1 \end{bmatrix} \qquad (9)$$

Where

$$\sigma_U^2 = \sigma_\varepsilon^2 / (1 - \rho^2)$$

To search for a suitable transformation matric P*, we consider the following (T-1)xT matrix P* defined by

$$\rho^* = \begin{bmatrix} -\rho & 1 & 0 & \dots & 0 & 0 \\ 0 & -\rho & 1 & \dots & 0 & 0 \\ 0 & 0 & -\rho & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -\rho & 1 \end{bmatrix}$$

Where

$$\Omega^{-1} = \frac{1}{1-\rho^2} \begin{bmatrix} 1 & -\rho & 0 & \dots & 0 & 0 \\ -\rho & 1+\rho^2 & \rho & \dots & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -\rho & 1 \end{bmatrix}_{(T-1)xT}$$

P*'P* gives $(1-\rho^2)\Omega^{-1}$ with ρ^2 instead of 1 as the first element. Next. We consider another transformation matrix P(TxT) obtained by adding a new first row with $\sqrt{1-\rho^2}$ in the first position and zero elsewhere

$$P = \begin{bmatrix} \sqrt{1-\rho^2} & 0 & 0 & \dots & 0 & 0 \\ -\rho & 1 & 0 & \dots & 0 & 0 \\ 0 & -\rho & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -\rho & 1 \end{bmatrix}_{TxT} \quad (12)$$

$P'P = (1-\rho^2)\Omega^{-1}$

P* and P differ only in the treatment of the first observation P* is much easier to use provided we are prepared to put up with its treatment of the first observation. It has been shown that when T is large, the difference is negligible but in small samples such as in this study, the difference is significant.

Such transformations give rise to different methods of estimation. These methods are broadly classified into those that use P* such as Cochrane-Orcutt (CORC) and Hildreth and LU (HILU) methods and those that use P for transformation such as Prais-Winstein (PW), Maximum Likelihood (ML) method of Beach and Mackinnon (1978) and Maximum Likelihood Gnd method (MLGRID) Nwabueze (2005a).

Many researchers have worked on autocorrelated errors. They include the early work of Cochran and Orcutt (1949). Durbin and Watson (1950, 1951, 1971), HIldreth and Lu (1960), Rao and Griliche (1969), Beach and Mackinnon (1978), Kramer (1980). Busse et al., (1994) and Kramer and Hassler (1998), to the recent works of Kleiber (2001), Kramer and Marmol (2002) and Olaomi and Ifederu (2006). Tests for detecting the presence of autocorrelation and alternative consistent methods of estimating linear models with autocorrelated disturbance terms have been proposed.

The effect of certain types of trends on explanatory variables on the relative performance of estimators has been recognized by Maeshiro (1976), Kramer (1998), Kramer and Marmol (2002), Nwabueze (2005b) and Ifederu (2006). However, some are mainly concerned with asymptotic properties. Asymtotically disregarding the first observation makes no difference but in small samples it may make a substantial difference.

However, in spite of these tests and estimation methods, a number of questions in connection with the estimation of the classical regression linear model with autocorrelated error terms and non-zero covariance between the explanatory variable and the error terms remained unanswered. These include the most appropriate method in the above named specification of the explanatory variable, the effect of the degree of correlation of the disturbance term, the effect of the degree of correlation of explanatory variable and the error terms, the effect of sample size and the sampling properties of the various estimation methods.

The answers to most of these questions would allow for correct influences to be made in linear models plagued by the scenario depicted earlier.

Materials and Methods

This study used the Bootstrap method for the investigation due to the non-zero covariance between the explanatory variable and the error terms. The problem is near intractable by analytical procedure.

The Bootstrapping is a relatively new statistical technique, which permits the assessment of variability in an estimate using the data at hand (see Efron 1979). The ideal is to resample the original observations in a suitable way, to construct "pseudo-data" on which the estimator of interest is exercised. More specifically, the theoretical distribution of a disturbance term is approximately by the empirical distribution of a set of residual. Measures of variability, confidence intervals, and even estimates of bias may then be calculated.

In the regression case, the bootstrap is useful for investigations when mathematical analysis can give only asymptotic result. Within the scope of the bootstrap are non normal errors, log structures, and generalized least squares with estimated covariance matrices. This work compares the performance of conventional asymptotic estimates of standard error to the performance of a bootstrap procedure in the setting of a single econometric equation.

The bootstrap procedure is appropriate better than the conventional asymptotic, when applied to the finite-sample situation. For a partial explanation, see Berran (1983) or Singh (1981).

In a bootstrap experiment, the experiment is artificially sets up a system and specified values for the parameters and all the independent variables, values are then generated for the random disturbances on the basis of the assumed model for some specified sample size using these values, estimates are the computed for the dependent variables at each sample point. Next, pretending as if the parameters are unknown, and using only the values of the dependent and independents variables at each sample points, one or more estimators are applied to obtain an associated estimate of parameters. The process of generating values for the disturbances, calculating values for the independent variables and estimating parameter build up empirical distribution of parameter estimates which are then used in evaluating the relative performance of the estimators in estimating the parameter values.

The design of the bootstrap studies may be summarized as follow:

1) The researcher specifies a model and ascribe specific numerical values to it parameter.
2) The researcher also specifies the distribution of the U's
3) The experimenter also specifies $X \sim U(0,1)$ for independent variable.
4) Then the experimenter uses the distribution of the U's and with random drawings from these distributions, he obtain values for all U's of each equation
5) Given the true parameters (assumed), the selected values of the explanatory variables and the chosen values of the random term, the experimenter solves the equations of the model and obtains value for the dependent variables (Y). These values are called generated values of dependent variable. For each randomly drawn value of the U's, a new generated value of the dependent variable is obtained with above procedure, the experimenter forms small samples (of generated observations of

dependent variables) which, together with the selected X values are used to estimate the coefficient by various econometric methods.

This experiments are repeated many times and thus the experimenter obtains a large number of estimates from which he obtains the coefficients for each estimation method. Furthermore, given that the true parameters are 'known' (because they have been defined in the first stage of the experiment) the bias, variances and mean square errors can be computed for the estimates of each method.

The following four Generalized Least Square (GLS) estimators, CORC, HILU, ML and MLGRID and OLS estimation methods, chose in the light of the earlier study are used. These estimators are equivalent with identical asymptotic properties. Kramer and Hassler (1998). But in small samples, such as in this study, Park and Mitchell (1980) have argued that those that use the T transformation matric (ML, MLGRID) are generally more efficient than those that use T* transformation matrix (CORC, HILU).

The degree of autocorrelation affects the efficiency of the estimators. Nwabueze (2000). Consequently, we investigated the sensitivity of the estimators to the degree of autocorrelation by varying rho (ρ) from 0.4 to 0.8 and 0.9. The effects of sample size on the estimators were also investigated by varying the samples size from 20 to 60 each replicated 50 times. Evaluation of the estimators was then done using the finite sampling properties of Bias (BIAS), Minimum Variance (VAR) and Minimum Root Mean Squared Error (RMSE).

The model

We assume a simple linear regression model:

$Y_t = \beta_0 + \beta_1 X_t + U_t$ where $U_t = \rho U_{t-1} + \varepsilon_t$,

$$/\rho/ < 1, X_t = \exp(0{,}4t)\,, U_t \to N\left(0, \frac{\sigma^2}{1-\rho^2}\right) \qquad (13)$$

t = 1, 2,….T, β_0 and β_1=(1,1)

Where

Y_t = The dependent variable and the exponential trended

X_t = The explanatory variable with U_t autoregressive of order one

ε_t = Normally distributed with zero mean and constant variance σ^2.

ρ= Stationarity parameter while the model parameters are assumed to be unity.

Nwabueze (2005b) and Olaomi and Ifederu (2006) had used this explanatory variable specification. It is chosen to allow for comparison of results.

Data Generation

A total of 27 data sets spread over three sample sizes (20 and 60) each replicated 50 times were used in generating the data for the study. Using model (13), a value U_0 was generated by drawing a random value ε_0 from N(0, 1) and dividing by $\sqrt{(1-\rho^2)}$. Successive values of ε_1 drawn from N(0,1) were used to calculate U_t X_t was generated as defined in (13). The procedure is repeated as many times as necessary to obtain 50 replications for a desired autocorrelation level, significance level and sample size. Olaomi eta'll (2004) had shown that in most Monte-carlo studies, magnitudes such as bias, variance and root mean square are not usually remarkable sensitive to the number of replications. Replication just shows the stability of estimates. Y_t is thus computed for the chosen U_t and X_t using eq. 13. The computations are made using the Microsoft Office Excel package, different estimation methods are then applied to the data using the AR procedure of the TSP (2005) package.

Performance Criteria and Tables

The summary of principal calculations for each model, estimation procedure, degree of autocorrelation of the error term and each sample size are presented in table 1 to 3

These include the bias (BIAS), variance (VAR) and the root mean square error (RMSE) of the fifty estimates of each parameter in each model for each autocorrelation coefficient e. Also calculated and displayed are the absolute sum of bias (SBIAS), the sum of variance (SVAR), sum of root mean squared error (SRMSE). These results are discussed below basing the comparison on the following properties: BIAS (SBIAS), VAR(SVAR),RMSE (SRMSE). For any parameter, the i[th] estimate is denoted by $\hat{\beta}_1$ and the true value by β_1.

Therefore, we have

$$BIAS\left(\hat{\beta}_i\right) = \frac{1}{50}\sum_{i=1}^{50}\left(\hat{\beta}_i - \beta_i\right) \quad \text{----------------------------------- } 14$$

$$RMSE\left(\hat{\beta}_i\right) = \left(\frac{1}{50}\sum_{i=1}^{50}\left(\hat{\beta}_i - \beta_i\right)^2\right)^{1/2} \quad \text{------------------------------}15$$

$$VAR\left(\hat{\beta}_i\right) = \frac{1}{50}\sum_{i=1}^{50}\left(\hat{\beta}_i - \bar{\hat{\beta}}_i\right)^2 \quad \text{-------------------------------------}16$$

Where $\bar{\hat{\beta}} = 1/50\sum_{i=1}^{50}\hat{\beta}_i$

In the model, the independent variable is specified as $X_t = \lambda X_t - 1$. Below is the summary of the experimental result for sample sizes 20 and 60 for ρ = 0.4, 0.8 and 0.9 in the tables below.

Table 1 BIAS FOR ESTIMATORS OF β

$$Y_t = \beta_0 + \beta_1 X_t, X_t = \lambda X_{t-1}, R = 50, \beta_0 \beta_1 = (1,1)$$

		N = 20			N = 60		
ρ	Estimator	β_0	β_1	SBIAS	β_0	β_1	SBIAS
0.4	OLS	-0.0029	0.00074768	0.007738	0.031080	-0.00022	0.03122
	COC	0.005556	0.050055	0.055611	0.044609	0.044609	0.089218
	HILU	0.005556	0.005556	0.055611	0.038838	0.825139	0.863971
	MLGRID	-3.5E-05	0.047164	0.0471675	0.031863	0.828411	0.860274
	ML	-3.5E-0.5	0.047164	0.0471675	0.031863	0.828411	0.860274
0.8	OLS	-1.09944	-0.3774	1.47684	-0.02230	-0.00300	0.0253
	COC	-1.20353	-0.16945	1.37298	-0.04029	-0.04027	0.08054
	HILU	-1.20207	-0.15142	1.35349	-0.04058	1.014889	1.055469
	MLGRID	-1.19772	-0.17639	1.37411	-0.04862	1.018774	1.067394
	ML	-1.18448	-0.2206	1.40508	-0.04862	1.018774	1.067394
0.9	OLS	-2.13227	-1.39017	3.52244	-0.10223	0.00340	0.1057
	COC	-2.19647	-1.16941	3.36588	-0.03676	-0.03676	0.07352
	HILU	-2.19647	-1.16941	3.36588	-0.03676	1.031726	1.065402
	MLGRID	-2.19201	-1.17633	3.36834	-0.2206	3.158013	3.378613
	ML	-2.15801	-1.22243	3.38231	-0.04681	1.036738	1.083548

Table 2: VARIANCE FOR ESTIMATORS OF β

$$Y_t = \beta_0 + \beta_1 X_t + U_t, X_t = \lambda X_{t-1}, R = 50, \beta_0 \beta_1 = (1,1)$$

		N = 20			N = 60		
ρ	Estimator	β_0	β_1	SVAR	β_0	β_1	SVAR
0.4	OLS	0.30923	0.844325	0.75248	0.312001	0.320201	0.632202
	COC	0.308315	0.622255	0.93057	0.317391	0.392708	0.710018
	HILU	0.096743	0.624467	0.72121	0.100634	0.151424	0.252058
	MLGRID	0.280878	0.770701	1.051579	0.321949	0.383823	0.705772
	ML	0.280878	0.170701	1.051579	0.321949	0.385823	0.705772
0.8	OLS	0.331236	0.844102	1.175338	0.412230	0.771002	1.183232
	COC	0.123845	0.736939	0.860784	0.411313	0.773283	1.184596
	HILU	0.124155	0.75311	0.877265	0.168919	0.584051	0.75297
	MLGRID	0.314068	0.83852	1.152568	0.415855	0.742647	1.058527
	ML	0.817771	0.311578	1.29349	0.415855	0.742647	1.58527
0.9	OLS	0.328301	0.336666	0.664961	0.512003	0.982231	1.494234
	COC	0.125081	0.733654	0.858735	0.506198	0.997864	1.503838
	HILU	0.125081	0.733654	0.858735	0.256237	1.013172	1.269409
	MLGRID	0.316182	0.836998	1.15318	0.512973	0.986582	1.499555
	ML	0.381775	0.819481	1.201256	0.512973	0.986582	1.49955

Table 3: RMSE FOR ESTIMATORS OF β

$$Y_t = \beta_0 + \beta_1 X_t + U_t, X_t = \lambda X_{t-1}, R = 50, \beta_0\beta_1 = (1,1)$$

		N = 20			N = 60		
ρ	Estimator	β_0	β_1	SRMSE	β_0	β_1	SRMSE
0.4	OLS	0.556092	0.921908	1.478000	0.30854	0.001200	0.30974
	COC	0.555289	0.790418	1.345707	0.565138	0.62825	1.193388
	HILU	0.311086	0.790252	1.101338	0.319598	0.912293	1.231891
	MLGRID	0.529979	0.879162	1.409141	0.568299	1.03445	1.602749
	ML	1.334372	1.229496	2.563868	0.568299	1.03445	1.602749
0.8	OLS	1.240966	0.993242	2.234208	0.75892	0.002011	0.760931
	COC	1.253929	0.875016	2.128945	0.642601	0.880287	1.522888
	HILU	1.252653	0.880931	2.133584	0.412997	1.270453	1.68348
	MLGRID	1.322344	0.932533	2.255973	0.646698	1.334372	1.98107
	ML	1.309412	0.930825	2.240237	0.646698	1.334372	1.98107
0.9	OLS	2.207909	1.664106	3.872015	1.523320	0.000015	1.523335
	COC	2.224763	1.449544	3.67307	0.712524	0.999608	1.723858
	HILU	2.224763	1.449544	3.674307	0.507532	1.441399	1.9491522
	MLGRID	2.26298	1.490218	3.75326	0.717749	1.43576	2.153509
	ML	2.244726	1.521123	3.765849	0.717749	1.43576	2.153509

Discussion of Results

In our estimates, we observe that the SBIAS of the OLS is lower than the SBIAS of the GLS methods for N=20 and 60 and for increasing value of ρ. Although the sum of bias of the GLS methods compare favourably with one another especially for large sample N=60, for small sample N=20 especially when ρ is large (ρ = 0.8 and 0.9), COC and HILU have edge over MLGRID and ML in respect of this property.

It is observed that both OLS and GLS are negatively biased for $\hat{\beta}_0(.)$ and $\hat{\beta}_1(.)$ when N = 20, N = 60 and ρ = 0.8 and 0.9. The pattern is mixed up for decreasing value of ρ For increased sample, the SBIAS for the estimators decreases and increases as the sample size increases. It is also observed that the estimates $\hat{\beta}_1$ are very close to the parameter values for both the OLS and the GLS estimators.

A more significant results is that the SVAR for HILU of OLS for N = 20 and ρ = 0.4 is much smaller than, SVAR of COC, MLGRID and ML from table 2, we find that at N = 20, ρ = 0.9, the SVAR of COC and HILU equal the same thing. While the SVAR of OLS = 0.512003, MLGRID = 1.1538 and ML = 1.201256.

We note that the SVAR of OLS is 0.664961 as against 0.858735 and 1.15318 for CO, HILU and MLGRID respectively. This implies that OLS dominate COC, HILU and MLGRID.

This result is not expected because one would expect COC or HILU to be more efficient than OLS in finite samples when there is autocorrelation in the error terms. However, the result confirms the works of Park and Mitchel (1980). In the sequel, the word dominate is used to qualify a magnitude which is smaller than another one and therefore it is preferred. The meaning is going to remain the same throughout this work.

For large sample, we also notice that the picture is quite similar as N – increases for (N = 60 and ρ = 0.9), the SVAR of COC = 1.503838, HILU = 1.269409, and OLS = 1.024431.

This implies that OLS dominates HILU and other estimator.

The HILU on the other hand have some edge over OLS for small value of ρ (ρ = 0.4 and 0.8) but as ρ increases, the other GLS estimators increases in their superiority over for all the sample sizes considered.

From table 3, for N = 60 and ρ = 0.4 we notice that on the basis of minimum RMSE, SRMSE of OLS has slight edge over other estimators.

From these estimates of SRMSE, we observe that COC and HILU mostly dominates OLS even when ρ = 0.9 and N = 20, which suggests that in this respect, the effect of decreasing autocorrelation on the comparative performance of the estimators is significant.

It can therefore, be reasonably assumed that small autocorrelation, COC and HILU significantly dominate OLS.

In our summary, the major conclusion which could be drawn from our experiments for the model are the following: -

a. For both large and small values of autocorrelation, the GLS methods are inferior to OLS in respect of SBIAS property.

b. On the basis of VAR and RMSE properties, the OLS dominates for large sample with ρ = 0.9 but for small value of autocorrelation COC and HILU dominate OLS especially as the degree of autocorrelation decreases.

c. Another important conclusion is that the scope coefficients of the estimators are very close to their parameter values for both the OLS and the GLS methods and they are less sensitive to the degree of autocorrelation than the constant term.

We found that the estimators conform to the asymptotic properties of estimates considered. This is seen at all levels of autocorrelation and at all significant levels. The estimators rank in the decreasing order of conformity with the observed asymptotic behavior as follows: OLS,ML, MLGRID, HILU and CORC. This ranking is contrary to that of Olaomi (2006).

We also note that ML and MLGRID have very similar behavioral pattern, the same for CORC and HILU as observed in the finite sampling properties of Bias, Variance and the RMSE. ML and MLGRID are better than both CORC and HILU as also observed by Park and Mitchell (1980).

References

Abrehamse, A.P.J and Kverts, J. (1971) “New estimators of Disturbances in Regression Analysis” **JASA**. **66**, 71-74

Anderson, T.W. and Mentz, R.P. (1980), "On the Structure of the Likelihood Function of Autoregressive and Moving Average Models," *Journal of Time Series*, 1, 83--94.

Ansley, C.F., Kohn, R., and Shively, T.S. (1992), "Computing *p*-values for the Generalized Durbin-Watson and Other Invariant Test Statistics," *Journal of Econometrics*, 54, 277--300.

Ayoola, F.J et’all (2007), ‘Estimation of Parameters of Linear Regression Model with Autocorrelated Error Terms: Using Bootsrap Approach’MSc. Dissertation (Unpublished)

Baillie, R.T. and Bollerslev, T. (1992), "Prediction in Dynamic Models with Time-Dependent Conditional Variances," *Journal of Econometrics*, 52, 91--113.

Balke, N.S. and Gordon, R.J. (1986) "Historical Data," in *The American Business Cycle*, ed. R.J. Gordon, Chicago: The University of Chicago Press, 781--850.

Beach, C.M. and MacKinnon, J.G. (1978), "A Maximum Likelihood Procedure for Regression with Autocorrelated Errors," *Econometrica*, 46, 51--58.

Beach, C.M and Mackinnon, J.S (1978) “A Maxium Likelihood Procedure for regression with Autocorrected Errors” **ECTRA 46**. No. 1, 51-57

David A. Freedmanard and Stephen C. Peters (1984) “Bootstrapping a Regression Equation” **JASA 79**: 97-105

Kramer W, (1980) “Finite Sample Efficiency of Ordinary Least Squares in the Linear Regression Model with Automcorrelated Errors” **JASA**, **73**, 1005 -1009.

Kramer and Hassler (1998) "Limiting efficiency of OLS Vs GLS when Regressors are Fractionally Integrated" Technical Report No 1, SFB475, Fachbreich Statistik, University Dorumund

Maeshiro Asatoshi (1976) "Autoregressive Transformation, Trended Independence Variables and autocorrelated Disturbance Terms" Review of Economic and Statistics, 58: 497 – 500

Martins Feldstein (1974) "Errors in Varieble" **JASA 69**: 900

Maddala, G.S (1977) Econometrics. New York, McGraw Hill

Maeshiro Asatoshi (1976) "Autoregresive transformation, treded independent variables and Autocorrelated disturbance terms" **Review of Economics and Statistics, 58, 497 – 500**

Mooney, C.Z and Rebert D.D (1993) Bootstrapping: A Non- Parametric Approximation to Statistical Inference (Sage Publications)

Nwabueze, J.C (2000) "Estimation of Parameters of Linear Regression Models with Autocorrelated Errors Terms" Unpublished Ph.D Thesis, University of Ibadan Nigeria.

Olaomi J.O. and A. Ifederu, 2008. Understanding Estimators of Linear Regression Model with AR(1) Error Which are Correlated with Exponential Regressor. *Asian Journal of Mathematics & Statistics,1:14-23.* **DOI:** 10.3923/ajms.2008.14.23 **URL:** http://scialert.net/abstract/?doi=ajms.2008.14.23

Olaomi, J.O (2004) "Estimation of Single Equation Models with Autogressive Errors which are also Correlated with the Regressor Unpublished Ph.D Thesis, University of Ibadan.

Olubusoye, O.E. (2005): Lecture Note on Advanced Econometric Methods. (Unpublished)

Park, R.E and Mitchell, B.M (1980) "Estimating the Autocorrelated Errors Model with Trended Data" Journal of Economics **JE, 13**, 185 – 201

Paris, S.J and Weinstein, C.B 1954 "Trend Estimators and Serial Correlation" Unpublished Cowles Commission, Discussions Paper, Chicago

Rao, P. and Grilliches, Z (1969) "Small Sample Properties of Several Two-Stage Regression Methods in the Context of Autocorralated Errors" Journal of American Statistical Association: 64: 251-272

Richard F.Gunstand Robert L. Manson (1977) "Biased Estimation in Regression" **JASA 72**, 616 -617.

Spitzer, J.J and Baillie, R.T. (1983) "Small Sample Properties of Predictions from the Regression Model with Autoregressive Errors" **JASA 78**: 258 -263.

Spitzer, J.J (1979) “Small Properties of Nonlinear Least Squares and Maximum Likelihood Estimators in the ontext of Autocorralated Errors”. Journal of Amerian Statistical Association, 74: 126-138

Theil, H (1957) “Specification Errors and the Estimation of Ecomics Relations” Review of the International Statistical Institute, Vol. 25, pg 41-41

Theil, H. (1971) Principles of Ecomonics, New Yolk, Wiley TSP (2005) Users Guid and Refrence Mannual Time Series Processor