Journal of Science and Multidisciplinary Research
Volume 5, No. 2, 2013
ISSN: 2277-0135

CONSEQUENCES OF VIOLATING THE ASSUMPTIONS OF OLS IN THE PRESENCE OF AUTOCORRELATION

Uchendu, Bartholomew .A.
Department of Maths/Statistics
Federal Polytechnic, Nekede, Owerri, Nigeria
E-mail: uchendubartholomew@yahoo.com

Abstract: The consequences of applying OLS to a relationship with autocorrected disturbances are qualitatively similar to those already derived for the heteroscedastic case, namely unbiased but inefficient estimation and invalid inference procedures. As in the case of heteroscedasticity, in the presence of autocorrelation, the OLS estimators are still linear unbiased as well as consistent and asymptotically normally distributed, but they are no longer efficient (i.e., minimum variance). In the case of heteroscedasticity, we distinguish two cases and the possible cause and sources of autocorrelation. The violation of the assumptions of normality may have significant consequences in applying OLS and such consequences include substantial loss in efficiency, inflating the precision or accuracy of the estimators by underestimating the standard error of β. Moreover, violating of the assumptions of normally of the error term is important in econometric analysis. If this assumption is violated, then the basis of hypothesis testing breaks down. In this direction, a large number of possible tests for normality and robust estimator have been suggested. The assumption of lack of autocorrelation or serial correlation of the error term implies that the disturbance covariance at all possible pairs of observation points are zero. Violation provides the basis of for this research because it affects the consistency of the OLS estimators. Models with such disturbances are widespread, as applied econometrics especially in modeling of economic data.

Keywords: Heteroscedasticity, Autocorrelation, OLS, Minimum Variance, Monte-Carlo

INTRODUCTION

A general source of autocorrelated disturbances is the fact that the disturbance represents the net influence of omitted explanatory variables. Economic theory cannot prescribe an exhaustive list of explanatory variables to be included in a relation and in any case, data limitations often curtail this number of variables that can be included. Exclusion of variables would not of itself impact autocorrelation to the disturbance term unless the excluded variables were autocorrelated. Even then autocorrelation in one explanatory variable might offset that in another. However, economic variables tend to be nonrandom over time and also to move roughly in phase so that excluded variables may impact autocorrelation to the disturbances term. A second source of autocorrelation may be a misspecification of the form of the relationship.

A third possible source of autocorrelated disturbances may be measurement error in the dependent variable. Economic statisticians typically have various formalized routines and procedures for estimating economic magnitudes. The sequential publication of revised estimates is eloquent testimony of the fact that the creators of the series believed their products to contain some error and indeed a series becomes definitive simply when the statisticians stop

revising it, which is not to say that it is then free of error. It is unlikely that the estimating procedures produces errors which are random from period to period and so, letting the y vector denote the observed Y values and y_* the true Y values generated by the mechanism $x\beta + \mu$, we have

$$y = y_* + v = x\beta + (u + v) \quad (1)$$

Where v is a vector of measurement errors. In the observed relationship, the disturbance term is u + v, which may exhibit autocorrelation through u or v or both.

What happens to the OLS estimators and their variances if we introduce autocorrelation in the disturbances by assuming that $E(\mu_t\, \mu_{t+s}) \neq 0 (s \neq 0)$ but retain all the other assumptions of the classical model? Noting that we are now using the subscript t on the disturbances to emphasize that we are dealing with time series data. In the two-variable regression mode

$$Y_t = \beta_1 + \beta_2 X_t + U_t \quad (2)$$

we must assume the mechanism that generate μ_t for $E(\mu_t\, \mu_{t+s}) \neq 0 (s \neq 0)$ is too general an assumption to be of any practical use. As a starting point, or first approximation, one can assume that the disturbances or error terms are generated by the following mechanism.

$$\mu_t = \rho\mu_{t-1} + \varepsilon_t \quad -1 < \rho < 1 \quad (3)$$

Where ρ is known as the coefficient of autocovariance and where ε_t is the stochastic disturbance term such that it satisfies the standard OLS assumptions namely;

$$\begin{aligned} &E(\varepsilon_t) = 0 \\ &var(\varepsilon_t) = \sigma_t^2 \\ &cov(\varepsilon_t, \varepsilon_{t+s}) = 0 \quad S \neq 0 \end{aligned} \quad (4)$$

In the engineering literature, an error term with the preceding properties is often called a white noise error term. What equation 3 postulate is that; the value of the disturbance term in period t is equal to ρ times its value in the previous period plus a purely random error term.

The scheme 3 is known as a Markov first-order autoregressive scheme, or simply a first - order autoregressive scheme, usually denoted as AR(1).The name autoregressive is appropriate because equation(3) can be interpreted as the regression of μ_t on itself lagged one period. it is first order because μ_t and its immediate past value are involved, that is the maximum lag is 1. If the model were $\mu_t = \rho_1\mu_{t-1} + \rho_2\mu_{t-2} + \varepsilon_t$, it would be an AR(2),or second-order autoregressive scheme and so on.

We note that ρ, the coefficient of Autocovariance in equation (5) can also be interpreted as the first-order coefficient of autocorrelation, or more accurately, the coefficient of autocorrelation at lag 1.given the AR(1) scheme, it can be shown that

$$var(\mu_t) = \in (\mu_t^2) = \frac{\sigma_\varepsilon^2}{1 - \rho^2} \quad (5)$$

$$cov(\mu_t, \mu_{t+s}) = \epsilon(\mu_t, \mu_{t-s}) = \rho^s \frac{\sigma_\varepsilon}{1-\rho^2} \quad (6)$$

$$cov(\mu_t, \mu_{t+s}) = \rho^s \quad (7)$$

Where $cov(\mu_t, \mu_{t+s})$ means covariance between error terms s periods apart and where $cor(\mu_t, \mu_{t+s})$ means correlation between error terms s period apart. We observe that because of symmetry property of covariances and correlations, $cov(\mu_t, \mu_{t+s}) = cov(\mu_t, \mu_{t-s})$ and $cor(\mu_t, \mu_{t+s}) = cor(\mu_t, \mu_{t-s})$

Since ρ is a constant between -1 and +1, equation (5) shows that under the AR(1) scheme, the variance of μ_t is still homoscedastic, but μ_t is correlated not only with its immediate past values but its values several periods in the past.

One reason we use the AR (1) process is not only because of its simplicity compared to higher-order AR schemes, but also because in many applications, it has proved to be quite useful. Additionally, a considerable amount of theoretical and empirical work has been done on the AR(1)scheme.

In the two variable regression model given in equation (2).the OLS estimator of the slope coefficient is

$$\hat{\beta}_2 = \frac{\sum x_t y_t}{\sum x_t^2} \quad (8)$$

and its variance is given by

$$var\ (\hat{\beta}_2) = \frac{\sigma^2}{\sum x_t^2} \quad (9)$$

Where the small letters as usual denote deviation from the mean values

Now under the AR (1) scheme, it can be shown that the variance of this estimator is

$$var\ (\hat{\beta}_2) = \frac{\sigma^2}{\sum x_t^2}\left[1 + 2\rho\frac{\sum x_t x_{t-1}}{\sum x_t^2} + 2\rho^2\frac{\sum x_t y_{t-2}}{\sum x_t^2} + \cdots + 2\rho^{n-2}\frac{\sum x_1 y_n}{\sum x_t^2}\right] \quad (10)$$

Where $var\ (\hat{\beta}_2)$ AR(1) means the variance of $\hat{\beta}_2$ under a first-order autoregressive scheme.

A comparison of equation (10) with equation (9) shows the former is equal to the latter times a term that depends on ρ as well as the sample autocorrelation between the values taken by the regressor x at various lags.

To give some idea about the difference between the variances given in equations (9) and equation (10), assume that the regressor x also follows the first-order autoregressive scheme with a coefficient of autocorrelation of r. Then it can be shown that equation (10) reduces to

$$var(\hat{\beta}_2)AR(1) = \frac{\sigma^2}{\sum x_t^2}\left(\frac{1+r\rho}{1-r\rho}\right) = var(\hat{\beta}_2)OLS\left(\frac{1+r\rho}{1-r\rho}\right) \quad (11)$$

If for example, r = 0.7 and ρ = 0.9, using equation (11), we can check that $var(\hat{\beta}_2)AR(1)$ = 4.405 $var\hat{\beta}_2 OLS$. To put it another way;

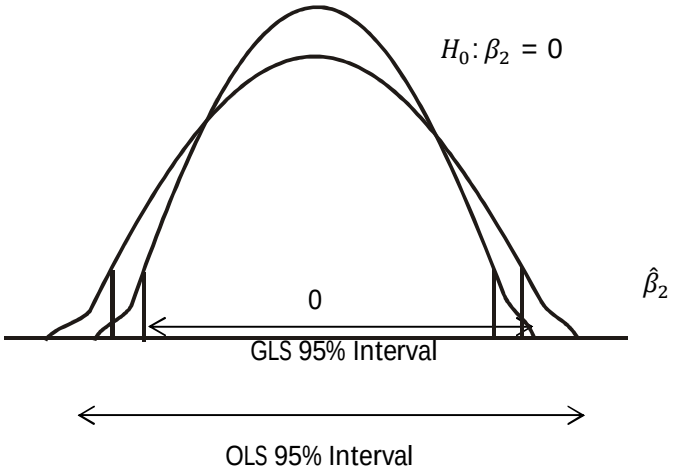
$$var(\hat{\beta}_2)OLS = \frac{1}{4.405} var(\hat{\beta}_2)AR(1) = 0.2270\ var(\hat{\beta}_2)AR(1)$$

That is the usual OLS formula (i.e. equation 9) will underestimate the variance of $\hat{\beta}_2 AR$ (1) by about (1 – 0.2270) = 77 percent. As can be seen, this result is specific for the given values of r and ρ. But the point of this exercise is to warn that a blind application of the OLS for procedures to compute the variances and standard errors of the OLS estimators could give serious misleading results and a major consequence.

OLS ESTIMATION ALLOWING FOR AUTOCORRELATION

$\hat{\beta}_2$ is not BLUE and even if we use $var(\hat{\beta}_2)_{AR1}$, the confidence intervals derived from there are likely to be wider than those based on the GLS procedures. This result is likely to be the case if the sample sizes increase indefinitely. That is $\hat{\beta}_2$ is not asymptotically efficient. The implication of this finding for hypothesis testing is clear. We are likely to declare a coefficient statistically insignificant (i.e. not different from zero), even though in fact (i.e. based on the correct GLS procedure) it may be. This difference can be seen clearly from figure 1. In this figure, we show the 99% OLS [AR (1)] and GLS confidence intervals assuming that true β_2 = 0. Consider a particular estimate of β_2, say b_2. Since b_2 lies in the OLS confidence interval, we could accept the hypothesis that true β_2 is zero with 95 percent confidence .But if we were to use the (correct) GLS confidence interval; we could reject the null hypotheses that true β_2 is zero, for b_2 lies in the region of rejection. To establish intervals and to test hypothesis, one should use GLS and not OLS even though the estimators derived from the later are unbiased and consistent.

FIG.1. GLS and OLS 95% Confidence intervals



OLS ESTIMATION DISREGARDING AUTOCORRELATION

The situation is potentially serious if we do not only use β_2 but also continue to use $var(\hat{\beta}_2) = \frac{\sigma^2}{\sum X_t^2}$ which completely disregards the problem of autocorrelation, that is, we mistakenly believe that the usual assumptions of the classical model hold true. Errors will arise for the following reasons:

A. The residual variance $\hat{\sigma}^2 = \frac{\sum \hat{\mu}_t^2}{n-2}$ is likely to underestimate the true σ^2

B. As a result, we are likely to overestimate R^2

C. Even if σ^2 is not underestimated, $var(\hat{\beta}_2)$ may underestimate $var(\hat{\beta}_2)$AR(1) equation (10), its variance under (first-order), autocorrelation, even though the latter is inefficient compared to $var(\hat{\beta}_2)_{GLS}$

D. Therefore, the usual t and F tests of significance are no longer valid and if applied, are likely to give seriously misleading conclusions about the statistical significance of the estimated regression coefficients.

To establish some of these properties under the two variable model, the classical assumption

$$\hat{\sigma}^2 = \frac{\sum \hat{U}_i^2}{n-2}$$

Provides an unbiased estimator of σ^2, that is, $E(\hat{\sigma}^2) = \sigma^2$. But if there is autocorrelation given AR(1),it can be shown that

$$E(\hat{\sigma}^2) = \frac{\sigma^2\{n - [2/(1-\rho)] - 2\rho r\}}{n-2} \qquad (12)$$

where

$$r = \sum_{t=1}^{n=1} x_t\, x_{t-1} \Big/ \sum_{t=1}^{n} x_t^2,$$

which can be interpreted as the (sample) correlation coefficient between successive values of the X's. If ρ and r are both positive (not an unlikely assumption for most economic time series), it is apparent from equation (12) that $E(\hat{\sigma}^2) < \sigma^2$, that is, the usual residual variance formula, on the average, will underestimate the true σ^2.

In otherwords, $\hat{\sigma}^2$ will be biased downward. Needless to say, this bias in $\hat{\sigma}^2$ will be transmitted to $var(\hat{\beta}_2)$ because in practice we estimate the latter by the formula $\hat{\sigma}^2/\sum x_t^2$. But even if σ^2 is not underestimated, $var(\hat{\beta}_2)$ is a biased estimator of $var(\hat{\beta}_2)AR(1)$, which can be readily seen by comparing equation (9) with equation (10), for the two formulars are not the same. As a matter of fact, if ρ is positive (which is true of most economic time series) and the x's are positively correlated (also true of most economic time series), then it is clear that

$$var(\hat{\beta}_2) < var(\hat{\beta}_2)AR1 \qquad (13)$$

that is, the usual OLS variance of $\hat{\beta}_2$ underestimates its variance under AR1 (see equation 11). Therefore, if we use $var(\hat{\beta}_2)$, we shall inflate the precession or accuracy (i.e. underestimate the standard error) of the estimator $\hat{\beta}_2$. As a result, in computing the t-ratio as $t = \hat{\beta}_2/se(\hat{\beta}_2)$

(Under the hypothesis that $\beta_2 = 0$), we shall be overestimating the t-value and hence the statistical significance of the estimated β_2. The situation is likely to get worse if additionally σ^2 is underestimated, as noted previously.

To see how OLS is likely to underestimate σ^2 and the variance of $\hat{\beta}_2$, let us conduct the following Monte Carlo experiment. Suppose in the two variable model, we "know" that the true $\beta_1 = 2\ and\ \beta_2 = 0.6$. Therefore, the stochastic PRF i.e. (Population Regression function) is

$$Y_t = 2.0 + 0.6X_t + \mu_t \qquad (14)$$

Hence

$$E(Y_t/X_t) = 2.0 + 0.6X_t \qquad (15)$$

which gives the true population regression line

Let us assume that μ_t are generated by the first-order autoregressive scheme as follows:

$$\mu_t = 0.8\mu_{t-1} + \varepsilon_t \qquad (16)$$

where ε_t satisfy all the OLS assumptions. We assume further for convenience that the ε_t are normally distributed with zero mean and a unit variance. Equation (16) postulates that the successive disturbances are positively correlated with a coefficient of autocorrelation of +0.8, a rather high degree of dependence.

Table 1: A Hypothetical Example of Positively Autocorrelated Error Terms

X_t	ε_t	$\mu_t = 0.8\mu_{t-1} + \varepsilon_t$
0	0	$\mu_0 = 7\ (assumed)$
1	0.185	$\mu_1 = 0.8(7) + 0.185 = 5.785$
2	-0.487	$\mu_2 = 0.8(5.785) - 0.487 = 4.141$
3	2.555	$\mu_3 = 0.8(4.141) + 2.555 = 5.867$
4	-1.811	$\mu_4 = 0.8(5.8678) - 1.811 = 2.883$
5	0.815	$\mu_5 = 0.8(2.883) + 0.815 = 3.121$
6	0.480	$\mu_6 = 0.8(3.121) + 0.480 = 2.977$
7	-0.942	$\mu_7 = 0.8(2.977) - 0.942 = 1.440$
8	2.312	$\mu_8 = 0.8(1.440) + 2.312 = 3.464$
9	0.244	$\mu_9 = 0.8(3.464) + 0.244 = 3.015$
10	0.059	$\mu_{10} = 0.8(3.015) + 0.059 = 2.471$

Now using a table of random normal numbers with zero mean and a unit variance, we generated 10 random numbers as shown in table 1 and then by the scheme (16),we generated μ_t. To start off the scheme, we need to specify the initial value of μ, say μ_0 =7

Suppose the value of x are fitted at 1, 2, 3, ..., 10. Then, given these x's, we can generate a sample of 10 Y values from equation (14) and the value of μ_t given in table in table 1

Table 2: **Generation of Y Sample Values**

X_t	μ_t	$Y_t = 2.0 + 0.6X_t + \mu_t$
1	5.785	Y_1 = 2.0 + 0.6(1) + 5.785 = 8.385
2	4.141	Y_2 = 2.0 + 0.6(2) + 4.141 = 7.341
3	5.8678	Y_3 = 2.0 + 0.6(3) + 5.8678 = 9.668
4	2.883	Y_4 = 2.0 + 0.6(4) + 2.883 = 7.283
5	3.121	Y_5 = 2.0 + 0.6(5) + 3.121 = 8.121
6	2.977	Y_6 = 2.0 + 0.6(6) + 2.977 = 8.577
7	1.440	Y_7 = 2.0 + 0.6(7) + 1.440 = 7.640
8	3.464	Y_8 = 2.0 + 0.6(8) + 3.464 = 10.264
9	3.015	Y_9 = 2.0 + 0.6(9) + 3.015 = 10.415
10	2.471	Y_{10} = 2.0 + 0.6(10) + 2.471 = 10.471

μ_t data are obtained from table 1.

If we regress Y on X, we obtain the following (sample) regression:

$$\hat{Y}_t = 7.3232 + 0.2715X_t \qquad (17)$$
$$r = 0.6421, r^2 = 0.4123$$
$$\hat{\sigma}^2 = 1.0838$$

Whereas the true regression line is given by equation (15).The fitted regression line distorts the true regression line. It seriously underestimates the true slope coefficient but overestimates the true intercept. We note that the OLS estimators are still unbiased.

The $\hat{\mu}_i$ are generally close to the fitted line which is due to the OLS procedure but deviate substantially from the true Population Regression Line (PRL). Hence, they do not give a correct picture of μ_i. To gain some insight into the extent of underestimation of true σ^2, suppose we conduct another sampling experiment Keeping the x_t and ε_t given in tables 1 and 2 let us assume that ρ = 0, that is, no autocorrelation. The new sample of Y values thus generated is given in table 3

Table 3: Sample of Y Values with Zero Serial Correlation

X_t	$\varepsilon_t = \mu_t$	$Y_t = 2.0 + 0.6X_t + \varepsilon_t$
1	0.185	2.785
2	-0.487	2.713
3	2.555	6.355
4	-1.811	2.589
5	0.815	5.815
6	0.480	6.080
7	-0.942	5.258
8	2.312	9.112
9	0.244	7.644
10	0.059	8.059

The regression based on table 3 is as follows:

$$\hat{Y}_t = 2.173 + 0.630\ X_t \qquad (18)$$
$$r = 0.8196, r^2 = 0.6717$$
$$\hat{\sigma}^2 = 2.004$$

This regression in (18) is much closer to the "truth "because the Y's are now essentially random. We observe that $\hat{\sigma}^2$ has increased from 1.0838 at ρ = 0.8 to 2.004 at ρ = 0. This result is in line with the theoretical result considered in Monte-Carlo experiment

CONCLUSIONS

The solution to be adopted in each particular case depends on the source of autocorrelation. Thus if the source is omitted variables the appropriate procedure is to include these variables in the set of explanatory variables. The simplest way to detect whether autocorrelation is due to omitted variables is to regress the residues $\varepsilon's$, against variables which, on a priori grounds might be relevant explanatory variables of the phenomenon being studied.

Similarly if the source of autocorrelation is the mis-specification of the mathematical form of the relationship, the relevant approach is to change the initial (linear) form. This can be investigated by regressing the residuals against higher powers of the explanatory variable(s),or by computing a "linear in logs" form and re-examining the resulting new residuals.

Only when the above sources of autocorrelation have been ruled out should we accept that the true μ's are temporally dependent. For this case (of true autocorrelation) the appropriate procedure is the transformation of the original data so as to produce a model whose random variable satisfies the assumptions of classical least squares, and consequently the parameters can be optimally estimated with this method. Once autocorrelation is detected by applying any relevant test, the appropriate corrective procedure is to obtain an estimate of the ρ's and apply OLS to a set of transformed data. The transformation of the original data depends on the pattern of the autoregressive structure.

REFERENCES

Damodar, N.G. (2009): Basic Econometrics 5th Edition, McGraw-Hill International Edition.

Johnson J. (1984): Econometric Methods McGraw-Hill Book Company University of California.

Koutsyiannis, A. (1977): Theory of Econometrics Second Edition Macmillan Publishers.

Reference to this paper should be made follows: Uchendu, Bartholomew .A. (2013), Consequences of Violating the Assumptions of OLS in the Presence of Autocorrelation. *J. of Sciences and Multidisciplinary Research,* Vol. 5, No. 2, Pp. 93 – 100.