GLOBAL JOURNAL OF PURE AND APPLIED SCIENCES VOL. 7, NO 3, JUNE 2001: 535-540
COPYRIGHT (C) BACHUDO SCIENCE CO. LTD. PRINTED IN NIGERIA. ISSN 1118-0579

WEIGHTED LEAST SQUARES METHOD AND ASSOCIATED TIME SERIES DATA PROBLEMS

A. C. EKE; D. ENI, and J. ATSU

(Received 10 May 2000; Revision accepted 16 March 2001)

ABSTRACT

An apparently high R^2 in any set of data and especially in a time series data which may suggest a good fit may not necessarily imply that the basic assumptions of a regression model have been well met. Even in the presence of the problems of heteroscedasticity, serial correlation and multi collinearity, it is possible to have a very high R^2. This paper seeks to address the problems arising from heteroscedasticity and serial correlation in time series data and to suggest how the problems can be dealt with so as to have data that will yield valid statistical inference.
Key Words: Time series, autocorrelation, multi-collinearity heteroscodasticity transformation.

INTRODUCTION

Studies and findings by Jokomba (1998), Doguwa (1994) have highlighted and emphasized how easily one can be led to accept an invalid statistical inference as the result of R^2 values. Many B.Sc and M.Sc projects display a lot of this wrong statistics inference. It is either the consequences of the errors in inference due to this is not known or not understood. Consideration is given to the linear regression model in this study.

THE REGRESSION MODEL

Given a linear model with stochastic regressors;

$$Y = \beta x + \epsilon \quad \ldots \quad (1)$$

Where Y is an (n x 1) vector of observations.
X is an (n x p) matrix of known form.
β is a (p x 1) vector of parameters.
ϵ is an (n x 1) vector of errors.
The relationship between Y and X is linear.
X values are on un-correlated with the error term.
The error term has constant variance for all observations. This is the assumption of homoscedasticity of error terms. The error term has zero expectation and is normally distributed.

The Generalized Least Squares (GLS) method as developed by Aitken (1934) is appropriate for the estimation of models in which the random term is heteroscedasticity or auto correlated. This method is simply the application of the Ordinary Least Square (OLS) methods to a set of transformed variables. The transformation required in **Y** and **X** depends on the form of heteroscedasticity or of serial correlation. If the type of heteroscedasticity or the relation between the successive values of ϵ is known a priori, or established by

EKE, A. C., Department of Statistics, The Polytechnic, Calabar (Died November 2000)
ENI D., Department of Mathematics, The Polytechnic, Calabar
ATSU J., Department of Mathematics, The Polytechnic, Calabar

experimentation, we use the information to transform the original relation, so that the standard assumptions of Least Squares are satisfied.

RELATIONSHIP BETWEEN OLS AND MAXIMUM LIKELIHOOD (ML)

A better understanding of GLS will be gained if a brief examination between the Ordinary Least Squares (OLS) and maximum likelihood methods are made. The OLS is based on the principle of minimization of the sum of squared residuals. It is to be observed that this sum of squared residuals is identical to the expression that appears in the exponent of the likelihood function of the sample observations except for the term $^1/_{2\sigma}{}^2$ which under the assumption of normality, zero mean, constant variance can be taken out of the summation in the exponent and left out of the differentiation. The joint probability of the n sample values of Y is the likelihood function.

$$\sigma^{-n} (2\pi)^{-n/2} e^{-\epsilon'\epsilon/2\sigma 2} \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad (2)$$

The likelihood function is maximized when the negative exponent is minimized. The expression is identical to the sum of squared residuals which we minimize when applying OLS. Therefore the OLS estimates are identical to the maximum likelihood estimates, provided that ϵ satisfies the standard assumptions.

Examining the residual sum of squares we see that all observations are treated symmetrically in the sense that all are given equal weights, each observation is given a functional weight. Thus we say that in OLS we minimize an un-weighted sum of squares. However, if heteroscedastic or serial correlation exists, the un-weighted sum of squares is inappropriate. Each sample observation should be given a different weight and the appropriate procedure is to minimize a weighted sum of squared residuals, where the weights are given to incorporate the effects of the various products. This is done by the method of Generalized Least Squares otherwise called Weighted Least Squares (WLS) method.

CONSEQUENCES OF THE VIOLATION OF THE ASSUMPTIONS OF THE REGRESSION MODEL.

Though estimates obtained will be unbiased if all other assumptions hold, the effects of serial correlation can be far and wide. In fact, it can destroy the optimal properties of OLS. Some of the possible effects can be summarized as follows (Jokomba, 1998).

a) Precision of parameter estimates may be impaired in any particular sample, especially in a small sample.
b) The variance of regression parameters are likely to understate the corresponding true variance, since the estimated variance of errors is likely to understate the true variance of error especially for positive serial correlation.
c) The application of the classical least squares technique yields a sample variance of error which is large relative to what can be obtained by some other estimating technique.
d) Predictions made from the estimated least squares are inefficient.
e) Consequently, inference procedures will be invalid.

If the assumption of homoscedastic disturbance is not fulfilled, we have the following consequences:

i. The formula of the variances of the coefficients to conduct tests of significance and construct confidence intervals will be wrong.

ii. If there is heteroscedasticity, the OLS estimates are not unbiased estimators; that is, they are inefficient in small samples. Furthermore, they are asymptotically inefficient.

iii. The coefficient estimates would still be statistically unbiased, that is even if the error terms are heteroscedastic; the estimates will have no statistical bias. Their expected value will be equal to the parameters.

iv. The prediction of Y (for a given value of X) based on the estimates from the original data, would have a high variance; that is the prediction would be inefficient.

With the obvious consequences of the violation of the two assumptions of non serial correlation and homoscedasticity, further attempt is made now to address the need for Weighted Least Squares method. In the case of positive heteroscedasticity, the variance of the error term increases with increasing values of the variables (Fig. 1). As the scatter of observations widens, values of X and Y give a less precise indication of where the true regression line lies. Therefore we should give less attention to these observations than to the more precise ones depicted on the left side of the diagram (Koutsoyianis, 1977).

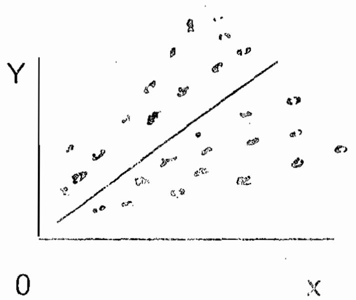

Fig 1: Deviation Sum of Squares

This is achieved by dividing the deviation by the value of the corresponding variance of the error term. Instead of applying OLS which minimizes the sum of the squares, we rather apply the Generalized Least Squares which minimizes the 'generalized' sum. In this sum each squared deviation is weighted by a factor σ^{-2}, so that as σ^2 increases, the weight σ^{-2} decreases, and thus less attention is given the less precise deviations on the right side. In this way, observations with large variances (less reliable observations), are "discounted" in the process of fitting the regression line.

This procedure is based on the maximum likelihood function, which of course assumes a different form when heteroscedasticity, and/or auto-correlation of the error term occurs. Thus if the variance of the error terms is not constant but changes with the observation we cannot take σ^2 out of the summation in the exponent of the likelihood function. With heteroscedasticity present in the structural equation, the likelihood function is

$$L = (2\pi)^{-n/2} (\sigma^{-n}) e^{-\epsilon'\epsilon/2\sigma 2} \quad \ldots \quad (3)$$

It can be seen that the sum in the exponent of the likelihood is identical with the 'generalized' sum of squared residuals which is minimized by the method of GLS.

Similarly, when the error terms are auto-correlated (but homoscedatic) the expression to be minimized changes (Kane, 1968; Wonnacott and Wonacott, 1970).

AITKEN TRANSFORMATION METHOD

A transformation procedure by Aitken (1935) is still vogue and is an alternative method. It involves the transformation of the variables of the original model so as to produce a new model which satisfies the standard assumptions of the random variable and to which GLS can be applied. When the random variable is heteroscedastic the original model is

$$Y = \beta X + \epsilon^* \quad \ldots \quad (4)$$

whose variance is none constant but varies with the X. If the form of $f(x_1, x_2, \ldots, x_k)$ is known in the original model, the required transformation

$$\frac{Y}{f(x_1, x_2, \ldots, x_k)} \doteq b_0 + \{f(x_1, x_2, \ldots, x_k)\}^{-1} + \epsilon' \quad \ldots (5)$$

can be achieved. This transformation gives rise to a homoscedastic error term. With auto-correlated error terms, the original model is

$$Y_t = b_0 + b_1 x_{1t} + \ldots + b_k x_{kt} + f(\epsilon_{t-1}, \epsilon_{t-2} \ldots) \quad \ldots (6)$$

As usual the appropriate transformation depends on the form of auto-correlation, that is on the relationship between the successive values of the error term. If this known transformation of the original model followed by the application of OLS, the new model whose error term will be corrected for serial correlation can be effected.

To appreciate the power and effect of the transformation, an example is presented below using data on savings and income as reported by Koutsoyiannis (1977) pp 192.

Table 1 Personal Savings and Income.

Period	1	2	3	4	5	6	7	8	9	10
Savings ($)	264	105	90	131	122	107	406	503	431	588
Income (x)	8,777	9,210	9,954	10,508	10,979	11,912	12,747	13,499	14,269	15,522
PERIOD	11	12	13	14	15	16	17	17	18	19
Savings ($)	898	950	779	819	1,222	1,702		1,578	1,654	1,400
Income (X)	27,670	28,300	27,430	29,560	28,150	32,100	32,500	35,250	33,500	
Period	29	30	31							
Savings ($)	1,900	2,100	38,200							
Income (X)	36,000	36,200	38,200							

*Adopted from Koutsoyiannis (1977). Pp. 192

applying OLS to the data of Table 1, we obtain

$$S_t = -644.1 + 0.085 X_t, \quad R^2 = 0.903$$
$$(117.6) \quad (0.005)$$

By the Goldfield and Quandt test, two subsets of data can obtained from the data. Application of OLS to each subset gives

$$S_1 = -738.84 + 0.088x, \quad R^2 = 0.787$$
$$(189.4) \quad (0.015)$$

$\sum e_1^2 = 144, 71.5$

$$\text{and } S_2 = \underset{(709.8)}{1141.07} + \underset{(0.022)}{0.029}x, \quad R^2 = 0.152$$

$$\sum e_2^2 = 769{,}899.2$$

The ratio of the unexplained variation gives

$$F^* = \frac{\sum e_2^2}{\sum e_1^2} = \frac{769{,}899.2}{144{,}771.5} = 5$$

The theoretical value of F at the 5 percent level of significance with.

$$V_1 = V_2 = \frac{n - c - 2k}{2} = \frac{31 - 9 - 2(2)}{2} = 9$$

degrees of freedom is 3.18. Since F* > F0.05, we reject the assumption of homoscedasticity. We assume that the pattern of heteroscedasticity is

$$\sigma^2 = K^2X^2$$

So that the appropriate transformation of the original model is

$$\frac{S_t}{X_t} = bo\frac{1}{X_t} + \frac{\epsilon_t}{X_t}$$

Applying OLS to the new variables we obtain

$$\frac{S_t}{X_t} = bo\frac{1}{X_t} + b_t = \frac{\underset{(71.27)}{-718.88}}{X_t} + \underset{(0.004)}{0.088}$$

wit! ^{2} = 0.770

T' value of the)earman correlation coefficient of the transformed equation r' = 0.22. Its standard error of 0.14 shows that the transformation has eliminated heteroscedasticity. Thus the new savings function is

$$S_t^* = \underset{(71.27)}{-718.88} + \underset{(0.004)}{0.088}X_t, \quad R^2 = 0.770$$

as compared with the old equation

$$S_t = -644.1 + 0.085X_t, \quad R^2 = 0.903$$

It is to be observed that the R^2 for the transformed data is lower than that of the original data as a result of weighting.

CONCLUSION

In general the method of GLS involves the minimization of some expression which includes the residuals properly weighted. If the assumption of homoscedasticity and or of serial independence of the error term does not hold in any particular equation the correct procedure is to weight the elements of the sum of squared deviations before minimizing this sum. The weights defined by the variances and co-variances of the error terms are not known in practice. Thus for the application of GLS Aitken's alternative approach should be adopted.

REFERENCES

Aitken, A.C., 1935. On least squares and linear combination of observations. Proceedings of the Royal Society of Edinburgh, 5: 42 – 48.

Doguwa S., 1994. The multiplier approach to money supply in Nigeria. Central Bank of Nigeria Economic and Financial, Review 32 (4): 55 – 59.

Jokoba H. Br., 1998. Serial correlation and the estimation of true economic relationships. First Bank of Nigeria PLC Bi-annual Review, 6 (12): 55 – 59.

Kane, E.J., 1968. Economic Statistics and Econometrics, Harper International, New York. 358 - 9

Koutsoyainnis A., 1977. Theory of econometrics: An introductory exposition of Econometric methods. Macmillan Press Ltd. London, 471 – 2.

Uchendu, O.A., 1993. Interest rate policy savings and investment in Nigeria. Central Bank of Nigeria and Financial Review 31 (1): 55 – 58

Uchendu, O.A., 1994: The determinants of external debt service in Africa. Central Bank of Nigeria, Economics and Financial Review 32, (1); 55 – 59

Wonnacott R.J. & Wonnacott T.H. 1970: Econometrics. John Wiley and Sons Inc., London, 322 – 6.