MONASH

WP 11/88

ISSN 0729-0683

ISBN 0 86746 735 5

ANCORA IMPARO

MONASH UNIVERSITY

EFFICIENCY OF OLS RELATIVE TO C-O FOR TRENDED x AND POSITIVE AUTOCORRELATION COEFFICIENT

Asraul Hoque

Working Paper No. 11/88

July 1988

DEPARTMENT OF ECONOMETRICS

ISSN 0729-0683

ISBN 0 86746 735 5

EFFICIENCY OF OLS RELATIVE TO C-O FOR TRENDED x AND POSITIVE AUTOCORRELATION COEFFICIENT

Asraul Hoque

Working Paper No. 11/88

July, 1988

DEPARTMENT OF ECONOMETRICS, FACULTY OF ECONOMICS AND POLITICS

MONASH UNIVERSITY, CLAYTON, VICTORIA 3168, AUSTRALIA.

Efficiency of OLS Relative to C-O for Trended x

and Positive Autocorrelation Coefficient

by Asraul Hoque*

Abstract: It is well known that the OLS estimator, though unbiased, is inefficient in the presence of autocorrelated disturbances. Further, it is also widely accepted that C-O (Cochrane-Orcutt) estimator is more efficient than OLS estimator. However, Kadiyala (1968) and Maeshiro (1976, 1978) have argued that OLS is more efficient than C-O when the independent variable is trended and the autocorrelation coefficient is positive. We re-examine this issue and show that C-O is more efficient than OLS for the model without an intercept term.

I. Introduction

It is widely accepted that the OLS estimator, though unbiased is, in general, inefficient in the presence of autocorrelated disturbances. In the case of AR(1) disturbances, Cochrane and Orcutt (1949) have suggested an effective transformation of the model. It is believed that OLS applied to the transformed model is more efficient than is OLS applied to the original model. However, some researchers have cast doubts on this belief, for example, Kadiyala (1968) and Maeshiro (1976, 1978). Kadiyala has shown analytically that when positive autocorrelation is present in the errors, there exist cases in which the efficiency of the estimator obtained by C-O (Cochrane-Orcutt) method is less than that of OLS. However, he is essentially estimating the population mean when he considers the independent variable x to be a

---

* Department of Econometrics, Monash University, Clayton, Victoria, Australia.

column of n ones, and this case is not very interesting. Maeshiro (1976) contends that if the independent variables are trended and the autocorrelation coefficient is positive then OLS is more efficient than the C-O estimator. The reason is as follows: after the transformation the total variation of the independent variable (i.e. $X'X$) is reduced when the autocorrelation coefficient is positive leading to some loss in efficiency; on the other hand, there is some gain in efficiency due to the transformation and Maeshiro claims that this trade-off works in favour of OLS. We re-examine this issue and show that C-O is more efficient than OLS in some cases.

However, this paper is not intended to argue that OLS is useless in finite sample analysis. On the contrary, it could be very useful especially in the context of dynamic models, for example, see Lahiri (1975), Hoque (1985a, 1985b), Hoque and Peters (1986), Hoque et al (1986), Carter and Ullah (1979), Maddala (1976), and Hoque et al (1988). We only argue that OLS may not be more efficient than C-O in all cases even when the explanatory variable is trended and the autocorrelation coefficient is positive.

Recently, Glezakos (1980) has made an attempt to show that the conclusions reached by Maeshiro (1976) are not generally valid. He made an experiment with the stochastic linear trend case and concluded that OLS is not uniformly more efficient than C-O. Then he compared the efficiency of OLS relative to GLS and found OLS to be uniformly less efficient than GLS. Since the difference between C-O and GLS lies in the treatment of just one observation, Glezakos argued that C-O should be more efficient than OLS. Maeshiro (1980) does not agree with Glezakos and mentions some further numerical calculations using US GNP series supporting his previous conclusions. Regarding the superiority of GLS to OLS he fully agrees with Glezakos but he does not support his view that lack of first observation in C-O method may not make much of a difference. I tend to support Maeshiro's view in this connection, because the inclusion of first observation in the transformed model will certainly enhance the total variation of the explanatory variable to a great extent.

My purpose here is to explore some cases where Maeshiro's conclusions may not be valid. All the basic assumptions of Maeshiro are maintained and both arithmetic and geometric trends are considered. The only difference is that we shall not consider the intercept term. But there are many economic models that do not consider the intercept term like Friedman's consumption function and other functions incorporating long term relationships. Since our results show that C-O is uniformly more efficient than OLS both in linear and geometric trend, Maeshiro's conclusions may not be useful in estimating the above class of models.

Outline of the paper: In section II, we examine the efficiency of OLS relative to C-O under AR(1) disturbances and arithmetic trend. In section III, we consider the efficiency of OLS relative to C-O for geometric trend. In section IV, the numerical results are interpreted. In section V, we provide a conclusion. Finally, we add an appendix in which some primary estimates have been illustrated so that one can check the results very easily.

II. Relative Efficiency for Arithmetic Trend

We consider the following model

$$Y_t = \beta x_t + u_t , \qquad t = 1, 2, \ldots, T \tag{1}$$

The following assumptions about u_t are made

(i) $u_t = \rho u_{t-1} + \epsilon_t$, $0 < \rho < 1$

(ii) $\epsilon_t \sim N(0,\sigma_\epsilon^2)$, for all t

(iii) $\mathrm{Var}(u_t) = \sigma^2$, for all t

where $\sigma^2 = \sigma_\epsilon^2/(1-\rho^2)$.

The OLS estimator of β in (1) is given by

$$\hat{\beta} = \sum_{t=1}^{T} x_t y_t \Big/ \sum_{t=1}^{T} x_t^2 \tag{2}$$

Now, the variance of $\hat{\beta}$ is

$$\mathrm{Var}(\hat{\beta}) = \frac{\sigma^2}{\left(\sum_1^T x_t^2\right)^2} \left[\sum_1^T x_t^2 + 2\rho \sum_2^T x_t x_{t-1} + 2\rho^2 \sum_3^T x_t x_{t-2} + 2\rho^3 \sum_4^T x_t x_{t-3} + \ldots + 2\rho^{T-1} x_T x_1 \right] \tag{3}$$

where $\sigma^2 = \sigma_\epsilon^2/(1-\rho^2)$.

Further, C-O estimator of β can be written as (assuming certain known values of ρ)

$$b = \sum_2^T x_t^* y_t^* \Big/ \sum_2^T x_t^{*2} \tag{4}$$

where $y_t^* = y_t - \rho y_{t-1}$ and $x_t^* = x_t - \rho x_{t-1}$.

Now, the variance of b is

$$\mathrm{Var}(b) = \sigma_\epsilon^2 / \sum_2^T (x_t^2 + \rho^2 x_{t-1}^2 - 2\rho x_t x_{t-1}) \tag{5}$$

Let us introduce the following two kinds of trend in x.

(a) $x_t = \gamma t, \quad \gamma > 0$ (6)

(b) $x_t = \theta x_{t-1} = \theta^i x_{t-i}, \quad \theta > 0$ (7)

To evaluate the efficiency of OLS relative to C-O we define the following ratio

$$e = \mathrm{var}(\hat{\beta}) / \mathrm{var}(b)$$

If e is greater than one, OLS is less efficient than C-O and vice versa.

We first consider the case when $x_t = \gamma t$. For simplicity, we take $\gamma = 1$; however, a particular value of γ does not affect our results. Equation (3) can now be written (after simplification) as

$$\mathrm{var}(\hat{\beta}) = \sigma^2 / \sum_1^T t^2 \,.\, (1 + \rho - C + C\rho - 2\rho^T) / (1-\rho) \tag{8}$$

where

$$C = [2\rho + 2\rho \sum_2^T t + 10\rho^2 + 4\rho^2 \sum_3^T t + 28\rho^3 + \ldots] / \sum_1^T t^2$$

Now, for reasonably large T, C and ρ^T can be ignored and thus,

$$\mathrm{var}(\hat{\beta}) \approx \sigma_\epsilon^2 / \sum_1^T t^2 \, (1-\rho)^2 \tag{9}$$

Again, from equation (5)

$$\mathrm{var}(b) = \sigma_\epsilon^2 / [(1-\rho)^2 + d] \sum_1^T t^2 \tag{10}$$

where

$$d = (2\rho \sum_{1}^{T} t - T^2 \rho^2 - 1) / \sum_{1}^{T} t^2$$

For large T, d can be ignored and thus,

$$v(b) \approx \sigma_{\epsilon}^2 / (1-\rho)^2 \sum_{1}^{T} t^2 \qquad (11)$$

Consequently, the efficienty of OLS relative to C-O is obtained by combining equations (9) and (11) which is

$$e \approx var(\hat{\beta}) / var(b) \approx 1$$

Therefore, for reasonably large T, the OLS and C-O estimators are approximately equally efficient when $x_t = \gamma t$ and ρ is positive. This is supported by Hannan (1970) and Chipman (1979). We present the <u>exact</u> numerical values of e for different ρ and T using equations (8) and (10) in Table A.

III. Relative Efficiency for Geometric Trend

We now consider the case when $x_t = \theta x_{t-1}$ which can be equivalently written as $x_{t-i} = x_t / \theta^i$.

Equation (3) can now be written as

$$var(\hat{\beta}) = \sigma^2 / \sum_{1}^{T} x_t^2 \,.\, (1 + \rho^* - C^* + C^*\rho^* - 2\rho^{*T}) / (1-\rho^*) \qquad (12)$$

where

$$C^* = [2\rho^* x_1^2 + 2\rho^{*2} \sum_{1}^{2} x_t^2 + 2\rho^{*T-1} \sum_{1}^{T-1} x_t^2] / \sum_{1}^{T} x_t^2$$

and $\rho^* = \rho/\theta$.

For reasonably large T, C* and ρ^{*T} can be ignored (assuming $\rho^* < 1$) and thus,

$$\mathrm{var}(\hat{\beta}) \approx \sigma_\epsilon^2 \,/\, \sum_1^T x_t^2 \,.\, (1+\rho^*) \,/\, (1-\rho^*)(1-\rho^2) \tag{13}$$

Further, from equation (5)

$$\mathrm{var}(b) = \sigma_\varepsilon^2 \,/\, \sum_1^T x_t^2 \,(1 + \rho^{*2} - 2\rho^* + d^*) \tag{14}$$

where

$$d^* = (2\rho^* - \rho^{*2} - 1)\, x_1^2 \,/\, \sum_1^T x_t^2$$

Now, for large T, d* can be ignored and thus,

$$\mathrm{var}(b) \approx \sigma_\epsilon^2 \,/\, \sum_1^T x_t^2 \,(1-\rho^*)^2 \tag{15}$$

Hence, combining equations (13) and (15), we have

$$e \approx (1-\rho^{*2}) \,/\, (1-\rho^2) \tag{16}$$

Here e is greater than 1 when θ is greater than 1 because in that case ρ^{*2} is less than ρ^2. However, when θ is less than 1, e will be less than 1 implying that OLS will be more efficient than C-O. Thus, in this special case ($\theta < 1$), Maeshiro's claim is absolutely valid. But this is relatively uninteresting in the sense that this implies geometrically declining time series which is very rare in economic variables. However, values of θ much away from unity to either side is unrealistic. For practical purposes, θ, very close to one but greater than one (such as $\theta = 1.1$) may be of interest.

We now present the <u>exact</u> numerical values of e in Table B for different values of ρ^* and T using equations (12) and (14). Here six values of θ and ρ have been combined to generate six values of ρ^*.

IV. Numerical Results

Looking at Table A we see that e is always greater than unity but its value decreases for all ρ as T increases. Further, OLS gets worse as ρ increases for a given T. We have already seen that for large T, the OLS and C-O estimators are approximately equally efficient when x follows an arithmetic trend. This is confirmed by Table A where e approaches unity as T grows large.

Table B gives two different pictures. For $\theta < 1$, e is always less than 1 confirming Maeshiro's claim that OLS is more efficient than C-O when x is geometrically trended and ρ is positive. This is also the conclusion one can reach from our equation (16). But this case is relatively less significant, since this implies geometrically declining time series which is not common in economic variables. Now, for $\theta > 1$, OLS becomes less efficient as ρ^* increases and becomes increasingly inferior to C-O when sample size increases unlike the arithmetic trend case. For $\theta = 1.1$ and 1.18, e is always greater than unity and increasing with sample size. But the values corresponding to $\theta = 1.18$ may not be taken very seriously. Because this gives rise to unrealistic economic time series. However, $\theta = 1.04$ and 1.1 could yield interesting and useful series. For $\theta = 1.04$, e is less than 1 for two sample sizes, 10 and 20. But from sample size 20 e gets larger and exceeds unity pretty quickly. Thus, Maeshiro may be correct for sample size up to 20 for $\theta = 1.04$. In general, his results do not hold for $\theta > 1$. Our approximation in equation (16) confirms this.

V. Conclusion

We have shown that C-O method leads to more efficient estimator than that of OLS even when x is trended and ρ is positive in the model without the intercept term. Both for arithmetic and geometric trends our results contradict both Kadiyala (1968) and Maeshiro (1976) who claim that OLS is more efficient than C-O for trended x and positive ρ. However, in the geometric trend when $\theta < 1$, Maeshiro's claim is valid. But this case is relatively uninteresting in the sense that this implies geometrically declining time series which is not common in economic variables. Chipman (1979) has also arrived at the same conclusion as ours in connection with arithmetic trend.

References

Carter, R.A.L., and Aman Ullah (1979), "The Finite Sample properties of OLS and IV estimators in special rational distributed lag models", Sankhya, Vol. 41D, 1-18.

Chipman, J.S. (1979), "Efficiency of LS estimation of linear trend when residuals are autocorrelated", Econometrica, Vol. 47, 115-28.

Cochrane, D., and G.H. Orcutt (1949), "Application of LS regression to relationships containing autocorrelated errors", JASA, Vol. 44, 32-61.

Glezakos, C. (1980), "Autocorrelation and Trended explanatory variables: A comment", Review of Economics and Statistics, Vol. 62, 484-87.

Hannan, E.J. (1970), Multiple Time Series, Chap. VII, New York, Wiley.

Hoque, A. (1985a), "The exact moments of forecast error in the general dynamic model", Sankhya, Vol. 47B, 128-43.

―――――――― (1985b), "Finite Sample Analysis of the first Order Autoregressive Model", Calcutta Statistical Assoc. Bulletin, Vol. 34, 51-63.

Hoque, A., Y. Kataoka and H. Miyashita (1986), "The exact moments of OLS estimators for Koyck distributed lag models", International Economic Review, Vol. 27, 245-60.

Hoque, A., and Tom A. Peters (1986), "Finite Sample Analysis of the ARMAX Models", Sankhya, Vol. 48B, 266-83.

Hoque, A., J.R. Magnus and B. Pesaran (1988), "The exact multiperiod mean square forecast error for the first order autoregressive model", Journal of Econometrics, Vol. 39, 000-000.

Kadiyala, K.R. (1968), "A Transformation used to circumvent the problem of Autocorrelation", Econometrica, Vol. 36, 93-96.

Lahiri, K. (1975), "Multiperiod predictions in dynamic models", International Economic Review, Vol. 16, 699-711.

Maddala, G.S. (1976), Econometrics, McGraw-Hill, New York.

Maeshiro, A. (1976), "Autoregressive Transformation, trended independent variables and Autocorrelated disturbance terms", Review of Economics and Statistics, Vol. 58, 497-500.

———— (1978), "A Total Reappraisal of the Small Sample Econometrics of AR(1) disturbances", University of Pittsburg (mimeo).

———— (1980), "Autocorrelation and Trended Explanatory Variables: A Reply", Review of Economics and Statistics, Vol. 62, 487-89.

Relative Efficiencies [V(OLS)/V(C-O]

Table A: Values of e for Arithmetic Trend

ρ / T	.2	.4	.6	.8	.9	.98
10	1.0056	1.0355	1.1027	1.4487	1.7105	5.3165
20	1.0048	1.0234	1.0669	1.1879	1.4071	3.3478
30	1.0035	1.0168	1.0482	1.1387	1.2971	2.5678
40	1.0027	1.0130	1.0379	1.1101	1.2326	2.1861
50	1.0023	1.0106	1.0306	1.0919	1.1953	1.9769
60	1.0019	1.0089	1.0263	1.0781	1.1725	1.8206

Table B: Values of e for Geometric Trend

ρ	.2	.4	.6	.8	.9	.98
θ	.8	.9	.98	1.04	1.1	1.18
T ρ^*	.25	.44	.61	.77	.82	.83
10	.5132	.5978	.6768	.7380	1.1310	5.5662
20	.5181	.6330	.7924	.9868	1.5798	7.4422
30	.5192	.6370	.8307	1.0725	1.6954	7.7903
40	.5194	.6372	.8484	1.1070	1.7187	7.8457
50	.5194	.6373	.8584	1.1211	1.7233	7.8547
60	.5194	.6373	.8602	1.1329	1.7283	7.8601

Appendix

We have used the following equation to calculate the exact values of e in Table A using equations (8) and (10):

$$e = V(\hat{\beta})/V(b) = (1+\rho-C+C\rho-2\rho^{T})(1+\rho^{2}-2\rho+d)/(1-\rho)(1-\rho^{2})$$

where C and d have been defined in the text. As an illustration, let us give the exact expressions for C and d for sample size 20. From equation (10) we have

$$d = (2\rho \sum_{1}^{20} t - 20^{2}.\rho^{2} - 1)/ \sum_{1}^{20} t^{2}$$

and from equation (8) we have

$$C = (2\rho \sum_{1}^{20} t + \rho^{2}(10 + 4 \sum_{3}^{20} t) + \rho^{3}(28 + 6 \sum_{4}^{20} t) + \rho^{4}(60 + 8 \sum_{5}^{20} t)$$

$$+ \rho^{5}(110 + 10 \sum_{6}^{20} t) + \rho^{6}(182 + 12 \sum_{7}^{20} t) + \rho^{7}(280 + 14 \sum_{8}^{20} t)$$

$$+ \rho^{8}(408 + 16 \sum_{9}^{20} t) + \rho^{9}(570 + 18 \sum_{10}^{20} t) + \rho^{10}(770 + 20 \sum_{11}^{20} t)$$

$$+ \rho^{11}(1012 + 22 \sum_{12}^{20} t) + \rho^{12}(1300 + 24 \sum_{13}^{20} t) + \rho^{13}(1638 + 26 \sum_{14}^{20} t)$$

$$+ \rho^{14}(2030 + 28 \sum_{15}^{20} t) + \rho^{15}(2480) + 30 \sum_{16}^{20} t) + \rho^{16}(2992 + 32 \sum_{17}^{20} t)$$

$$+ \rho^{17}(3570 + 34 \sum_{18}^{20} t) + \rho^{18}(4218 + 36 \sum_{19}^{20} t) + \rho^{19}(4940$$

$$+ 38x20)) / \sum_{1}^{20} t^{2}$$

A sample calculation of C and d for T = 20 is given below:

ρ	.2	.4	.6	.8	.9	.98
C	.0457	.1617	.5388	2.5488	7.1550	18.4263
d	.0233	.0359	.0373	.0275	.0184	.0092

Further, for Table B we have used the following equation using equations (12) and (14):

$$e = V(\hat{\beta})/V(b) = (1+\rho^*-C^*+C^*\rho^*-2\rho^{*T})(1+\rho^{*2}-2\rho^*+d^*)/(1-\rho^*)(1-\rho^2)$$

From equation (14) we have (for sample size 20, let us say)

$$d^* = (2\rho^* - \rho^{*2} - 1)\, x_1^2 \Big/ \sum_1^{20} x_t^2 .$$

and from equation (12) we have

$$\begin{aligned} C^* = \Big(&2\rho^* x_1^2 + 2\rho^{*2} \sum_1^{2} x_t^2 + 2\rho^{*3} \sum_1^{3} x_t^2 + 2\rho^{*4} \sum_1^{4} x_t^2 + 2\rho^{*5} \sum_1^{5} x_t^2 \\ &+ 2\rho^{*6} \sum_1^{6} x_t^2 + 2\rho^{*7} \sum_1^{7} x_t^2 + 2\rho^{*8} \sum_1^{8} x_t^2 + 2\rho^{*9} \sum_1^{9} x_t^2 \\ &+ 2\rho^{*10} \sum_1^{10} x_t^2 + 2\rho^{*11} \sum_1^{11} x_t^2 + 2\rho^{*12} \sum_1^{12} x_t^2 + 2\rho^{*13} \sum_1^{13} x_t^2 \\ &+ 2\rho^{*14} \sum_1^{14} x_t^2 + 2\rho^{*15} \sum_1^{15} x_t^2 + 2\rho^{*16} \sum_1^{16} x_t^2 + 2\rho^{*17} \sum_1^{17} x_t^2 \\ &+ 2\rho^{*18} \sum_1^{18} x_t^2 + 2\rho^{*19} \sum_1^{19} x_t^2\Big) \Big/ \sum_1^{20} x_t^2 \end{aligned}$$

A sample calculation for C^* and d^* for $T = 20$ is given below:

ρ	.2	.4	.6	.8	.9	.98
θ	.8	.9	.98	1.04	1.1	1.18
ρ^*	.25	.44	.61	.77	.82	.83
C^*	.285	.4708	.5391	.7906	.5934	.2771
d^*	-.2025	-.0604	-.0108	-.0011	-.0001	-.00001

Recent Working Papers
(Available on Request)

1987

1/87 Maxwell L. King and Merran A. Evans, "Locally Optimal Properties of the Durbin-Watson Test".

2/87 R.D. Snyder, "Computational Aspects of Kalman Filtering with a Diffuse Prior Distribution".

3/87 Merran A. Evans, "The Twelfth Order Analogue of the Durbin-Watson Test".

4/87 Radhey S. Singh, "A Family of Improved Estimators in Linear Regression Models with Errors Having Multi-variate Student-t Distribution".

5/87 Radhey S. Singh, "Estimation of Error Variance in Linear Regression Models with Errors Having Multi-variate Student-t Distribution with Unknown Degrees of Freedom."

6/87 Radhey S. Singh, "Melo and Empirical Bayes Estimators for Reciprocals of Population Means and Regression Coefficients."

7/87 A.R. Tremayne, "A Synthesis of Econometrics and Time Series Methods to Model and Forecast UK Demands for Imports".

8/87 Merran A. Evans and Maxwell L. King, "A Further Class of Tests for Heteroscedasticity".

9/87 Asraul Hoque, "Size of Holding and efficiency in Bangladesh: An Application of a Random Coefficient Model".

10/87 Brett A. Inder, "Bias in the Ordinary Least Squares Estimator in the Dynamic Linear Regression Model with Autocorrelated Disturbances".

11/87 Russel J. Cooper and Keith R. McLaren, "Approaches to the Solution of Intertemporal Consumer demand Models".

12/87 Esme Preston, "On Choosing a Technical Wordprocessor".

1988

1/88 John Preston and Esme Preston, "Two Papers on Simple Word Processors for the PC".

2/88 Esme Preston, "Multilingual and Mathematical Text Processing".

3/88 Esme Preston and John Preston, "The Accountant's PC: Getting Started".

4/88 Maxwell L. King, "Towards a Theory of Point Optimal Testing".

5/88 Ralph D. Snyder, "Statistical Foundations of Exponential Smoothing".

6/88 Grant H. Hillier, "On the Interpretation of Exact Results for Structural Equation Estimators".