The Korean Communications
in Statistics Vol. 4, No. 2, 1997
pp. 353 - 358

Pitfalls in the Application of the COTE in a Linear Regression Model with Seasonal Data

Seuck Heun Song[1] and YouSung Park[2]

Abstract

When the disturbances in the linear regression model are generated by a seasonal autoregressive scheme the Cochrane Orcutt transformation estimator (COTE) is a well known alternative to Generalized Least Squares estimator (GLSE). In this paper it is analyzed in which situation the Ordinary Least Squares estimator (OLSE) is always better than COTE for positive autocorrelation in terms of efficiency which is here defined as the ratio of the total variances.

1. Introduction

For a linear regression model with first order autocorrelated disturbances, a variety of estimators for regression coefficients have been proposed in the literature (Judge et. al (1985)). One of most commonly used estimator for this situation is the Cochrane Orcutt transformation estimator (COTE) as a well known alternative to GLSE due to its intuitive and computational simplicity. However, several authors including Kadiyala (1968), Poirier (1978), Maeshiro (1976, 1979, 1980), Park and Mitchell (1980), Doran (1981), Kraemer (1982), Puterman (1988), Hoque (1989) and Stemann and Trenkler (1993) have studied the effect of omitting the first transformed observation and they have pointed out that the efficiency of COTE can be very adversely affected if the sample is relatively small. All these papers deal with the case of strongly trended data.

In this paper we show that the OLSE is better than COTE for positive autocorrelation in terms of efficiency which is here defined as the ratio of the total variances. In section 2, we consider a regression model with seasonal data and define the efficiency. We derive the efficiency function of OLSE relative to the COTE and investigate the propeties of efficiency function in section 3.

1) Assistant Professor, Department of Statistics, Duksung Women's University, Seoul 132-174, Korea.
2) Associate Professor, Department of Statistics, Korea University, Seoul, 136-701, Korea.

2. Model and Cochrane Orcutt Estimator

We consider the following linear regression model

$$y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + \cdots + \beta_s X_{st} + u_t, \quad t = 1, 2, \cdots, T, \tag{2.1}$$

where $X_{2t}, \cdots, X_{st}$ denote seasonal dummy-variables.
The model (2.1) can be written in matrix notation as

$$y = X\beta + u, \tag{2.2}$$

where $y = (y_1, y_2, \cdots, y_T)'$ is a $(T \times 1)$ vector of observations on the dependent variables, $X = (X_1, X_2, \cdots, X_s)'$ is a $(T \times s)$ nonstochastic regressor in which

$$X = \begin{bmatrix} X^* \\ X^* \\ \vdots \\ X^* \end{bmatrix} \quad \text{with} \quad X^* = \begin{bmatrix} 1 & 0 & 0 & \cdot & \cdots & \cdot & 0 & 0 \\ 1 & 1 & 0 & \cdot & \cdots & \cdot & 0 & 0 \\ 1 & 0 & 1 & \cdot & \cdots & \cdot & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & \cdot & \cdots & \cdot & 1 & 0 \\ 1 & 0 & 0 & \cdot & \cdots & \cdot & 0 & 1 \end{bmatrix},$$

$\beta = (\beta_1, \beta_2, \cdots, \beta_s)'$ is a $(s \times 1)$ vector of unknown regression coefficients to be estimated and $u = (u_1, u_2, \cdots, u_T)'$ is a $(T \times 1)$ vector of disturbances. The disturbances are assumed to follow the s-th order autoregressive process, AR(s)-processes, as

$$u_t = \rho u_{t-s} + \varepsilon_t, \tag{2.3}$$

where s denotes the ''seasons'' per year, $|\rho| < 1$, $E(\varepsilon_t) = 0$, $Var(\varepsilon) = \sigma_\varepsilon^2$ and $E(\varepsilon_t, \varepsilon_s) = 0$ for $s \neq t$, $s, t = 1, 2, \cdots, T$. For observations of m years, $T = ms$. The autocovariance $E(u_t, u_{t-j})$ is zero unless j is an integer multiple of s, in which case $E(u_t, u_{t-j}) = \sigma_u^2 \rho^{j/s}$, where σ_u^2 is $\sigma_\varepsilon^2/(1-\rho^2)$. Thus we have the $(T \times T)$ covariance matrix of the disturbances

$$Cov(u) = E(uu') = \sigma_u^2 V_s = \sigma_u^2 (V_1 \otimes I_s), \tag{2.4}$$

where
$$V_1 = \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{m-1} \\ \rho & 1 & \rho & \cdots & \rho^{m-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{m-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{m-1} & \rho^{m-2} & \rho^{m-3} & \cdots & 1 \end{bmatrix}.$$

where I_s is the $(s \times s)$ identity matrix and $\otimes$ denotes the Kronecker product. It can be shown that the inverse of V_s is

$$V_s^{-1} = (V_1^{-1} \otimes I_s) \tag{2.5}$$

with
$$V_1^{-1} = \frac{1}{1-\rho^2}\begin{bmatrix} 1 & -\rho & 0 & \cdots & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & \cdots & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & \cdots & -\rho & 1 \end{bmatrix}.$$

With this covariance matrix, COTE for β in model (2.2) can be obtained as (cf. Fomby et. al (1984), pp. 208-209)

$$\widehat{\beta_{co}} = (X'R'RX)^{-1}X'R'Ry, \tag{2.6}$$

where $R = (R_1 \otimes I_s)$, $R_1 = \begin{bmatrix} -\rho & 1 & 0 & \cdot & \cdots & 0 & 0 \\ 0 & -\rho & 1 & \cdot & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdot & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdot & \cdots & -\rho & 1 \end{bmatrix}.$

Note that (2.6) is an altenative to GLSE, in particular, is as good as FGLSE (feasible GLSE) when the sample is large and ρ is unknown.

The covariance of COTE given in (2.6) is

$$\begin{aligned} Cov(\widehat{\beta_{co}}) &= \sigma_u^2(X'R'RX)^{-1}X'R'RV_sR'RX(X'R'RX)^{-1} \\ &= \sigma_u^2(1-\rho^2)(X'R'RX)^{-1}. \end{aligned} \tag{2.7}$$

3. Efficiency of OLSE relative to COTE

In the following we are interested in the efficiency of OLSE relative to the COTE for β. There are several approaches to define relative efficiency of OLSE to GLSE in Iwasaki (1988). Since for any unbiased (vector-valued) estimator the mean square error (MSE) is equal to the trace of its covariance matrix, we compare the MSE of both estimators using the following definition by Kraemer (1980):

Definition: The relative efficiency function is defined by

$$e(\rho, m) := \frac{tr[Cov(\hat{\beta})]}{tr[Cov(\widehat{\beta_{co}})]},$$

where $Cov(\widehat{\beta_{co}}) = \sigma_u^2(1-\rho^2)(X'R'RX)^{-1}$ and $Cov(\hat{\beta}) = \sigma_u^2(X'X)^{-1}X'V_sX(X'X)^{-1}$.

This efficiency clearly depends on the correlation coefficient ρ. Since both estimator are unbiased, $e(\rho, m)$ can also be expressed as

$$e(\rho, m) := \frac{MSE(\hat{\beta})}{MSE(\widehat{\beta}_{co})}, \tag{3.1}$$

where MSE of an estimator for β is $MSE(\hat{\beta}) = E((\hat{\beta}-\beta)'(\hat{\beta}-\beta))$.

The following result establishes the efficiency function of OLSE relative to COTE:

Theorem 3.1: The efficiency function of COTE relative to OLSE is for model (2.1)

$$e(\rho, m) = \frac{(m-1)[m(1-\rho^2)-2\rho(1-\rho^m)]}{m^2(1-\rho^2)}. \tag{3.2}$$

Proof: First, we consider $Cov(\hat{\beta}) = \sigma_u^2(X'X)^{-1}X'V_sX(XX)^{-1}$.

$$\text{Since} \quad (X'X) = m\begin{bmatrix} s & 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & 0 & 0 & \cdots & 1 & 0 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} := mF, \tag{3.3}$$

$$\text{so} \quad (X'X)^{-1} = \frac{1}{m}\begin{bmatrix} 1 & -1 & -1 & \cdots & -1 & -1 \\ -1 & 2 & 1 & \cdots & 1 & 1 \\ -1 & 1 & 2 & \cdots & 1 & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -1 & 1 & 1 & \cdots & 2 & 1 \\ -1 & 1 & 1 & \cdots & 1 & 2 \end{bmatrix} := \frac{1}{m}Q. \tag{3.4}$$

Further, after the some calculation we have

$$X'V_sX = \frac{[m(1-\rho^2)-2\rho(1-\rho^m)]}{(1-\rho)^2}F. \tag{3.5}$$

Therefore, we have by (3.3), (3.4) and (3.5)

$$\begin{aligned} Cov(\hat{\beta}) &= \sigma_u^2(X'X)^{-1}X'V_sX(XX)^{-1} \\ &= \frac{\sigma_u^2[m(1-\rho^2)-2\rho(1-\rho^m)]}{m^2(1-\rho)^2}Q. \end{aligned} \tag{3.6}$$

Next, consider that $Cov(\widehat{\beta_{co}}) = \sigma_u^2(1-\rho^2)(X'R'RX)^{-1}$ as given in (2.7).

$$\text{Since} \quad RX = \begin{bmatrix} X^{\Delta} \\ X^{\Delta} \\ \vdots \\ X^{\Delta} \end{bmatrix} \quad \text{with} \quad X^{\Delta} = \begin{bmatrix} 1-\rho & 0 & 0 & 0 & \cdots & 0 & 0 \\ 1-\rho & 1-\rho & 0 & 0 & \cdots & 0 & 0 \\ 1-\rho & 0 & 1-\rho & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1-\rho & 0 & 0 & 0 & \cdots & 1-\rho & 0 \\ 1-\rho & 0 & 0 & 0 & \cdots & 0 & 1-\rho \end{bmatrix},$$

so $\quad (RX)'(RX) = (m-1)(1-\rho)^2 F.$

Therefore, we have

$$Cov(\widehat{\beta_{co}}) = \sigma_u^2(1-\rho^2)(1-\rho)^2(X'R'RX)^{-1}$$
$$= \frac{\sigma_u^2(1-\rho^2)}{(m-1)(1-\rho)^2} Q. \tag{3.7}$$

The relative efficiency function (3.2) is immediately proved by substituting (3.6) and (3.7) into (3.1).

From the efficiency function defined in (3.2), MSE's of OLSE and COTE are same regardless ρ for large sample since $\lim_{m\to\infty} e(\rho,m) = 1, \ -1 < \rho < 1$. We now turn to the efficiency function $e(\rho,m)$ whose behavior obviousely depends on ρ and m. The next results provide the conditions that OLSE is better than COTE in the values of ρ and m.

Corollary 3.2: For the efficiency function $e(\rho,m)$ given in (3.2),

(i) for nonnegative ρ, $e(\rho,m)$ is a strictly decreasing function in ρ with

$$\max_{0\le\rho<1} e(\rho,m) = e(0,m) = \frac{m-1}{m}, \quad \text{for all } m.$$

(ii) for any ρ,

$$\text{if } m \ge -\frac{2\rho(1-\rho^m)}{1-\rho^2}, \text{ then } e(\rho,m) \le 1 \text{ is.}$$

Proof: From (3.2), (i) and (ii) follow immediately by showing that the first derivative with respective to ρ is negative and that $e(\rho,m) \le 1$ is equivalent to the condition (ii).

By Corollary, we can say that in a seasonal regression model with AR(s) type disturbance as model (2.2), OLSE for β is always better than COTE and moreover, OLSE becomes superior to COTE as ρ goes to 1 regardless m where $\rho \ge 0$.

References

[1] Doran, H.E. (1981). Omission of an observation from a regression analysis, *Journal of Econometrics*, Vol. 16, 367-374.

[2] Fomby, J.B., Hill, R.C. and Johnson, S.R. (1984). *Advanced Econometrics Methods*, Springer Verlag, New York, Berlin.

[3] Hoque, A. (1989). The efficiency of the cochrane-orcutt procedure, *Australian Journal of Statistics*, Vol. 31, 385-392.

[4] Iwasaki, M. (1988). Efficiency of least squares in a linear model with autocorrelated disturbances, *Statistical Theory and Data Analysis II*, K. Matusita (Hrsg.) North Holland, 511-523.

[5] Judge, G.G., Griffiths, W.E., Hill, R.C., Luekepohl, H. and Lee, T.C. (1985). *The Theory and Practice of Econometrices*, Wiley, New York.

[6] Kadiyala, K.R. (1968). A transformation used to circumvent the problem of autocorrelation, *Econometrica*, Vol. 36, 93-96.

[7] Kraemer, W. (1980). Finite sample efficiency of ordinary least squares in the linear regression model with autocorrelated errors, *Journal of American Statistical Association*, Vol. 62, 1018-1021.

[8] Kraemer, W. (1982). Note on estimating linear trend when residuals are autocorrelated, *Econometrica*, Vol. 50, 1065-1067.

[9] Maeshiro, A. (1976). Autoregressive transformation, trended independent variables and autocorrelated disturbance term, *Review of Economics and Statistics*, Vol. 58, 497-500.

[10] Maeshiro, A. (1979). On the retention of first observation in serial correlation adjustment of regression models, *International Economic Review*, Vol. 20, 259-265.

[11] Maeshiro, A. (1980). Autocorrelation and trended explanatory variables: a reply, *Review of Economics and Statistics*, Vol. 62, 487-489.

[12] Park, R.E. and Mitchell, B.M. (1980). Estimating the autocorrelated error model with trended data, *Journal of Econometrics*, Vol. 13, 185-201.

[13] Poirier, D.J. (1978). The effect of the first observation in regression models with first-order autoregressive disturbance, *Applied Statistics*, Vol. 27, 67-68.

[14] Puterman, M.L. (1988). Leverage and influence in autocorrelated regressionmodels, *Applied Statistics*, Vol. 37, 76-86.

[15] Stemann, D. and Trenkler, G. (1993). Leverage and Cochrane Orcutt estimation in the linear regression, *Communications in Statistics -Theory and Method*, Vol. 22, 1315-1333.