The Korean Communications
in Statistics Vol. 6, No. 3, 1999
pp. 667-676

The Effect of First Observation in Panel Regression Model with Serially Correlated Error Components[1)]

Seuck Heun Song[2)]

Abstract

We investigate the effects of omission of initial observations in each individuals in the panel data regression model when the disturbances follow a serially correlated one way error components. We show that the first transformed observation can have a relative large hat matrix diagonal component and a large influence on parameter estimates when the correlation coefficient is large in absolute value.

1. Introduction

In the usual linear regression model with first-order autoregressive(AR(1)) disturbances, the importance of the initial observation has been well documented. See Kadiyala(1968), Poirier(1978), Maeshiro(1976, 1979), Doran(1981) and Kraemer(1982). Recently, Puterman(1988) considered simple regression models with AR(1) disturbances and sensitivity of leverage with respect to a varying correlation coefficient. It is shown there that in case of one regressor the first transformed observation frequently has a large hat matrix diagonal and consequently its deletion might have a major impact on the parameter estimates.

In this paper we extend the results of Puterman(1988) to the general linear regression model. In the following we consider the panel regression model with serially correlated one way error components. To investigate the effects of omission of initial observations in each individuals, we will demonstrate that presence or absence of a constant term is highly influential on the first leverage when the correlation coefficient of the disturbances is large in absolute value.

2. The Model

We consider the following panel data regression model:

$$y_{it} = \sum_{j=1}^{k}\beta_j x_{jit} + u_{it}, \qquad i=1,2,\cdots,N, \quad t=1,2,\cdots,T, \tag{2.1}$$

1) This research was supported by Duksung Women's University Fund 1999.

2) Associate Professor, Department of Statistics, Duksung Women's University, 419 Ssangmun Dong Tobong Ku, Seoul, 132-714, Korea.

where y_{it} is an observation on a dependent variable for the ith cross sectional unit (firms, individuals) for the tth time period, x_{jit} is an observation on the jth nonstochastic regressor for the ith cross sectional unit and tth time period. The model (2.1) can be written in matrix notation as

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

where y is an $NT\times 1$ observation vector, X is an $NT\times k$ regressor matrix, β is a $k\times 1$ vector of regression coefficients to be estimated, and u is an $NT\times 1$ disturbance vector. Both N and T are assumed to be larger than k. A popular specification of the disturbances is the error components model, see Hsiao(1986). This paper focuses on an one-way serially correlated error component model:

$$u_{it} = \mu_i + \nu_{it}, \qquad i = 1, 2, \cdots, N, \quad t = 1, 2, \cdots, T, \tag{2.3}$$

where the μ_i are the unobservable individual specific effects which are assumed to be $i.i.d.\ (0, \sigma_\mu^2)$. The ν_{it} are the remainder disturbances which are also assumed to be generated by AR(1) process (see, Lillard and Willis(1978) and Lillard and Weiss(1979)):

$$\nu_{it} = \rho\nu_{it-1} + \varepsilon_{it}, \quad |\rho| < 1, \quad i = 1, 2, \cdots, N, \quad t = 1, 2, \cdots, T, \tag{2.4}$$

where the ε_{it} are $i.i.d.\,(0, \sigma_\varepsilon^2)$ and $Var(\nu_{it}) = \sigma_\nu^2 = \sigma_\varepsilon^2/(1-\rho^2)$ and σ_ε^2 is held constant in what follows. The μ_i's and the ν_{it}'s are independent of each other. Under these assumptions, the $NT\times NT$ covariance matrix of disturbances can be written as

$$E(uu') = \Omega = I_N \otimes \Sigma = I_N \otimes (\sigma_\mu^2 \iota_T \iota_T' + \sigma_\nu^2 V), \tag{2.5}$$

where I_N is an $N\times N$ identity matrix, ι_T is a $T\times 1$ vector of ones, $\otimes$ denotes the Kronecker Product and V is the $AR(1)$ correlation matrix of order T:

$$V = \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{T-1} \\ \rho & 1 & \rho & \cdots & \rho^{T-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \cdots & 1 \end{bmatrix}. \tag{2.6}$$

In this model, when Σ is known, the best linear unbiased estimator of β is given by generalized least squares(GLS) estimator:

$$\tilde{\beta} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y \ . \tag{2.7}$$

The GLS-estimator for β can be also obtained by applying ordinary least squares(OLS) to the transformed observation matrix $[Py, PX]$ (see Baltagi and Li(1991) and Baltagi(1993)). That is,

$$Py = PX\beta + Pu, \tag{2.8}$$

where $P = I_N \otimes [(I_T - \theta \overline{J_T}^{a})R_T]$, (2.9)

$$\text{with } R_T = \begin{bmatrix} (1-\rho^2)^{1/2} & 0 & \cdots & 0 & 0 \\ -\rho & 1 & \cdots & 0 & 0 \\ 0 & -\rho & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & -\rho & 1 \end{bmatrix}, \ \overline{J_T}^{a} = \frac{\iota_T^{a}\iota_T^{a\prime}}{d^2}, \ \iota_T^{a\prime} = (a, \ \iota_{T-1}'), \tag{2.10}$$

$$a = \sqrt{\frac{1+\rho}{1-\rho}}, \quad \theta = 1 - \left\{\frac{\sigma_\varepsilon^2}{d^2(1-\rho)^2\sigma_\mu^2 + \sigma_\varepsilon^2}\right\}^{1/2} \text{ and } d^2 = \frac{1+\rho}{1-\rho} + T - 1 . \tag{2.11}$$

The error terms of the transformed model in (2.8) are $i.i.d.$ and thus OLS applied to (2.8) produces the GLSE of β

$$\tilde{\beta} = [(PX)'(PX)]^{-1}(PX)'(Py). \tag{2.12}$$

If we wish to detect influential observations the diagonal entries of the so-called hat matrix play an important role for each individual i. In the context of our GLSE (2.12) it is given by

$$H_i^* = X_i^*(X_i^{*\prime}X_i^*)^{-1}X_i^{*\prime} = P_iX_i(X_i'P_i'P_iX_i)^{-1}X_i'P_i', \tag{2.13}$$

where H_i^* denotes the hat matrix for each individual i and X_i^* and X_i denote also submatrix of X^* and X for each individual i, respectively, and $P_i = (I_T - \theta \overline{J_T}^{a})R_T$. It is well known that the diagonal elements of H_i^* satisfy $0 \le h_{ij}^* \le 1$. According to Belsley et al. (1980, pp. 17) we call y_{ij}^* a high leverage point if $h_{ij}^* \ge 2k/T$. In the following we shall concentrate on the first element of H_i^* which is

$$h_{i1}^* = x_{i1}^{*\prime}(X_i^{*\prime}X_i^*)^{-1}x_{i1}^* = (1-\rho^2)(1-\frac{T\theta}{d^2})^2 x_{i1}'(X_i'P_i'P_iX_i)^{-1}x_{i1}, \tag{2.14}$$

where $x^*_{i1}{}'$ denotes the first row of X^*_i, and $x_{i1}{}'$ denotes the first row of X_i. It will be seen that in some cases h^*_{i1} will be high value, dependent only on ρ and thus it is questionable to delete the corresponding observation from the data set as a high leverage point.

3. Main Results

Puterman(1988) considered simple regression models with AR(1)-disturbances and sensitivity of h^*_{i1} with respect to a varying correlation coefficient. In the following, we will extend this approach to the panel regression model with serially correlated error components. In Theorem 1 and Theorem 2 we distinguish between the situations where a constant term is included and where it is excluded from the model. The importance of the constant term for usual regression models with AR(1)-disturbances has been mentioned before by several authors including Berenblut and Webb(1973), Kraemer(1980, 1982) and Bhargava(1989). In the presence of a constant term in (2.1), we will show that the leverage h^*_{i1} increases to 1 as $\rho \to 1$. If there is no constant term in the model the first leverage value h^*_{i1} decreases to 0 as $\rho \to 1$. As a result, we have the following theorems:

Lemma: Let $C(X_i)$ denote the range (column space) of a matrix.

If $\dot{\iota}_T \notin C(X_i)$, then $Q_i = X_i' D' D X_i$ is nonsingular

Proof: Suppose Q_i is singular. Then there exists a nonzero vector v such that $Q_i v = 0$ which implies $D'DX_i v = 0$. Since the null space of $D'D$ is generated by the vector $\dot{\iota}_T$ we see that $Xv = \lambda_0 \dot{\iota}_T$ for some scalar λ_0. Furthermore, as $v \neq 0$ and X is of full column rank we have $\lambda_0 \neq 0$. This entails $\dot{\iota}_T \in C(X_i)$ which by assumption is excluded.

Theorem 1: Let $C(X_i)$ denotes the range (column space) of a matrix.

1) If $\dot{\iota}_T \notin C(X_i)$, then $\lim_{\rho \to 1} h^*_{i1} = 0$.

2) If $\dot{\iota}_T$ is a column of X_i, then $\lim_{\rho \to 1} h^*_{i1} = 1$.

Proof: 1. For the matrix R_T from (2.10) we have

$$\lim_{\rho\to 1} R_T = \begin{bmatrix} 0 & \cdots & 0 \\ & D & \end{bmatrix}, \tag{3.1}$$

where the $(T-1)\times T$ matrix D is given by

$$D = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{bmatrix}. \tag{3.2}$$

Further it can be easily shown that

$$\lim_{\rho\to 1} \theta = \lim_{\rho\to 1} \left\{ 1 - \left[\frac{\sigma_\varepsilon^2}{d^2(1-\rho)^2\sigma_\mu^2 + \sigma_\varepsilon^2} \right]^{1/2} \right\} = 0. \tag{3.3}$$

Making use of

$$\lim_{\rho\to 1} \frac{\theta}{d^2} = 0, \quad \lim_{\rho\to 1} \frac{\theta\alpha}{d^2} = 0, \quad \lim_{\rho\to 1} \frac{\theta\alpha^2}{d^2} = 0, \tag{3.4}$$

we obtain

$$\lim_{\rho\to 1} \theta \overline{J_T^\alpha} = 0. \tag{3.5}$$

Hence we obtain

$$\begin{aligned} \lim_{\rho\to 1} (X_i^{*\prime} X_i^*) &= \lim_{\rho\to 1} (X_i' P_i' P_i X_i) \\ &= \lim_{\rho\to 1} X_i' [(I_T - \theta \overline{J_T}^\alpha) R_T]' [(I_T - \theta \overline{J_T}^\alpha) R_T] X_i \\ &= X_i' D' D X_i, \end{aligned} \tag{3.6}$$

where

$$D'D = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & -1 \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{bmatrix}. \tag{3.7}$$

According to the Lemma the matrix $Q_i = X_i' D' D X_i$ is nonsingular if $\iota_T \notin C(X_i)$. Consequently we get

$$\begin{aligned} \lim_{\rho\to 1} h_{i1}^* &= \lim_{\rho\to 1} x_{i1}^{*\prime} (X_i^{*\prime} X_i^*)^{-1} x_{i1}^* \\ &= \lim_{\rho\to 1} (1-\rho^2)(1 - \frac{T\theta}{d^2})^2 x_{i1}' (X_i' P_i' P_i X_i)^{-1} x_{i1} = 0. \end{aligned} \tag{3.8}$$

2. When the first column of X_i consists of ones, $X_i = [\, \iota_T, X_{i1} \,]$, where X_{i1} is a submatrix of size $T\times(k-1)$. Let us partitioned as

$$(X_i'P_i'P_iX_i)^{-1} = \begin{bmatrix} \dot{\iota}_T'P_i'P_i\dot{\iota}_T & \dot{\iota}_T'P_i'P_iX_{i1} \\ X_{i1}'P_i'P_i\dot{\iota}_T & X_{i1}'P_i'P_iX_{i1} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ A_{12}' & A_{22} \end{bmatrix}. \tag{3.9}$$

By a well known inversion formula of partitioned matrices (Dhrymes(1995), p. 485) we derive

$$A_{11} = \psi(1 + \psi\dot{\iota}_T'P_i'P_iX_{i1}A_{22}X_{i1}'P_i'P_i\dot{\iota}_T), \tag{3.10}$$

$$A_{12} = -\psi\,\dot{\iota}_T'P_i'P_iX_{i1}A_{22}, \tag{3.11}$$

$$A_{22} = (X_{i1}'P_i'P_iX_{i1} - \psi X_{i1}'P_i'P_i\dot{\iota}_T\dot{\iota}_T'P_i'P_iX_{i1})^{-1}, \tag{3.12}$$

where $\psi = 1/\dot{\iota}_T'P_i'P_i\dot{\iota}_T$.

Let us write the first row of X_i in the form $x_{i1}' = (1, \bar{x}_{i1}')$, where $\bar{x}_{i1}$ is a column vector with $k-1$ entries. Then we have

$$h_{i1}^* = (1-\rho^2)(1-\frac{T\theta}{d^2})^2 x_{i1}'(X_i'P_i'P_iX_i)^{-1}x_{i1}$$

$$= (1-\rho^2)(1-\frac{T\theta}{d^2})^2(A_{11} + 2A_{12}\bar{x}_{i1} + \bar{x}_{i1}'A_{22}\bar{x}_{i1}). \tag{3.13}$$

Our first task is to show

$$\lim_{\rho\to 1}(1-\rho^2)(1-\frac{T\theta}{d^2})^2\,\bar{x}_{i1}'A_{22}\bar{x}_{i1} = 0. \tag{3.14}$$

In (3.12) we observe that

$$P_i\dot{\iota}_T\dot{\iota}_T'P_i' = (I_T - \theta\overline{J_T^a})R_T\dot{\iota}_T\dot{\iota}_T'R_T'(I_T - \theta\overline{J_T^a})'$$

$$= R_T\dot{\iota}_T\dot{\iota}_T'R_T' - 2\theta\overline{J_T^a}'R_T\dot{\iota}_T\dot{\iota}_T'R_T' + \theta^2\overline{J_T^a}R_T\dot{\iota}_T\dot{\iota}_T'R_T'\overline{J_T^a}'.$$

Using $\lim_{\rho\to 1}\theta\overline{J_T^a} = 0$,

$$R_T\dot{\iota}_T\dot{\iota}_T'R_T' = (1-\rho)\begin{bmatrix} 1+\rho & (1-\rho^2)^{1/2} & \cdots & (1-\rho^2)^{1/2} \\ (1-\rho^2)^{1/2} & 1-\rho & \cdots & 1-\rho \\ \vdots & \vdots & \ddots & \vdots \\ (1-\rho^2)^{1/2} & 1-\rho & \cdots & 1-\rho \end{bmatrix} \tag{3.15}$$

and

$$\dot{\iota}_T'R_i'R_i\dot{\iota}_T = 1-\rho^2+(T-1)(1-\rho)^2, \tag{3.16}$$

we can obtain

$$\lim_{\rho\to 1} \psi\, P_i \iota_T \iota_T' P_i' = e_1 e_1', \tag{3.17}$$

where $e_1 = (1, 0, \cdots, 0)'$, and

$$\lim_{\rho\to 1} P_i X_{i1} = \lim_{\rho\to 1} (I_T - \theta\, \overline{J_T^{\alpha}}) R_T X_{i1} = \begin{bmatrix} 0 & \cdots & 0 \\ & D & \end{bmatrix} X_{i1}. \tag{3.18}$$

From (3.17) and (3.18) it follows that

$$\begin{aligned}
\lim_{\rho\to 1} \psi X_{i1}' P_i' P_i \iota_T \iota_T' P_i' P_i X_{i1} &= \lim_{\rho\to 1} (P_i X_{i1})' \psi P_i \iota_T \iota_T' P_i' (P_i X_{i1}) \\
&= X_{i1}' \begin{bmatrix} 0 & \\ \vdots & D' \\ 0 & \end{bmatrix} e_1 e_1' \begin{bmatrix} 0 & \cdots & 0 \\ & D & \end{bmatrix} X_{i1} \\
&= X_{i1}' \begin{bmatrix} 0 & \\ \vdots & D' \\ 0 & \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} 0 & \cdots & 0 \\ & D & \end{bmatrix} X_{i1} = 0
\end{aligned} \tag{3.19}$$

which gives

$$\lim_{\rho\to 1} A_{22}^{-1} = \lim_{\rho\to 1} X_{i1}' P_i' P_i X_{i1} = X_{i1}' \begin{bmatrix} 0 & \\ \vdots & D' \\ 0 & \end{bmatrix} \begin{bmatrix} 0 & \cdots & 0 \\ & D & \end{bmatrix} X_{i1} = X_{i1}' D' D X_{i1} = Q_1, \tag{3.20}$$

where the matrix $Q_1 = X_{i1}' D' D X_{i1}$ is positive definite by the fact that $\iota_T \notin C(X_i)$ (otherwise $X_i = (\iota_T, X_{i1})$ would be singular). Then we obtain

$$\lim_{\rho\to 1} (1-\rho^2)\, \bar{x}_{i1}' A_{22}\, \bar{x}_{i1} = 0. \tag{3.21}$$

Let us now consider the limit corresponding to the second expression in (3.13). Making use of (3.1), (3.7) and (3.20) we obtain

$$\begin{aligned}
&\lim_{\rho\to 1} -(1-\rho^2)\left(1-\frac{T\theta}{d^2}\right)^2 \psi \iota_T' P_i' P_i X_{i1} A_{22}\, \bar{x}_{i1} \\
&= -\lim_{\rho\to 1} (1-\rho^2)\left(1-\frac{T\theta}{d^2}\right)^2 \psi \iota_T' (I_T - \theta \overline{J_T^{\alpha}})' R_T' R_T (I_T - \theta \overline{J_T^{\alpha}}) X_{i1} A_{22}\, \bar{x}_{i1} \\
&= -\lim_{\rho\to 1} \iota_T' (I_T - \theta \overline{J_T^{\alpha}})' R_T' R_T (I_T - \theta \overline{J_T^{\alpha}}) X_{i1} A_{22}\, \bar{x}_{i1} \\
&= -\iota_T' D' D X_{i1} Q_1^{-1}\, \bar{x}_{i1} = 0.
\end{aligned} \tag{3.22}$$

The latter identify follows from $\dot{\iota}_T' D' D = 0$. To find out the first limit in (3.13) observe that

$$\lim_{\rho\to 1}(1-\rho^2)(1-\frac{T\theta}{d^2})^2\psi = 1. \tag{3.23}$$

$$\lim_{\rho\to 1}\psi^{1/2}P_i\dot{\iota}_T = \lim_{\rho\to 1}\psi^{1/2}R_T(I_T-\theta\overline{J_T})\dot{\iota}_T = e_1. \tag{3.24}$$

Hence we have

$$\lim_{\rho\to 1}(1-\rho^2)(1-\frac{T\theta}{d^2})^2 A_{11} = 1. \tag{3.25}$$

We will still investigate the case $\rho\to -1$. For this the vector $\dot{\iota}_T^* = (1,-1,\cdots,(-1)^{T-1})'$ plays a crucial role.

Theorem 2:

1) If $\dot{\iota}_T^* \notin C(X_i)$, then $\lim_{\rho\to -1} h_{i1}^* = 0$.

2) If $\dot{\iota}_T^*$ is a column of X_i, then $\lim_{\rho\to -1} h_{i1}^* = 1$.

Proof: 1. Consider again the matrix R_T from (2.9)

$$\lim_{\rho\to -1} R_T = \begin{bmatrix} 0 & \cdots & 0 \\ & D^* & \end{bmatrix}, \tag{3.26}$$

where the $(T-1)\times T$ matrix D^* is given by

$$D^* = \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & 0 & \cdots & 1 & 1 \end{bmatrix} \tag{3.27}$$

and

$$\lim_{\rho\to -1}\theta = \lim_{\rho\to -1}\left\{1-\left[\frac{\sigma_\varepsilon^2}{d^2(1-\rho)^2\sigma_\mu^2+\sigma_\varepsilon^2}\right]^{1/2}\right\} \neq 0. \tag{3.28}$$

Thus

$$\begin{aligned}\lim_{\rho\to -1} X_i' P_i' P_i X_i &= \lim_{\rho\to -1} X_i' [(I_T-\theta\overline{J_T}^a)R_T]'[(I_T-\theta\overline{J_T}^a)R_T]X_i \\ &= X_i' D^{**'} D^{**} X_i, \end{aligned}\tag{3.29}$$

where $D^{**} = (I_T-\theta\overline{J_T}^a)D^*$ and analogous to our Lemma, since $\dot{\iota}_T^* \notin C(X_i)$, the matrix $Q_i^* = X_i' D^{**'} D^{**} X_i$ is nonsingular. Then we find

$$\lim_{\rho\to -1} h_{i1}^* = \lim_{\rho\to -1} (1-\rho^2)(1-\frac{T\theta}{d^2})^2 x_{i1}'(X_i'P_i'P_iX_i)^{-1}x_{i1} = 0 \tag{3.30}$$

2. Consider $X_i = (\dot{\iota}_T^*, X_{i1})$, and compute the inverse of $X_i'P_i'P_iX_i$ partitioned as:

$$(X_i'P_i'P_iX_i)^{-1} = \begin{bmatrix} \dot{\iota}_T^{*\prime}P_i'P_i\dot{\iota}_T^* & \dot{\iota}_T^{*\prime}P_i'P_iX_{i1} \\ X_{i1}'P_i'P_i\dot{\iota}_T^* & X_{i1}'P_i'P_iX_{i1} \end{bmatrix}^{-1} = \begin{bmatrix} B_{11} & B_{12} \\ B_{12}' & B_{22} \end{bmatrix}, \tag{3.31}$$

where
$$B_{11} = \tilde{\psi}(1+\tilde{\psi}\,\dot{\iota}_T^{*\prime}P_i'P_iX_{i1}B_{22}X_{i1}'P_i'P_i\dot{\iota}_T^*), \tag{3.32}$$
$$B_{22} = -\tilde{\psi}\,\dot{\iota}_T^{*\prime}P_i'P_iX_{i1}B_{22}, \tag{3.33}$$
$$B_{22} = (X_{i1}'P_i'P_iX_{i1} - \tilde{\psi}X_{i1}'P_i'P_i\dot{\iota}_T^*\dot{\iota}_T^{*\prime}P_i'P_iX_{i1})^{-1}, \tag{3.34}$$

where $\tilde{\psi} = (1/\dot{\iota}_T^{*\prime}P_i'P_i\dot{\iota}_T^*)$.

It suffices to show

$$\lim_{\rho\to -1}(1-\rho^2)(1-\frac{T\theta}{d^2})^2 B_{11} = 1, \tag{3.35}$$
$$\lim_{\rho\to -1}(1-\rho^2)(1-\frac{T\theta}{d^2})^2 B_{12}\bar{x}_{i1} = 0, \tag{3.36}$$
$$\lim_{\rho\to -1}(1-\rho^2)(1-\frac{T\theta}{d^2})^2 \bar{x}_{i1}'B_{22}\bar{x}_{i1} = 0. \tag{3.37}$$

However, these identities follow along the same lines as in the second part of Theorem 1, and their proof can therefore be omitted.

4. Conclusions

In this paper, it was shown that presence or absence of a constant term is highly influential on the first leverage when the correlation coefficient is large in absolute value. In presence of a constant term we have shown that the leverage h_{i1}^* increases to 1 as $|\rho| \to 1$. Thus in models with an intercept term and as $|\rho| \to 1$, the first transformed observation has a large hat matrix diagonal, and consequently its deletion can have a large impact on parameter estimates. If there is no constant term in the model the first leverage value h_{i1}^* decreases to 0 as $|\rho| \to 1$. Hence the effect of the first transformed observation becomes arbitrary small when ρ approaches 1.

References

[1] Baltagi, B.H. (1993). Useful matrix transformations for panel data analysis: a survey, *Statistical Papers,* Vol. 34, 281-301.

[2] Baltagi, B.H. and Li, Q. (1991). A transformation that will circumvent the problem of autocorrelation in an error component model, *Journal of Econometics,* Vol. 48, 385-393.

[3] Belsley, D.A., Kuh, E. and Welsch, R.E. (1980). *Regression Diagnostics: Identifying Influential Data and Source of Collinearity*, Wiley, New York.

[4] Berenblut, I.I. and Webb, G.I. (1973). A new test for autocorrelated errors in the linear regression model, *Journal of the Royal Statistical Society,* B, Vol. 35, 33-50.

[5] Bhargava, A. (1989). Missing observations and the use of the Durbin-Watson statistic, *Biometika,* Vol. 76, 823-831.

[6] Dhrymes, P.J. (1995). *Introductory Econometrics*, Springer Verlag, Heidelberg.

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

[8] Hsiao C (1986). *Analysis of Panel Data,* Cambridge University Press, Cambridge.

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

[10] 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.

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

[12] Lillard, L.A. and Willis, R.J. (1978). Dynamic aspects of earning mobility, *Econometrica,* Vol. 46, 985-1012.

[13] Lillard, L.A. and Weiss, Y. (1979). Components of variation in panel earnings data: american scientists 1960-1970, *Econometrica*, Vol. 47, 437-454.

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

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

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

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