

On the Estimation of a Linear Time Trend Regression with a One-Way Error Component Model in the Presence of Serially Correlated Errors

Chihwa Kao* and Jamie Emerson
Syracuse University

First Draft: May 25, 1998

This Draft: July 2, 1998

Abstract

In this paper, we study the limiting distributions for the ordinary least squares (OLS), the fixed effects (FE), first difference (FD), and the generalized least squares (GLS) estimators in a linear time trend regression with a one-way error component model in the presence of serially correlated errors. We show that when the error term is $I(0)$, the FE is asymptotically equivalent to GLS. However, when the error term is $I(1)$, the GLS could be less efficient than FD or FE estimators and FD is the most efficient estimator. However, when the intercept is included in the model and the error term is $I(0)$, the OLS, FE, and GLS are asymptotically equivalent. The limiting distribution of the GLS depends on the initial condition significantly when the error term is $I(1)$ and an intercept is included in the regression. Monte Carlo experiments are employed to compare the performance of these estimators in finite samples. The main findings are: (1) the two-steps GLS estimators perform well if the variance component, Δ , is small and close to zero when $\rho < 1$, (2) the FD estimator dominates the other estimators when $\rho = 1$ for all values of Δ , and (3) the FE estimator is recommended in practice since it performs pretty well for all values of ρ and Δ .

1 Introduction

In this paper we study the limiting distributions for the ordinary least squares (OLS), the fixed effects (FE), first difference (FD), and the generalized least squares (GLS) estimators in a linear time trend regression with a one-way error component model in the presence of serially correlated errors. There are two popular ways of estimating a regression with error components, i.e., the FE model and the random-effect model. A FE model can be estimated by OLS by conditioning on the error components, while the random-effect model is usually estimated by GLS unconditionally. One advantage of using the FE estimator is that we do not need to invert the variance-covariance matrix which could be computationally involved especially

*An electronic version of the paper in postscript format can be retrieved from <http://web.syr.edu/~cdkao>. Address correspondence to: Chihwa Kao, Center for Policy Research, 426 Eggers, Syracuse University, Syracuse, NY 13244-1020; e-mail: cdkao@maxwell.syr.edu.

when the error terms are serially correlated (e.g., Baltagi and Li, 1991). However, as we will show later in Theorems 1-4, the GLS will be the asymptotically efficient estimator when the error term is $I(0)$ and almost asymptotically efficient when the error term is $I(1)$. On the other hand, econometricians have been concerned with conditions for ordinary least squares (OLS) estimator to be asymptotically efficient, e.g., Grenander and Rosenblatt (1957), Kruskal (1968), Chipman (1979), Krämer (1982), Baltagi (1989), Phillips and Park (1988), Canjels and Watson (1997), and Vogelsang (1998).

In this paper we show that when the error term is $I(0)$, the FE is asymptotically equivalent to GLS and the OLS is less efficient than the GLS. However, when the error term is $I(1)$, GLS could be less efficient than the FE and FD, and FD is the most efficient estimator. When the error term is $I(0)$, the OLS, FE, and GLS are all asymptotically equivalent if the intercept is included in the model. On the other hand, the limiting distribution of the GLS depends on the initial condition significantly when the error term is $I(1)$ and an intercept is included in the regression.

Section 2 develops the asymptotic theory for OLS, FE, FD, and GLS estimators with $I(0)$ error term. Section 3 gives the limiting distributions of OLS, FE, FD, and GLS estimators with $I(1)$ error term. Section 4 discusses the effects of a fitted intercept. Section 5 discusses the feasible GLS estimators. Section 6 presents Monte Carlo results to evaluate the finite sample properties of the proposed estimators. Section 6 summarizes the findings. All proofs are in the Appendix.

A word on notation. We use \xrightarrow{d} to denote convergence in distribution, \xrightarrow{p} to denote convergence in probability, $[x]$ to denote the largest integer $\leq x$, and $I(0)$ and $I(1)$ to signify a time series that is integrated of order zero and one, respectively.

2 OLS, FE, FD, and GLS Estimators

Consider the following simple linear trend with one way error component model

$$y_{it} = \mu_i + \beta t + v_{it}, \tag{1}$$

$i = 1, \dots, N$, $t = 1, \dots, T$, where $\{y_{it}\}$ are 1×1 , β is the slope parameters, $\{\mu_i\}$ are the unobservable individual effects with $\mu_i \sim iid(0, \sigma_\mu^2)$, and $\{v_{it}\}$ are AR(1) stationary disturbance terms with

$$v_{it} = \rho v_{it-1} + \varepsilon_{it}, |\rho| < 1, \tag{2}$$

where $\varepsilon_{it} \sim iid(0, \sigma_\varepsilon^2)$. The μ_i are assumed to be independent of v_{it} and $v_{it} \sim (0, \sigma_v^2)$, where $\sigma_v^2 = \frac{\sigma_\varepsilon^2}{1-\rho^2}$. Let $u_{it} = \mu_i + v_{it}$. We follow Canjels and Watson (1997) to assume the following initial conditions:

Assumption 1 $v_{i1} = \sum_{j=0}^{\lfloor \kappa T \rfloor} \rho^j \varepsilon_{i1-j}$, where κ is a parameter that governs the variance of the initial condition.

Remark 1 When $\kappa = 0$, v_{i1} is $O_p(1)$. When $\kappa > 0$, v_{i1} is $O_p(1)$ when v_{it} is $I(0)$ but is $O_p(T^{1/2})$ when v_{it} is $I(1)$.

Remark 2 Model (1) can be seen as a panel regression with a non-zero drift $I(1)$ regressor.

Remark 3 Many data sets have both a large time-series and a large cross-section dimension, e.g., Summers and Heston (1991) data.

Our interest is on the estimates of the trend coefficient, β , and the estimators to be considered are the OLS estimator, the FE estimator, the FD estimator, and the GLS estimator. The OLS, $\widehat{\beta}_{OLS}$, FE, $\widehat{\beta}_{FE}$, and FD, $\widehat{\beta}_{FD}$, are:

$$\widehat{\beta}_{OLS} = \frac{\sum_{i=1}^N \sum_{t=1}^T ty_{it}}{\sum_{i=1}^N \sum_{t=1}^T t^2}, \quad (3)$$

$$\widehat{\beta}_{FE} = \frac{\sum_{i=1}^N \sum_{t=1}^T (t - \bar{t})y_{it}}{\sum_{i=1}^N \sum_{t=1}^T (t - \bar{t})^2}, \quad (4)$$

and

$$\widehat{\beta}_{FD} = \frac{\sum_{i=1}^N (y_{iT} - y_{i1})}{N(T-1)}. \quad (5)$$

where

$$\bar{t} = \frac{1}{T} \sum_{t=1}^T t = \frac{T+1}{2}.$$

Next we consider the GLS, $\widehat{\beta}_{GLS}$, (1) can be written in vector form

$$y = \beta X + u \quad (6)$$

where y is $NT \times 1$, X is a vector of $x' = (1, 2, \dots, T)$ of dimension $NT \times 1$, and u is $NT \times 1$. In order to obtain the GLS estimator we need to know the variance-covariance matrix of u , Ω . It is known that

$$\begin{aligned} \Omega &= E(uu') \\ &= I_N \otimes \Sigma, \end{aligned}$$

where

$$\Sigma = \left(\sigma_v^2 A + \sigma_\mu^2 \iota_T \iota_T' \right),$$

$$A = \begin{bmatrix} 1 & \rho & \dots & \rho^{T-1} \\ \rho & 1 & \ddots & \vdots \\ \rho^{T-1} & \dots & \dots & 1 \end{bmatrix},$$

ι_T is a $T \times 1$ ones, I_N is an identity matrix and \otimes denotes Kronecker product. Then the GLS estimator is

$$\widehat{\beta}_{GLS} = \left(X' \Omega^{-1} X \right)^{-1} X' \Omega^{-1} y. \quad (7)$$

The limiting distributions of $\widehat{\beta}_{OLS}$, $\widehat{\beta}_{FE}$, $\widehat{\beta}_{FD}$, and $\widehat{\beta}_{GLS}$ are summarized as follows. All limits in Theorems 1 – 4 are taken as $T \rightarrow \infty$ followed by $N \rightarrow \infty$ sequentially.

Theorem 1 *Let y_{it} be generated from a simple time trend model in (1) where*

$$v_{it} = \rho v_{it-1} + \varepsilon_{it},$$

with $\varepsilon_{it} \sim iid(0, \sigma_\varepsilon^2)$ and initial condition in Assumption 1. Then

- (a) $\sqrt{NT} \left(\widehat{\beta}_{OLS} - \beta \right) \xrightarrow{d} N \left(0, \frac{9}{4} \sigma_\mu^2 \right),$
- (b) $\sqrt{NT^3} \left(\widehat{\beta}_{FE} - \beta \right) \xrightarrow{d} N \left(0, \frac{12\sigma_\varepsilon^2}{(1-\rho)^2} \right),$
- (c) $\sqrt{NT} \left(\widehat{\beta}_{FD} - \beta \right) \xrightarrow{d} N \left(0, \sigma_\varepsilon^2 \frac{2-\rho^{2[\kappa T]+2}}{1-\rho^2} \right),$
- (d) $\sqrt{NT^3} \left(\widehat{\beta}_{GLS} - \beta \right) \xrightarrow{d} N \left(0, \frac{12\sigma_\varepsilon^2}{(1-\rho)^2} \right).$

Remark 4 *The results in Theorem 1 still hold if the error term is assumed to be a martingale difference sequence as in Canjels and Watson (1997).*

Remark 5 *When the error term, v_{it} , is iid, then the equivalence of the GLS and FE estimators can be shown easily. To see this, note that (e.g., Baltagi, 1995, p.16) the GLS is a weighted average of the $\widehat{\beta}_{FE}$ and the between estimator, $\widehat{\beta}_B$:*

$$\widehat{\beta}_{GLS} = W \widehat{\beta}_{FE} + (1 - W) \widehat{\beta}_B, \quad (8)$$

where W is a weight. Note that $W = 1$ and $\widehat{\beta}_B = 0$ since the time trend, t , in (1) does not vary across i . It is clear that $\widehat{\beta}_{FE}$ and $\widehat{\beta}_{GLS}$ are identical.

Remark 6 *If $[\kappa T] \rightarrow \infty$, then $\sqrt{NT} \left(\widehat{\beta}_{FD} - \beta \right) \xrightarrow{d} N \left(0, \frac{2\sigma_\varepsilon^2}{1-\rho^2} \right).$*

Remark 7 We expect that the equivalence results in Theorems 1-4 will continue to hold if we replace the time trend by the $I(1)$ regressor, though the speed of convergence of estimators will be slower. Also a fully modified version of the estimators may be needed if the regressor is correlated to the regression error with the $I(1)$ regressor. The results will be reported by the authors in different papers. A fully modified FE estimator has been studied by Kao and Chiang (1997).

Remark 8 Note that

$$\Omega^{-1}X = [I_N \otimes \Sigma^{-1}] X = \begin{bmatrix} \Sigma^{-1}x \\ \vdots \\ \Sigma^{-1}x \end{bmatrix},$$

where

$$\Sigma^{-1} = \frac{1}{\sigma_v^2} \left(A^{-1} - \frac{\sigma_\mu^2}{\sigma_v^2 + \theta\sigma_\mu^2} a_T a_T' \right),$$

$$A^{-1} = \frac{1}{1 - \rho^2} \begin{bmatrix} 1 & -\rho & 0 & 0 & 0 & \cdots & \cdots & 0 \\ -\rho & 1 + \rho^2 & -\rho & 0 & 0 & \cdots & \cdots & 0 \\ 0 & -\rho & 1 + \rho^2 & -\rho & 0 & \cdots & \cdots & 0 \\ & & & \vdots & & & & \\ 0 & 0 & 0 & 0 & \cdots & -\rho & 1 + \rho^2 & -\rho \\ 0 & 0 & 0 & 0 & \cdots & 0 & -\rho & 1 \end{bmatrix},$$

$$a_T = A^{-1} \iota_T = \frac{1}{1 + \rho} \begin{bmatrix} 1 \\ 1 - \rho \\ 1 - \rho \\ \vdots \\ 1 - \rho \\ 1 \end{bmatrix},$$

and

$$\theta = \iota_T' A^{-1} \iota_T = \frac{(1 - \rho)T + 2\rho}{1 + \rho}.$$

Then

$$\begin{aligned}
A^{-1}x &= \frac{1}{1-\rho^2} \begin{bmatrix} 1 & -\rho & 0 & 0 & 0 & \cdots & \cdots & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & 0 & \cdots & \cdots & 0 \\ 0 & -\rho & 1+\rho^2 & -\rho & 0 & \cdots & \cdots & 0 \\ & & & \vdots & & & & \\ 0 & 0 & 0 & 0 & \cdots & -\rho & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & 0 & \cdots & 0 & -\rho & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ \vdots \\ T-1 \\ T \end{bmatrix} \\
&= \frac{1}{1-\rho^2} \left\{ (1-\rho)^2 x + \begin{bmatrix} \rho^2 \\ 0 \\ 0 \\ \vdots \\ 0 \\ \rho(T-T\rho+1) \end{bmatrix} \right\},
\end{aligned}$$

and

$$a_T a'_T x = \left[\frac{1 + (1-\rho) \left(\sum_{t=2}^{T-1} t \right) + T}{(1+\rho)^2} \right] \begin{bmatrix} 1 \\ 1-\rho \\ 1-\rho \\ \vdots \\ 1-\rho \\ 1 \end{bmatrix}.$$

It follows that

$$\begin{aligned}
\Sigma^{-1}x &= \frac{1}{\sigma_v^2} \left(A^{-1} - \frac{\sigma_\mu^2}{\sigma_v^2 + \theta \sigma_\mu^2} a_T a'_T \right) x \\
&= \frac{1}{\sigma_\varepsilon^2} \left\{ (1-\rho)^2 x + \begin{bmatrix} \rho^2 \\ 0 \\ 0 \\ \vdots \\ 0 \\ \rho(T-T\rho+1) \end{bmatrix} \right\} -
\end{aligned}$$

$$\frac{1}{\sigma_\varepsilon^2} \frac{1-\rho}{1+\rho} \frac{\sigma_\mu^2}{\sigma_v^2 + \theta \sigma_\mu^2} \left[1 + (1-\rho) \left(\sum_{t=2}^{T-1} t \right) + T \right] \left\{ (1-\rho) \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{bmatrix} + \rho \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Clearly $R(\Sigma^{-1}x) \neq R(x)$ and hence $R(\Omega^{-1}X) \neq R(X)$, where R signifies the range space of a matrix. It follows by Kruskal's theorem (1968) that GLS and OLS are not equivalent even asymptotically.

Assumption 2 $\varepsilon_{it} = d(L)\varepsilon_{it}$, with $d(L) = \sum_{j=0}^{\infty} d_j L^j$, $\sum_{j=0}^{\infty} j |d_j| < \infty$ and $\varepsilon_{it} \sim iid(0, \sigma_\varepsilon^2)$.

Then we have following Corollary:

Corollary 1 Suppose ε_{it} follows Assumption 2 and let y_{it} be generated from a simple time trend model in (1) where

$$v_{it} = \rho v_{it-1} + \varepsilon_{it}$$

and initial condition in Assumption 1. Then

- (a) $\sqrt{NT} \left(\widehat{\beta}_{OLS} - \beta \right) \xrightarrow{d} N \left(0, \frac{9}{4} \sigma_\mu^2 \right)$,
- (b) $\sqrt{NT^3} \left(\widehat{\beta}_{FE} - \beta \right) \xrightarrow{d} N \left(0, \frac{d^2(1)12\sigma_\varepsilon^2}{(1-\rho)^2} \right)$,
- (c) $\sqrt{NT} \left(\widehat{\beta}_{FD} - \beta \right) \xrightarrow{d} N \left(0, d^2(1) \sigma_\varepsilon^2 \frac{2-\rho^{2[\kappa T]+2}}{1-\rho^2} \right)$,
- (d) $\sqrt{NT^3} \left(\widehat{\beta}_{GLS} - \beta \right) \xrightarrow{d} N \left(0, \frac{d^2(1)12\sigma_\varepsilon^2}{(1-\rho)^2} \right)$.

Remark 9 The GLS estimator in Corollary 1 ignores the $I(0)$ serial correlation associated with $d(L)$, due to the results in Grenander and Rosenblatt (1957).

3 Asymptotics of OLS, FE, FD, and GLS Estimators when $\rho = 1$

Model (2) is restrictive because it excludes $\rho = 1$. We investigated the asymptotic properties of the OLS, FE, FD and GLS estimators in Section 2 and found that the FE is asymptotically equivalent to the GLS estimator when v_{it} is $I(0)$. In this section, we assume $\rho = 1$ in (2), i.e., v_{it} is $I(1)$. We will show that the previous conclusions in Section 2 are substantially altered when $\rho = 1$. Note that $v_{it} = \sum_{j=0}^t \varepsilon_{ij}$ so

$$\Omega = E \left(uu' \right) = I_N \otimes \Sigma,$$

where $\Sigma = (\sigma_\varepsilon^2 A + \sigma_\mu^2 \iota_T \iota_T')$ and

$$A = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 2 & \cdots & 2 \\ 1 & 2 & 3 & \cdots & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \cdots & T \end{bmatrix}.$$

Now

$$A^{-1} = \begin{bmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \cdots & 0 \\ 0 & -1 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix},$$

$$a_T = A^{-1} \iota_T = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

$$a_T a_T' = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & & 0 & 0 \\ 0 & 0 & 0 & & 0 & 0 \\ \vdots & & & \cdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix},$$

and $\theta = \iota_T' A^{-1} \iota_T = 1$. Then

$$\Sigma^{-1} = \frac{1}{\sigma_\varepsilon^2} \left(A^{-1} - \frac{\sigma_\mu^2}{\sigma_\varepsilon^2 + \theta \sigma_\mu^2} a_T a_T' \right)$$

$$\begin{aligned}
&= \frac{1}{\sigma_\varepsilon^2} \left(\begin{bmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \cdots & 0 \\ 0 & -1 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} - \frac{\sigma_\mu^2}{\sigma_\varepsilon^2 + \sigma_\mu^2} \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & & 0 & 0 \\ 0 & 0 & 0 & & 0 & 0 \\ \vdots & & & \cdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \right) \\
&= \frac{1}{\sigma_\varepsilon^2} \begin{bmatrix} 2 - \Delta & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \cdots & 0 \\ 0 & -1 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix},
\end{aligned}$$

where $\Delta = \frac{\sigma_\mu^2}{\sigma_\varepsilon^2 + \sigma_\mu^2}$. Next, let's look at the GLS. First we note,

$$\begin{aligned}
X' \Omega^{-1} X &= X' (I_N \otimes \Sigma^{-1}) X \\
&= \begin{bmatrix} x' & x' & \cdots & x' & x' \end{bmatrix} \begin{bmatrix} \Sigma^{-1} & 0 & 0 & \cdots & 0 \\ 0 & \Sigma^{-1} & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \Sigma^{-1} \end{bmatrix} \begin{bmatrix} x \\ x \\ \vdots \\ x \\ x \end{bmatrix} = \sum_{i=1}^N x' \Sigma^{-1} x,
\end{aligned}$$

where

$$\begin{aligned}
x' \Sigma^{-1} x &= \frac{1}{\sigma_\varepsilon^2} \begin{bmatrix} 1 & 2 & 3 & \cdots & T \end{bmatrix} \begin{bmatrix} 2 - \Delta & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \cdots & 0 \\ 0 & -1 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ \vdots \\ T \end{bmatrix} \\
&= \frac{1}{\sigma_\varepsilon^2} \begin{bmatrix} -\Delta & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ \vdots \\ T \end{bmatrix} = \frac{1}{\sigma_\varepsilon^2} [T - \Delta].
\end{aligned}$$

Next,

$$\begin{aligned}
X' \Omega^{-1} y &= X' (I_N \otimes \Sigma^{-1}) Y \\
&= \begin{bmatrix} x' & x' & \cdots & x' & x' \end{bmatrix} \begin{bmatrix} \Sigma^{-1} & 0 & 0 & \cdots & 0 \\ 0 & \Sigma^{-1} & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \Sigma^{-1} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_N \end{bmatrix} \\
&= \begin{bmatrix} x' \Sigma^{-1} & x' \Sigma^{-1} & \cdots & x' \Sigma^{-1} & x' \Sigma^{-1} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_N \end{bmatrix} = \sum_{i=1}^N x' \Sigma^{-1} y_i,
\end{aligned}$$

where

$$\begin{aligned}
x' \Sigma^{-1} y_i &= \frac{1}{\sigma_\varepsilon^2} \begin{bmatrix} 1 & 2 & 3 & \cdots & T \end{bmatrix} \begin{bmatrix} 2 - \Delta & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \cdots & 0 \\ 0 & -1 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \\ \vdots \\ y_{iT} \end{bmatrix} \\
&= \frac{1}{\sigma_\varepsilon^2} \begin{bmatrix} -\Delta & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \\ \vdots \\ y_{iT} \end{bmatrix} = \frac{1}{\sigma_\varepsilon^2} [y_{iT} - \Delta y_{i1}].
\end{aligned}$$

It follows that GLS can be written as

$$\begin{aligned}
\widehat{\beta}_{GLS} &= (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} y \\
&= \left[\sum_{i=1}^N (x' \Sigma^{-1} x) \right]^{-1} \sum_{i=1}^N (x' \Sigma^{-1} y_i) \\
&= \left[\sum_{i=1}^N (T - \Delta) \right]^{-1} \left[\sum_{i=1}^N (y_{iT} - \Delta y_{i1}) \right].
\end{aligned}$$

Remark 10 If $\Delta = \frac{\sigma_\mu^2}{\sigma_\varepsilon^2 + \sigma_\mu^2} = 1$, then GLS is reduced to the FD estimator.

Remark 11 Since v_{it} is $I(1)$, then for all i ,

$$\frac{1}{\sqrt{T}} v_{i1} = \frac{1}{\sqrt{T}} \sum_{j=0}^{[\kappa T]} \varepsilon_{i1-j} \xrightarrow{d} N(0, \kappa \sigma_\varepsilon^2).$$

Remark 12 The μ_i can not be consistently estimated when the error term is $I(1)$.

The limiting distributions of OLS, FE, FD, and GLS estimators are given in the next Theorem.

Theorem 2 Let y_{it} be generated from a simple time trend model in (1) where

$$v_{it} = v_{it-1} + \varepsilon_{it},$$

with $\varepsilon_{it} \sim iid(0, \sigma_\varepsilon^2)$ and $v_{i1} = \sum_{j=0}^{[\kappa T]} \varepsilon_{i1-j}$. Then

- (a) $\sqrt{NT} (\hat{\beta}_{OLS} - \beta) \xrightarrow{d} N\left(0, \frac{3\sigma_\varepsilon^2(8+15\kappa)}{20}\right)$,
- (b) $\sqrt{NT} (\hat{\beta}_{FE} - \beta) \xrightarrow{d} N\left(0, \frac{6}{5}\sigma_\varepsilon^2\right)$,
- (c) $\sqrt{NT} (\hat{\beta}_{FD} - \beta) \xrightarrow{d} N(0, \sigma_\varepsilon^2)$,
- (d) $\sqrt{NT} (\hat{\beta}_{GLS} - \beta) \xrightarrow{d} N\left(0, \left[1 + (1 - \Delta)^2 \kappa\right] \sigma_\varepsilon^2\right)$.

Remark 13 The limiting distribution of the GLS depends on the variance component, $\Delta = \frac{\sigma_\mu^2}{\sigma_\varepsilon^2 + \sigma_\mu^2}$ and the parameter that governs the variance of the initial error, κ .

Remark 14 If $\Delta < 1$ then $\frac{1}{5} + \frac{5}{4}\kappa + 2\kappa\Delta - \kappa\Delta^2 > 0$. Hence the GLS estimator is also more efficient than the OLS estimator with $\rho = 1$. However, the GLS could be less efficient than FD or FE estimators. For example, the GLS is less efficient than FD unless $\kappa = 0$ or $\Delta = 1$. Also the GLS is more efficient than the FE only if $(1 - \Delta)^2 \kappa < \frac{1}{5}$.

Remark 15 Theorem 2 confirms the results of Baltagi and Chang (1992), i.e., the relative performance of the estimators depends on the variance of the initial error in panel data.

Remark 16 The FD estimator is the most efficient estimator with $\rho = 1$.

Remark 17 In Appendix E, we derive the limiting distribution of FE estimator when the error term, v_{it} , is nearly $I(1)$.

Remark 18 *It can be shown that*

$$\Sigma^{-1}x = \frac{1}{\sigma_\varepsilon^2} \begin{bmatrix} -\Delta \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

and then $R(\Sigma^{-1}x) \neq R(x)$ and $R(\Omega^{-1}X) \neq R(X)$. It follows that by Theorem 1 in Kruskal (1968) that the OLS and GLS are not equivalent even asymptotically.

Remark 19 *From Theorems 1 and 2 we know that when the error term is $I(0)$ the FE estimator is asymptotically efficient and when the error term is $I(1)$, the FD estimator is asymptotically efficient. Hence the inference on β in (1) can be carried out using the t -statistic from the FE when the error term is $I(0)$ and from the FD estimator when the error term is $I(1)$. The tests on whether the error term is $I(0)$ or $I(1)$ in (2) can be found in Kao (1997) and McCoskey and Kao (1998).*

4 The Effects of a Fitted Intercept

An intercept is not included in (1) since the intercept can not be consistently estimated by any method when the error term is $I(1)$. However, it is usual in empirical work for panel regression to include an intercept (e.g., Baltagi, 1995). Consider the following model in place of (1):

$$y_{it} = \alpha + \beta t + u_{it}, \tag{9}$$

where α is the intercept,

$$u_{it} = \mu_i + v_{it}$$

and

$$v_{it} = \rho v_{it-1} + \varepsilon_{it}, |\rho| < 1.$$

We note that

$$\widehat{\beta}_{OLS} = \widehat{\beta}_{FE} = \frac{\sum_{i=1}^N \sum_{t=1}^T (t - \bar{t}) y_{it}}{\sum_{i=1}^N \sum_{t=1}^T (t - \bar{t})^2}.$$

From the proof of Theorem 1 we know the inclusion of a fitted intercept in (9) does not alter the asymptotic distribution of $\widehat{\beta}_{FE}$ that is given in Theorem 1. Thus, the FE estimator has the same limiting distribution whether or not an intercept is included in the regression. However the limiting distribution of $\widehat{\beta}_{OLS}$ will be

affected by a fitted intercept. In this case, the limiting distribution of $\widehat{\beta}_{OLS}$ is

$$\sqrt{NT^3} \left(\widehat{\beta}_{OLS} - \beta \right) \xrightarrow{d} N \left(0, \frac{12\sigma_\varepsilon^2}{(1-\rho)^2} \right)$$

since the OLS is identical to the FE estimator under model (9). Also from the proof of Theorem 1 we note that the presence of the intercept does not influence the limiting distribution of the GLS estimator. The following two theorems provide the limiting distributions of the FE, FD and GLS estimators under model (9) with stationary error term and nonstationary error term.

Theorem 3 *Let y_{it} be generated from a simple time trend model in (9) where*

$$v_{it} = \rho v_{it-1} + \varepsilon_{it},$$

with $\varepsilon_{it} \sim iid(0, \sigma_\varepsilon^2)$ and initial condition in Assumption 1. Then

- (a) $\sqrt{NT^3} \left(\widehat{\beta}_{FE} - \beta \right) \xrightarrow{d} N \left(0, \frac{12\sigma_\varepsilon^2}{(1-\rho)^2} \right),$
- (b) $\sqrt{NT} \left(\widehat{\beta}_{FD} - \beta \right) \xrightarrow{d} N \left(0, \sigma_\varepsilon^2 \frac{2-\rho^{2(\kappa T+2)}}{1-\rho^2} \right),$
- (c) $\sqrt{NT^3} \left(\widehat{\beta}_{GLS} - \beta \right) \xrightarrow{d} N \left(0, \frac{12\sigma_\varepsilon^2}{(1-\rho)^2} \right).$

Remark 20 *Theorem 3 shows that, when the error term is $I(0)$, the presence of the intercept does not change the limiting distribution of the GLS estimator.*

Theorem 4 *Let y_{it} be generated from a simple time trend model in (9) where*

$$v_{it} = v_{it-1} + \varepsilon_{it},$$

with $\varepsilon_{it} \sim iid(0, \sigma_\varepsilon^2)$ and initial condition in Assumption 1. Then

- (a) $\sqrt{NT} \left(\widehat{\beta}_{FE} - \beta \right) \xrightarrow{d} N \left(0, \frac{6}{5} \sigma_\varepsilon^2 \right),$
- (b) $\sqrt{NT} \left(\widehat{\beta}_{FD} - \beta \right) \xrightarrow{d} N \left(0, \sigma_\varepsilon^2 \right),$
- (c) $\sqrt{NT} \left(\widehat{\beta}_{GLS} - \beta \right) \xrightarrow{d} N \left(0, 4\kappa \sigma_\varepsilon^2 \right)$ if $\kappa > 0,$
- (d) $\sqrt{NT} \left(\widehat{\beta}_{GLS} - \beta \right) \xrightarrow{d} N \left(0, 4 \left(\sigma_\mu^2 + \sigma_\varepsilon^2 \right) \right)$ if $\kappa = 0.$

Remark 21 *Theorem 4 shows that the limiting distribution of the GLS estimator depends on the initialization parameter κ when the error term is $I(1)$ and the presence of the intercept. The GLS estimator is the most efficient estimator since it is \sqrt{NT} consistent while the other estimators are \sqrt{NT} consistent if $\kappa = 0$. However, the GLS estimator could be less efficient than the FE and FD if $\kappa > 0$ and large. For example, the GLS is less efficient than the FD if $\kappa > \frac{1}{4}$ and is less efficient than the FE if $\kappa > \frac{3}{10}$.*

Remark 22 If v_{it} is $I(1)$, i.e., $\rho = 1$, then the FE and FD estimators have the same limiting distributions that are given in Theorem 2.

5 Feasible GLS Estimators

For the feasible GLS estimators we need the estimation of the variance components, Δ , and the autocorrelation coefficient ρ . The parameter ρ can be estimated easily, i.e.,

$$\hat{\rho} = \frac{\sum_{i=1}^N \sum_{t=2}^T \hat{u}_{it} \hat{u}_{it-1}}{\sum_{i=1}^N \sum_{t=2}^T (\hat{u}_{it-1})^2}, \quad (10)$$

where \hat{u}_{it} is the estimated residual, taken from the FE estimation of the model in (1). It can be shown that $\hat{\rho}$ in (10) is a consistent estimator of ρ by using (b) in Theorem 1. The variance component can be estimated in the same way as for the models without autocorrelation by using the variance decomposition and the Prais-Winsten (PW) transformation as Baltagi and Li (1991) pointed out.

On the other hand, the efficiency of the GLS estimator also relies on κ as we know from Theorem 2. Maeshiro (1976) pointed out that the Cochrane-Orcutt (CO) procedure, which ignores the information contained in the first observation, performs worse than OLS for smoothly trended regressors. Beach and MacKinnon (1978) and Park and Mitchell (1980) suggested that when the regressors are trended, estimation using the PW transformation is more efficient than using the CO procedure. The importance of the initial observation in panel data regression with AR(1) error terms has been studied recently by Baltagi and Chang (1992). However, when $\rho = 1$, μ_i can no longer be consistently estimated by any method and hence σ_μ^2 is no longer identifiable. It limits the usefulness of the GLS estimator when the error term is $I(1)$.

6 Finite Sample Simulations of Estimators

In this section we will evaluate the finite sample properties of the OLS, FE, FD, GLS-CO, GLS-PW and infeasible GLS estimators and ask whether the $I(0)$ and $I(1)$ asymptotic variance provide a useful guide for choosing among the estimators in small samples. The model is set as follows:

$$y_{it} = \mu_i + \beta t + u_{it}, i = 1, \dots, N, t = 1, \dots, T,$$

where $\beta = 1$ and with $\mu_i \stackrel{iid}{\sim} N(0, \sigma_\mu^2)$ and v_{it} follows an AR(1) with

$$v_{it} = \rho v_{it-1} + \varepsilon_{it},$$

$\varepsilon_{it} \stackrel{iid}{\sim} N(0, \sigma_\varepsilon^2)$ and $v_{i1} = \sum_{j=0}^{[\kappa T]} \rho^j \varepsilon_{i1-j}$, where $\kappa = (0, 0.1, 0.25, 1.0)$ and $\rho = (0, 0.2, 0.4, 0.6, 0.8, 0.9, 0.95, 1.0)$. We fix $\sigma_\mu^2 + \sigma_\varepsilon^2 = 10$ and let $\Delta = \frac{\sigma_\mu^2}{\sigma_\varepsilon^2 + \sigma_\mu^2}$ take the values $(0, 0.2, 0.4, 0.6, 0.8)$. The following sample size combinations are used: $N = 25, 50$ and $T = 25, 50$. Each experiment involves 10,000 replications. For each replication we estimated the model using OLS, FE, FD, GLS-CO, GLS-PW and infeasible GLS estimators. The simulations were performed by a Ultra Enterprise 3000. GAUSS 3.2.31 was used to perform the simulations. Random numbers for μ_i and ε_{it} were generated by the GAUSS procedure RNDNS. At each replication, we generated an $N(T + 1000)$ length of random numbers and then split it into N series so that each series had the same mean and variance. The first 1,000 observations were discarded for each series.

Tables 1-2 give the mean square error (MSE) of the various estimators of β relative to the infeasible GLS estimator for various values of ρ , Δ , and κ when $N = T = 25$. We do not report the cases when N and T are more than 25 since the results are similar to Tables 1 and 2. We also do not report the bias which was negligibly small for all the estimators. The FE, GLS-CO and GLS-PW estimators are essentially efficient and are preferred to the FD and OLS estimators as predicted from Theorem 1 with $\rho < 1$ but not too close to 1. However, when $\rho = 1$, the FD estimator is efficient and is preferred to the OLS and FE estimators as predicted from Theorem 2. The OLS, GLS-CO and GLS-PW estimators perform poorly for large values of Δ . Interestingly, the FE and FD estimators perform well for large values of Δ . This observation was also noted by Baltagi (1981, p. 43). In general, the FD estimator is better than the OLS when Δ is large and is better than the FE estimator when ρ is closer to 1. However, the FD estimator is worse than the OLS when Δ is small and also worse than the FE when ρ is small. The relative performance of the OLS, FE and FD also depends critically on κ when $\rho = 1$. When $\kappa = 0$, the FD estimator is as efficient as the infeasible GLS estimator and dominates the OLS and FE estimators. As we expected from Theorem 2, the FD estimator is more efficient than the infeasible GLS estimator when $\kappa > .1$, though the better performance of the FD estimator decreases as Δ increases. Overall, the FE estimator performs relatively well when compared to other methods.

7 Conclusion

This paper considers a linear time trend model with one-way error component with serially correlated error term and studies the limiting distributions of the OLS, FE, FD, GLS-CO, and GLS-PW estimators. The results are confirmed by means of Monte Carlo experiments. The main findings are:

1. The GLS-CO and GLS-PW estimators perform well if the variance component, Δ , is small and close to zero when $\rho < 1$.

2. The FD estimator dominates the other estimators when $\rho = 1$ for all values of Δ .
3. The FE estimator is recommend in practice since it performs pretty well for all values of ρ and Δ .

Appendix

A Proof of Theorem 1

Proof. (b) Equation (3) can be expressed as

$$\widehat{\beta}_{FE} - \beta = \frac{\sum_{i=1}^N \sum_{t=1}^T (t - \bar{t}) u_{it}}{\sum_{i=1}^N \sum_{t=1}^T (t - \bar{t})^2}. \quad (11)$$

We multiply (11) by $\sqrt{NT^3}$, resulting in

$$\sqrt{NT^3} (\widehat{\beta}_{FE} - \beta) = \frac{\frac{1}{\sqrt{NT^3}} \sum_{i=1}^N \sum_{t=1}^T (t - \bar{t}) u_{it}}{\frac{1}{NT^3} \sum_{i=1}^N \sum_{t=1}^T (t - \bar{t})^2}. \quad (12)$$

It is straightforward to show that

$$\begin{aligned} \sum_{i=1}^N \sum_{t=1}^T (t - \bar{t})^2 &= \frac{NT(T^2 - 1)}{12} \\ &= O(N^{-1}T^3). \end{aligned} \quad (13)$$

Thus, the leading term in $\sum_{i=1}^N \sum_{t=1}^T (t - \bar{t})^2$ is $\frac{NT^3}{12}$, that is

$$\frac{1}{NT^3} \sum_{i=1}^N \sum_{t=1}^T (t - \bar{t})^2 \rightarrow \frac{1}{12}.$$

We turn next to the numerator in (12). It can be shown that (e.g., Hamilton, 1994) if $\psi(L)$ is a possible infinite polynomial in the lag operator L , such that $\psi(z)$ has all of its roots outside the unit circle then

$$T^{-3/2} \sum_{t=1}^T t \psi(L) \varepsilon_t \xrightarrow{d} N \left(0, \frac{\sigma_\varepsilon^2 \psi(1)^2}{3} \right)$$

and

$$T^{-3/2} \sum_{t=1}^T (t - \bar{t}) \psi(L) \varepsilon_t \xrightarrow{d} N \left(0, \frac{\sigma_\varepsilon^2 \psi(1)^2}{12} \right).$$

Note that $v_{it} = \sum_{j=0}^{\infty} \rho^j L^j \varepsilon_t$. Choose $\psi(L) = \sum_{j=0}^{\infty} \rho^j L^j$ and $\psi(1) = \sum_{j=0}^{\infty} \rho^j = \frac{1}{1-\rho}$ we have

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N T^{-3/2} \sum_{t=1}^T (t - \bar{t}) u_{it} = \frac{1}{\sqrt{N}} \sum_{i=1}^N T^{-3/2} \sum_{t=1}^T (t - \bar{t}) (\mu_i + v_{it})$$

$$\begin{aligned}
&= \frac{1}{\sqrt{N}} \sum_{i=1}^N T^{-3/2} \left\{ \mu_i \sum_{t=1}^T (t - \bar{t}) + \sum_{t=1}^T (t - \bar{t}) v_{it} \right\} \\
&= \frac{1}{\sqrt{N}} \sum_{i=1}^N T^{-3/2} \sum_{t=1}^T (t - \bar{t}) v_{it} \\
&= \frac{1}{\sqrt{N}} \sum_{i=1}^N T^{-3/2} \sum_{t=1}^T (t - \bar{t}) \psi(L) \varepsilon_t.
\end{aligned}$$

Now for fixed N as $T \rightarrow \infty$ we have

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N T^{-3/2} \sum_{t=1}^T (t - \bar{t}) \psi(L) \varepsilon_t \xrightarrow{d} \frac{1}{\sqrt{N}} \sum_{i=1}^N N \left(0, \frac{\sigma_\varepsilon^2}{12} \frac{1}{(1 - \rho)^2} \right)$$

and

$$\frac{\frac{1}{\sqrt{N} T^3} \sum_{i=1}^N \sum_{t=1}^T (t - \bar{t}) u_{it}}{\frac{1}{N T^3} \sum_{i=1}^N \sum_{t=1}^T (t - \bar{t})^2} \xrightarrow{d} \frac{1}{\sqrt{N}} \sum_{i=1}^N N \left(0, \frac{12 \sigma_\varepsilon^2}{(1 - \rho)^2} \right).$$

Obviously

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N N \left(0, \frac{12 \sigma_\varepsilon^2}{(1 - \rho)^2} \right) \xrightarrow{d} N \left(0, \frac{12 \sigma_\varepsilon^2}{(1 - \rho)^2} \right)$$

as $N \rightarrow \infty$ (in fact, it is true for all N) proving (b). (a) can be shown similarly by following the proof of (b).

To see this, first we note that

$$\hat{\beta}_{OLS} - \beta = \frac{\sum_{i=1}^N \sum_{t=1}^T t u_{it}}{\sum_{i=1}^N \sum_{t=1}^T t^2}.$$

Then for a fixed N , we obtain

$$\begin{aligned}
\sum_{i=1}^N \sum_{t=1}^T t u_{it} &= \sum_{i=1}^N \sum_{t=1}^T t (\mu_i + v_{it}) \\
&= \sum_{i=1}^N \mu_i \sum_{t=1}^T t + \sum_{i=1}^N \sum_{t=1}^T t v_{it} \\
&= \sum_{i=1}^N \mu_i \frac{1}{2} O(T^2) + O(T^{3/2}).
\end{aligned}$$

It follows that for a fixed N as $T \rightarrow \infty$,

$$\frac{1}{T^2} \sum_{i=1}^N \sum_{t=1}^T t u_{it} = \frac{1}{2} \sum_{i=1}^N \mu_i + o_p(1).$$

Hence,

$$\frac{1}{\sqrt{N}} \frac{1}{T^2} \sum_{i=1}^N \sum_{t=1}^T t u_{it} \xrightarrow{d} N \left(0, \frac{1}{4} \sigma_\mu^2 \right)$$

as $N \rightarrow \infty$. Finally,

$$\begin{aligned}\sqrt{NT} (\hat{\beta}_{OLS} - \beta) &= \frac{\frac{1}{\sqrt{N}} \frac{1}{T^2} \sum_{i=1}^N \sum_{t=1}^T tu_{it}}{\frac{1}{N} \frac{1}{T^3} \sum_{i=1}^N \sum_{t=1}^T t^2} \\ &= 3 \frac{1}{\sqrt{N}} \frac{1}{T^2} \sum_{i=1}^N \sum_{t=1}^T tu_{it} + o(1) \xrightarrow{d} N \left(0, \frac{9}{4} \sigma_\mu^2 \right)\end{aligned}$$

proving (a). Note that

$$\begin{aligned}x' A^{-1} x &= \frac{1}{1 - \rho^2} x' \left\{ (1 - \rho)^2 x + \begin{bmatrix} \rho^2 \\ 0 \\ 0 \\ \vdots \\ 0 \\ \rho(T - T\rho + 1) \end{bmatrix} \right\} \\ &= \frac{(1 - \rho)^2}{1 - \rho^2} \sum_{t=1}^T t^2 + \frac{1}{1 - \rho^2} (\rho^2 + T\rho(T - T\rho + 1)),\end{aligned}$$

$$\begin{aligned}x' a_T a_T' x &= \left[\frac{1 + (1 - \rho) \left(\sum_{t=2}^{T-1} t \right) + T}{(1 + \rho)^2} \right] x' \begin{bmatrix} 1 \\ 1 - \rho \\ 1 - \rho \\ \vdots \\ 1 - \rho \\ 1 \end{bmatrix} \\ &= \frac{1}{(1 + \rho)^2} \left[1 + (1 - \rho) \left(\sum_{t=2}^{T-1} t \right) + T \right]^2,\end{aligned}$$

$$x' A^{-1} y_i = \frac{1}{1 - \rho^2} x' \begin{bmatrix} 1 & -\rho & 0 & 0 & 0 & \cdots & \cdots & 0 \\ -\rho & 1 + \rho^2 & -\rho & 0 & 0 & \cdots & \cdots & 0 \\ 0 & -\rho & 1 + \rho^2 & -\rho & 0 & \cdots & \cdots & 0 \\ & & & \vdots & & & & \\ 0 & 0 & 0 & 0 & \cdots & -\rho & 1 + \rho^2 & -\rho \\ 0 & 0 & 0 & 0 & \cdots & 0 & -\rho & 1 \end{bmatrix} \begin{bmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \\ \vdots \\ y_{iT-1} \\ y_{iT} \end{bmatrix}$$

$$\begin{aligned}
&= \frac{1}{1-\rho^2} \begin{bmatrix} 1-2\rho & 2(1-\rho)^2 & 3(1-\rho)^2 & \cdots & (T-1)(1-\rho)^2 & T-(T-1)\rho \end{bmatrix} \begin{bmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \\ \vdots \\ y_{iT-1} \\ y_{iT} \end{bmatrix} \\
&= \frac{1}{1-\rho^2} \left\{ y_{i1}(1-2\rho) + y_{i2}2(1-\rho)^2 + y_{i3}3(1-\rho)^2 + \cdots + y_{iT}[T-(T-1)\rho] \right\} \\
&= \frac{1}{1-\rho^2} \left\{ y_{i1}(1-2\rho) + (1-\rho)^2 \sum_{t=2}^{T-1} ty_{it} + y_{iT}[T-(T-1)\rho] \right\},
\end{aligned}$$

and

$$\begin{aligned}
x' a_T a_T' y_i &= \left(\frac{1}{1+\rho} \right)^2 x' \begin{bmatrix} 1 \\ 1-\rho \\ 1-\rho \\ \vdots \\ 1-\rho \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1-\rho & 1-\rho & \cdots & 1-\rho & 1 \end{bmatrix} y_i \\
&= \left(\frac{1}{1+\rho} \right)^2 \left[1+T+(1-\rho) \sum_{t=2}^{T-1} t \right] \left[y_{i1} + (1-\rho) \sum_{t=2}^{T-1} y_{it} + y_{iT} \right].
\end{aligned}$$

Then

$$\begin{aligned}
x' \Sigma^{-1} x &= \frac{1}{\sigma_v^2} x' \left(A^{-1} - \frac{\sigma_\mu^2}{\sigma_v^2 + \theta \sigma_\mu^2} a_T a_T' \right) x \\
&= \frac{1}{\sigma_v^2} x' A^{-1} x - \frac{1}{\sigma_v^2} \frac{\sigma_\mu^2}{\sigma_v^2 + \theta \sigma_\mu^2} x a_T a_T' x \\
&= \frac{(1-\rho)^2}{\sigma_\varepsilon^2} \sum_{t=1}^T t^2 + \frac{1}{\sigma_\varepsilon^2} [\rho^2 + T\rho(T-T\rho+1)] \\
&\quad - \frac{1}{\sigma_v^2} \frac{\sigma_\mu^2}{\sigma_v^2 + \theta \sigma_\mu^2} \frac{1}{(1+\rho)^2} \left[1 + (1-\rho) \left(\sum_{t=2}^{T-1} t \right) + T \right]^2, \\
\frac{1}{T^3} x' \Sigma^{-1} x &= \frac{1}{3} \frac{(1-\rho)^2}{\sigma_\varepsilon^2} - \frac{1}{4} \frac{(1-\rho)^2}{\sigma_\varepsilon^2} \frac{1-\rho^2}{(1+\rho)^2} \frac{1+\rho}{1-\rho} + o(1) \\
&= \frac{1}{3} \frac{(1-\rho)^2}{\sigma_\varepsilon^2} - \frac{1}{4} \frac{(1-\rho)^2}{\sigma_\varepsilon^2} + o(1)
\end{aligned}$$

$$= \frac{1}{12} \frac{(1-\rho)^2}{\sigma_\varepsilon^2} + o(1),$$

$$\begin{aligned} x' \Sigma^{-1} y_i &= \frac{1}{\sigma_v^2} x' \left(A^{-1} - \frac{\sigma_\mu^2}{\sigma_v^2 + \theta \sigma_\mu^2} a_T a_T' \right) y_i \\ &= \frac{1}{\sigma_v^2} x' A^{-1} y_i - \frac{1}{\sigma_v^2} \frac{\sigma_\mu^2}{\sigma_v^2 + \theta \sigma_\mu^2} x' a_T a_T' y_i \\ &= \frac{1}{\sigma_v^2} \frac{1}{1-\rho^2} \left\{ y_{i1} (1-2\rho) + (1-\rho)^2 \sum_{t=2}^{T-1} t y_{it} + y_{iT} [T - (T-1)\rho] \right\} \\ &\quad - \frac{1}{\sigma_v^2} \frac{\sigma_\mu^2}{\sigma_v^2 + \theta \sigma_\mu^2} \left(\frac{1}{1+\rho} \right)^2 \left[1 + T + (1-\rho) \sum_{t=2}^{T-1} t \right] \left[y_{i1} + (1-\rho) \sum_{t=2}^{T-1} y_{it} + y_{iT} \right], \end{aligned}$$

$$\begin{aligned} x' \Sigma^{-1} u_i &= \frac{1}{\sigma_v^2} x' \left(A^{-1} - \frac{\sigma_\mu^2}{\sigma_v^2 + \theta \sigma_\mu^2} a_T a_T' \right) u_i \\ &= \frac{1}{\sigma_\varepsilon^2} \left\{ u_{i1} (1-2\rho) + (1-\rho)^2 \sum_{t=2}^{T-1} t u_{it} + u_{iT} [T - (T-1)\rho] \right\} \\ &\quad - \frac{1}{\sigma_v^2} \frac{\sigma_\mu^2}{\sigma_v^2 + \theta \sigma_\mu^2} \left(\frac{1}{1+\rho} \right)^2 \left[1 + T + (1-\rho) \sum_{t=2}^{T-1} t \right] \left[u_{i1} + (1-\rho) \sum_{t=2}^{T-1} u_{it} + u_{iT} \right] \\ &= \frac{1}{\sigma_\varepsilon^2} \left\{ (\mu_i + v_{i1}) (1-2\rho) + (1-\rho)^2 \sum_{t=2}^{T-1} t (\mu_i + v_{it}) + (\mu_i + v_{iT}) [T - (T-1)\rho] \right\} \\ &\quad - \frac{1}{\sigma_v^2} \frac{\sigma_\mu^2}{\sigma_v^2 + \theta \sigma_\mu^2} \left(\frac{1}{1+\rho} \right)^2 \\ &\quad \left[1 + T + (1-\rho) \sum_{t=2}^{T-1} t \right] \left[(\mu_i + v_{i1}) + (1-\rho) \sum_{t=2}^{T-1} (\mu_i + v_{it}) + (\mu_i + v_{iT}) \right], \end{aligned} \tag{14}$$

and

$$\begin{aligned} \frac{1}{T^{3/2}} x' \Sigma^{-1} u_i &= \frac{(1-\rho)^2}{\sigma_\varepsilon^2} \frac{1}{T^{3/2}} \left(\mu_i \sum_{t=2}^{T-1} t + \sum_{t=2}^{T-1} t v_{it} \right) \\ &\quad - \frac{1}{\sigma_\varepsilon^2} \left(\frac{1-\rho}{1+\rho} \right) (1-\rho)^2 \left[\frac{\sigma_\mu^2}{\sigma_v^2 + \theta \sigma_\mu^2} (T-2) \right] \frac{1}{T^{3/2}} \mu_i \sum_{t=2}^{T-1} t \\ &\quad - \frac{1}{\sigma_\varepsilon^2} \left(\frac{1-\rho}{1+\rho} \right) (1-\rho)^2 \left[\frac{1}{T} \frac{\sigma_\mu^2}{\sigma_v^2 + \theta \sigma_\mu^2} \left(\sum_{t=2}^{T-1} t \right) \right] \frac{1}{T^{1/2}} \sum_{t=2}^{T-1} v_{it} \\ &\quad + \frac{1}{\sigma_\varepsilon^2} \left(\frac{1-\rho}{1+\rho} \right) (1-\rho) \left[\frac{1}{T} \frac{\sigma_\mu^2}{\sigma_v^2 + \theta \sigma_\mu^2} \left(\sum_{t=2}^{T-1} t \right) \right] \left[\frac{1}{T^{1/2}} (\mu_i + v_{iT}) \right] + o_p(1) \\ &= \frac{(1-\rho)^2}{\sigma_\varepsilon^2} \frac{1}{T^{3/2}} \sum_{t=2}^{T-1} t v_{it} + \frac{(1-\rho)^2}{\sigma_\varepsilon^2} \frac{1}{T^{3/2}} \mu_i \sum_{t=2}^{T-1} t \end{aligned}$$

$$\begin{aligned}
& -\frac{(1-\rho)^2}{\sigma_\varepsilon^2} \frac{1}{T^{3/2}} \mu_i \sum_{t=2}^{T-1} t \\
& -\frac{1}{\sigma_\varepsilon^2} \frac{1}{2} \left(\frac{1-\rho}{1+\rho} \right) (1-\rho)^2 \frac{1+\rho}{1-\rho} \frac{1}{T^{1/2}} \sum_{t=2}^{T-1} v_{it} \\
& -\frac{1}{\sigma_\varepsilon^2} \frac{1}{2} \left(\frac{1-\rho}{1+\rho} \right) (1-\rho) \frac{1+\rho}{1-\rho} \left[\frac{1}{T^{1/2}} (\mu_i + v_{iT}) \right] + o(1) \\
= & \frac{(1-\rho)^2}{\sigma_\varepsilon^2} \frac{1}{T^{3/2}} \sum_{t=2}^{T-1} t v_{it} - \frac{(1-\rho)^2}{\sigma_\varepsilon^2} \frac{1}{2} \frac{1}{T^{1/2}} \sum_{t=2}^{T-1} v_{it} + o_p(1) \\
= & \frac{(1-\rho)^2}{\sigma_\varepsilon^2} \left[\frac{1}{T^{3/2}} \sum_{t=2}^{T-1} t v_{it} - \frac{1}{2} \frac{1}{T^{1/2}} \sum_{t=2}^{T-1} v_{it} \right]
\end{aligned}$$

since

$$\begin{aligned}
\lim_{T \rightarrow \infty} \frac{1}{T^3} \frac{\sigma_\mu^2}{\sigma_v^2 + \theta \sigma_\mu^2} \left(\sum_{t=2}^{T-1} t \right)^2 &= \frac{1}{4} \frac{1+\rho}{1-\rho}, \\
\lim_{T \rightarrow \infty} \frac{\sigma_\mu^2}{\sigma_v^2 + \theta \sigma_\mu^2} (T-2) &= \frac{1+\rho}{1-\rho},
\end{aligned}$$

and

$$\lim_{T \rightarrow \infty} \frac{1}{T} \frac{\sigma_\mu^2}{\sigma_v^2 + \theta \sigma_\mu^2} \left(\sum_{t=2}^{T-1} t \right) = \frac{1}{2} \frac{1+\rho}{1-\rho}.$$

It is known that

$$\begin{bmatrix} T^{-1/2} \sum_{t=2}^{T-1} v_{it} \\ T^{-3/2} \sum_{t=2}^{T-1} t v_{it} \end{bmatrix} \xrightarrow{d} N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \frac{\sigma_\varepsilon^2}{(1-\rho)^2} \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{bmatrix} \right)$$

and then

$$\frac{1}{T^{3/2}} x' \Sigma^{-1} u_i \xrightarrow{d} \frac{(1-\rho)^2}{\sigma_\varepsilon^2} N \left(0, \frac{\sigma_\varepsilon^2}{12(1-\rho)^2} \right).$$

It follows that for a fixed N as $T \rightarrow \infty$,

$$\begin{aligned}
\frac{1}{N} \frac{1}{T^3} X \Omega^{-1} X &= \frac{1}{N} \frac{1}{T^3} \sum_{i=1}^N x' \Sigma^{-1} x \\
&= \frac{1}{12} \frac{(1-\rho)^2}{\sigma_\varepsilon^2} + o(1)
\end{aligned}$$

and

$$\begin{aligned}
\frac{1}{\sqrt{N}} \frac{1}{T^{3/2}} X' \Omega^{-1} u_i &= \frac{1}{\sqrt{N}} \frac{1}{T^{3/2}} \sum_{i=1}^N x' \Sigma^{-1} u_i \\
&\xrightarrow{d} \frac{(1-\rho)^2}{\sigma_\varepsilon^2} N \left(0, \frac{1\sigma_\varepsilon^2}{12(1-\rho)^2} \right).
\end{aligned}$$

Then

$$\begin{aligned}
\sqrt{NT^3} \left(\widehat{\beta}_{GLS} - \beta \right) &= \frac{\frac{1}{\sqrt{N}} \frac{1}{T^{3/2}} \sum_{i=1}^N x' \Sigma^{-1} u_i}{\frac{1}{N} \frac{1}{T^3} \sum_{i=1}^N x' \Sigma^{-1} x} \\
&= 12 \frac{\sigma_\varepsilon^2}{(1-\rho)^2} \frac{1}{\sqrt{N}} \frac{1}{T^{3/2}} \sum_{i=1}^N x' \Sigma^{-1} u_i + o_p(1) \\
&\xrightarrow{d} N \left(0, \frac{12\sigma_\varepsilon^2}{(1-\rho)^2} \right)
\end{aligned}$$

proving (d). Finally, we show the limiting distribution of the FD estimator.

$$\begin{aligned}
\sqrt{NT} \left(\widehat{\beta}_{FD} - \beta \right) &= \frac{\frac{1}{\sqrt{N}} \sum_{i=1}^N (u_{iT} - v_{i1})}{\frac{1}{NT} N(T-1)} \\
&= \frac{\frac{1}{\sqrt{N}} \sum_{i=1}^N (v_{iT} - v_{i1})}{\frac{1}{NT} N(T-1)}.
\end{aligned}$$

Using

$$v_{iT} - v_{i1} = \sum_{j=0}^{\infty} \rho^j \varepsilon_{T-j} - \sum_{j=0}^{[\kappa T]} \rho^j \varepsilon_{1-j}$$

and

$$\begin{aligned}
Var(v_{iT} - v_{i1}) &= Var \left(\sum_{j=0}^{\infty} \rho^j \varepsilon_{T-j} \right) + Var \left(\sum_{j=0}^{[\kappa T]} \rho^j \varepsilon_{1-j} \right) \\
&= \frac{\sigma_\varepsilon^2}{1-\rho^2} + \frac{\sigma_\varepsilon^2 (1 - \rho^{2T+2})}{1-\rho^2} \\
&= \frac{\sigma_\varepsilon^2}{1-\rho^2} \left(2 - \rho^{2[\kappa T]+2} \right).
\end{aligned}$$

Then

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N (v_{iT} - v_{i1}) \xrightarrow{d} N \left(0, \sigma_\varepsilon^2 \frac{2 - \rho^{2[\kappa T]+2}}{1 - \rho^2} \right)$$

and hence

$$\sqrt{NT} \left(\widehat{\beta}_{FD} - \beta \right) \xrightarrow{d} N \left(0, \sigma_\varepsilon^2 \frac{2 - \rho^{2[\kappa T]+2}}{1 - \rho^2} \right)$$

proving (c). ■

B Proof of Theorem 2

Following Canjels and Watson (1997) we note that when $\rho = 1$

$$\frac{1}{\sqrt{T}} v_{i1} = \sum_{j=1}^{[T\kappa]} \varepsilon_{i1-j} \xrightarrow{d} \sigma_\varepsilon \widetilde{W}(\kappa)$$

and

$$\frac{1}{\sqrt{T}}v_{it} = \frac{1}{\sqrt{T}} \sum_{j=1}^{[Tr]} \varepsilon_{ij} \xrightarrow{d} \sigma_\varepsilon \left[W(r) + \widetilde{W}(\kappa) \right]$$

for all i , where $W(r)$ and $\widetilde{W}(\kappa)$ are independent standard Brownian motion. We write the FE as

$$\begin{aligned} \sqrt{NT} \left(\widehat{\beta}_{FE} - \beta \right) &= \frac{\frac{1}{\sqrt{N}} \sum_{i=1}^N \frac{1}{\sqrt{T^5}} \sum_{t=1}^T (t - \bar{t}) u_{it}}{\frac{1}{N} \sum_{i=1}^N \frac{1}{T^3} \sum_{t=1}^T (t - \bar{t})^2} \\ &= \frac{\frac{1}{\sqrt{N}} \sum_{i=1}^N \frac{1}{\sqrt{T^5}} \sum_{t=1}^T (t - \bar{t}) v_{it}}{\frac{1}{N} \sum_{i=1}^N \frac{1}{T^3} \sum_{t=1}^T (t - \bar{t})^2}. \end{aligned} \quad (15)$$

Let us begin with the numerator in (15). Using Proposition 17.1 in Hamilton (1994), we see that for a fixed N as $T \rightarrow \infty$

$$T^{-5/2} \sum_{t=1}^T (t - \bar{t}) v_{it} \xrightarrow{d} \sigma_\varepsilon \int_0^1 \left(r - \frac{1}{2} \right) \left[W(r) + \widetilde{W}(\kappa) \right] dr \sim N \left(0, \frac{\sigma_\varepsilon^2}{120} \right).$$

Then

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N \frac{1}{\sqrt{T^5}} \sum_{t=1}^T (t - \bar{t}) v_{it} \xrightarrow{d} N \left(0, \frac{\sigma_\varepsilon^2}{120} \right)$$

as $N \rightarrow \infty$. Thus

$$\sqrt{NT} \left(\widehat{\beta}_{FE} - \beta \right) \xrightarrow{d} N \left(0, \frac{6}{5} \sigma_\varepsilon^2 \right)$$

proving (b). Similarly,

$$\begin{aligned} \frac{1}{\sqrt{N}} \sum_{i=1}^N \frac{1}{\sqrt{T^5}} \sum_{t=1}^T t u_{it} &= \frac{1}{\sqrt{N}} \sum_{i=1}^N \frac{1}{\sqrt{T^5}} \sum_{t=1}^T t (\mu_i + v_{it}) \\ &= \frac{1}{\sqrt{N}} \sum_{i=1}^N \frac{1}{\sqrt{T^5}} \left[\mu_i \sum_{t=1}^T t + \sum_{t=1}^T t v_{it} \right] \\ &= \frac{1}{\sqrt{N}} \sum_{i=1}^N \frac{1}{\sqrt{T^5}} \left[\sum_{t=1}^T t v_{it} \right] + o_p(1). \end{aligned}$$

It follows that

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N \frac{1}{\sqrt{T^5}} \sum_{t=1}^T t u_{it} \xrightarrow{d} \sigma_\varepsilon \int_0^1 r \left[W(r) + \widetilde{W}(\kappa) \right] dr \sim N \left(0, \frac{2\sigma_\varepsilon^2}{15} + \frac{\kappa\sigma_\varepsilon^2}{4} \right)$$

as $T \rightarrow \infty$ and then $N \rightarrow \infty$.

$$\begin{aligned} \sqrt{NT} \left(\widehat{\beta}_{OLS} - \beta \right) &= 3 \frac{1}{\sqrt{N}} \sum_{i=1}^N \frac{1}{\sqrt{T^5}} \sum_{t=1}^T t v_{it} + o_p(1) \\ &\xrightarrow{d} N \left(0, \frac{3\sigma_\varepsilon^2(8 + 15\kappa)}{20} \right) \end{aligned}$$

as $T \rightarrow \infty$ and then $N \rightarrow \infty$ proving (a). Next we prove (c).

$$\sqrt{NT} \left(\widehat{\beta}_{GLS} - \beta \right) = \left[\frac{1}{N} \sum_{i=1}^N \frac{1}{T} (T - \Delta) \right]^{-1} \left[\frac{1}{\sqrt{N}} \sum_{i=1}^N \frac{1}{\sqrt{T}} (u_{iT} - \Delta u_{i1}) \right].$$

It is easy to see that for a fixed N as $T \rightarrow \infty$ that

$$\frac{1}{T} (T - \Delta) \rightarrow 1$$

and

$$\begin{aligned} \frac{1}{\sqrt{T}} (u_{iT} - \Delta u_{i1}) &= \frac{1}{\sqrt{T}} u_{iT} - \Delta \frac{1}{\sqrt{T}} u_{i1} \\ &= \frac{1}{\sqrt{T}} (\mu_i + v_{iT}) - \Delta \frac{1}{\sqrt{T}} (\mu_i + v_{i1}) \\ &= \frac{1}{\sqrt{T}} v_{iT} - \Delta \frac{1}{\sqrt{T}} v_{i1} + \frac{1}{\sqrt{T}} (\mu_i - \Delta \mu_i) \\ &= \frac{1}{\sqrt{T}} \tilde{v}_{iT} + \frac{1}{\sqrt{T}} v_{i1} - \Delta \frac{1}{\sqrt{T}} v_{i1} + o_p(1) \\ &= \frac{1}{\sqrt{T}} \tilde{v}_{iT} + (1 - \Delta) \frac{1}{\sqrt{T}} v_{i1} \\ &= \sigma_\varepsilon \left(W(1) + (1 - \Delta)^2 \widetilde{W}(\kappa) \right) \xrightarrow{d} N \left(0, \sigma_\varepsilon^2 \left[1 + (1 - \Delta)^2 \kappa \right] \right), \end{aligned}$$

where $\tilde{v}_{iT} = v_{iT} - v_{i1}$. Then

$$\sqrt{NT} \left(\widehat{\beta}_{GLS} - \beta \right) \xrightarrow{d} N \left(0, \sigma_\varepsilon^2 \left[1 + (1 - \Delta)^2 \kappa \right] \right)$$

as $N \rightarrow \infty$ proving (d). It can be shown easily that the FD estimator has the following limiting distribution:

$$\sqrt{NT} \left(\widehat{\beta}_{FD} - \beta \right) \xrightarrow{d} N \left(0, \sigma_\varepsilon^2 \right)$$

proving (c). ■

C Proof of Theorem 3

Proof. It will be convenient to center the observation so that $z = t - \bar{t}$. From the proof of Theorem 1 we know

$$z' A^{-1} z = \frac{1}{1 - \rho^2} z' \begin{bmatrix} 1 & -\rho & 0 & 0 & 0 & \cdots & \cdots & 0 \\ -\rho & 1 + \rho^2 & -\rho & 0 & 0 & \cdots & \cdots & 0 \\ 0 & -\rho & 1 + \rho^2 & -\rho & 0 & \cdots & \cdots & 0 \\ & & & \vdots & & & & \\ 0 & 0 & 0 & 0 & \cdots & -\rho & 1 + \rho^2 & -\rho \\ 0 & 0 & 0 & 0 & \cdots & 0 & -\rho & 1 \end{bmatrix} z$$

$$\begin{aligned}
&= z' \left(\frac{1-\rho}{1+\rho} \right) z + z' \frac{\rho}{2(1-\rho^2)} [(T-1) - (T-1)\rho] \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ -1 \end{bmatrix} \\
&= \frac{(1-\rho)^2}{1-\rho^2} \sum_{t=1}^T (t-\bar{t})^2 + \frac{\rho}{2(1-\rho^2)} [(T-1) - (T-1)\rho] [1-T],
\end{aligned}$$

$$z' a_T a_T' z = z' \begin{bmatrix} 1 \\ 1-\rho \\ \vdots \\ 1-\rho \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1-\rho & \cdots & 1-\rho & 1 \end{bmatrix} z$$

$$\begin{aligned}
&= (1-\bar{t}, 2-\bar{t}, \dots, T-1-\bar{t}, T-\bar{t}) \begin{bmatrix} 1 \\ 1-\rho \\ \vdots \\ 1-\rho \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1-\rho & \cdots & 1-\rho & 1 \end{bmatrix} \begin{bmatrix} 1-\bar{t} \\ 2-\bar{t} \\ \vdots \\ T-1-\bar{t} \\ T-\bar{t} \end{bmatrix} \\
&= \left[1-\bar{t} + T-\bar{t} + (1-\rho) \sum_{t=2}^{T-1} (t-\bar{t}) \right]^2 = 0,
\end{aligned}$$

$$\begin{aligned}
z' a_T a_T' y_i &= z' \begin{bmatrix} 1 \\ 1-\rho \\ \vdots \\ 1-\rho \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1-\rho & \cdots & 1-\rho & 1 \end{bmatrix} y_i \\
&= \left[1-\bar{t} + T-\bar{t} + (1-\rho) \sum_{t=2}^{T-1} (t-\bar{t}) \right] \left[y_{i1} + y_{iT} + (1-\rho) \sum_{t=2}^{T-1} y_{it} \right] \\
&= 0,
\end{aligned}$$

$$z' A^{-1} u_i = \frac{1}{1-\rho^2} \left\{ u_{i1} ((1-\bar{t}) - (2-\bar{t})\rho) + (1-\rho)^2 \sum_{t=2}^{T-1} (t-\bar{t}) u_{it} + u_{iT} [(T-\bar{t}) - (T-1-\bar{t})\rho] \right\}$$

$$= \frac{1}{1-\rho^2} \left\{ (\mu_i + v_{i1}) \left(\frac{-T+1}{2} - \frac{-T+3}{2} \rho \right) + (1-\rho)^2 \sum_{t=2}^{T-1} (t-\bar{t}) v_{it} + (\mu_i + v_{iT}) \left[\frac{T-1}{2} - \frac{T-3}{2} \rho \right] \right\},$$

and

$$\frac{(1-\rho)^2}{1-\rho^2} \sum_{t=1}^T (t-\bar{t})^2 \rightarrow \frac{1}{12} \frac{(1-\rho)^2}{1-\rho^2}.$$

It follows that

$$\begin{aligned} z' \Sigma^{-1} z &= \frac{1}{\sigma_v^2} z' \left(A^{-1} - \frac{\sigma_\mu^2}{\sigma_v^2 + \theta \sigma_\mu^2} a_T a_T' \right) z \\ &= \frac{1}{\sigma_v^2} z' A^{-1} z - \frac{1}{\sigma_v^2} \frac{\sigma_\mu^2}{\sigma_v^2 + \theta \sigma_\mu^2} z a_T a_T' z \\ &= \frac{1}{\sigma_v^2} \frac{(1-\rho)^2}{1-\rho^2} \sum_{t=1}^T (t-\bar{t})^2 + \frac{\rho}{2(1-\rho^2)} [(T-1) - (T-1)\rho] [1-T] \\ &= \frac{(1-\rho)^2}{\sigma_\varepsilon^2} \sum_{t=1}^T (t-\bar{t})^2 + \frac{\rho}{2(1-\rho^2)} [(T-1) - (T-1)\rho] [1-T], \\ \frac{1}{T^3} z' \Sigma^{-1} z &= \frac{1}{12} \frac{(1-\rho)^2}{\sigma_\varepsilon^2} + o(1), \end{aligned}$$

$$\begin{aligned} z' \Sigma^{-1} y_i &= \frac{1}{\sigma_v^2} z' \left(A^{-1} - \frac{\sigma_\mu^2}{\sigma_v^2 + \theta \sigma_\mu^2} a_T a_T' \right) y_i \\ &= \frac{1}{\sigma_v^2} z' A^{-1} y_i - \frac{1}{\sigma_v^2} \frac{\sigma_\mu^2}{\sigma_v^2 + \theta \sigma_\mu^2} z' a_T a_T' y_i \\ &= \frac{1}{\sigma_v^2} \frac{1}{1-\rho^2} \left\{ y_{i1} ((1-\bar{t}) - (2-\bar{t})\rho) + (1-\rho)^2 \sum_{t=2}^{T-1} (t-\bar{t}) y_{it} + y_{iT} [(T-\bar{t}) - (T-1-\bar{t})\rho] \right\}, \end{aligned}$$

$$\begin{aligned} z' \Sigma^{-1} u_i &= \frac{1}{\sigma_v^2} x' \left(A^{-1} - \frac{\sigma_\mu^2}{\sigma_v^2 + \theta \sigma_\mu^2} a_T a_T' \right) u_i \\ &= \frac{1}{\sigma_v^2} \frac{1}{1-\rho^2} \left\{ u_{i1} ((1-\bar{t}) - (2-\bar{t})\rho) + (1-\rho)^2 \sum_{t=2}^{T-1} (t-\bar{t}) u_{it} + u_{iT} [(T-\bar{t}) - (T-1-\bar{t})\rho] \right\} \\ &= \frac{1}{\sigma_v^2} \frac{1}{1-\rho^2} \left\{ (\mu_i + v_{i1}) \left(\frac{-T+1}{2} - \frac{-T+3}{2} \rho \right) + (1-\rho)^2 \sum_{t=2}^{T-1} (t-\bar{t}) v_{it} + (\mu_i + v_{iT}) \left[\frac{T-1}{2} - \frac{T-3}{2} \rho \right] \right\} \end{aligned}$$

and

$$\begin{aligned} \frac{1}{T^{3/2}} x' \Sigma^{-1} u_i &= \frac{1}{T^{3/2}} \frac{1}{\sigma_v^2} \frac{1}{1-\rho^2} (1-\rho)^2 \sum_{t=2}^{T-1} (t-\bar{t}) v_{it} + o_p(1) \\ &= \frac{1}{T^{3/2}} \frac{(1-\rho)^2}{\sigma_\varepsilon^2} \sum_{t=2}^{T-1} (t-\bar{t}) v_{it} + o_p(1). \end{aligned}$$

Hence,

$$\begin{aligned}
\sqrt{NT^3} \left(\hat{\beta}_{GLS} - \beta \right) &= \frac{\frac{1}{\sqrt{N}} \frac{1}{T^{3/2}} \sum_{i=1}^N z' \Sigma^{-1} u_i}{\frac{1}{N} \frac{1}{T^3} \sum_{i=1}^N z' \Sigma^{-1} z} \\
&= \frac{\frac{1}{\sqrt{N}} \frac{1}{T^{3/2}} \sum_{i=1}^N \frac{(1-\rho)^2}{\sigma_\varepsilon^2} \sum_{t=2}^{T-1} (t-\bar{t}) v_{it}}{\frac{1}{N} \sum_{i=1}^N \frac{1}{12} \frac{(1-\rho)^2}{\sigma_\varepsilon^2}} + o(1) \\
&= 12 \frac{1}{\sqrt{N}} \sum_{i=1}^N \frac{1}{T^{3/2}} \sum_{t=2}^{T-1} (t-\bar{t}) v_{it} + o(1) \xrightarrow{d} N \left(0, \frac{12\sigma_\varepsilon^2}{(1-\rho)^2} \right)
\end{aligned}$$

since

$$\frac{1}{T^{3/2}} \sum_{t=2}^{T-1} (t-\bar{t}) v_{it} \xrightarrow{d} N \left(0, \frac{\sigma_\varepsilon^2}{12(1-\rho)^2} \right)$$

proving (c). ■

D Proof of Theorem 4

Proof. If $\rho = 1$, then from Section 2 the GLS can be written as

$$\hat{\beta}_{GLS} = \frac{\sum_{i=1}^N z' \Sigma^{-1} y_i}{\sum_{i=1}^N z' \Sigma^{-1} z},$$

where

$$\begin{aligned}
z' \Sigma^{-1} z &= \frac{1}{\sigma_\varepsilon^2} \begin{bmatrix} 1-\bar{t} & 2-\bar{t} & 3-\bar{t} & \dots & T-\bar{t} \end{bmatrix} \begin{bmatrix} 2-\Delta & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \dots & 0 \\ 0 & -1 & 2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1-\bar{t} \\ 2-\bar{t} \\ 3-\bar{t} \\ \vdots \\ T-\bar{t} \end{bmatrix} \\
&= \frac{1}{\sigma_\varepsilon^2} \begin{bmatrix} -\bar{t} - (1-\bar{t})\Delta & 0 & 0 & \dots & 1 \end{bmatrix} \begin{bmatrix} 1-\bar{t} \\ 2-\bar{t} \\ 3-\bar{t} \\ \vdots \\ T-\bar{t} \end{bmatrix} \\
&= \frac{1}{\sigma_\varepsilon^2} \left[-\bar{t}(1-\bar{t}) - (1-\bar{t})^2 \Delta + T-\bar{t} \right] \\
&= \frac{1}{\sigma_\varepsilon^2} \left[\frac{T^2-1}{4} - \frac{(T-1)^2}{4} \Delta + \frac{T-1}{2} \right]
\end{aligned}$$

and

$$\begin{aligned}
z' \Sigma^{-1} y_i &= \frac{1}{\sigma_\varepsilon^2} \begin{bmatrix} 1 - \bar{t} & 2 - \bar{t} & 3 - \bar{t} & \cdots & T - \bar{t} \end{bmatrix} \begin{bmatrix} 2 - \Delta & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \cdots & 0 \\ 0 & -1 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \\ \vdots \\ y_{iT} \end{bmatrix} \\
&= \frac{1}{\sigma_\varepsilon^2} \begin{bmatrix} -\bar{t} & -(1 - \bar{t}) \Delta & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \\ \vdots \\ y_{iT} \end{bmatrix} \\
&= \frac{1}{\sigma_\varepsilon^2} [y_{iT} - (\bar{t} + (1 - \bar{t}) \Delta) y_{i1}] \\
&= \frac{1}{\sigma_\varepsilon^2} \left[y_{iT} - \left(\frac{T+1}{2} + \left(\frac{-T+1}{2} \right) \Delta \right) y_{i1} \right].
\end{aligned}$$

Then for $\kappa > 0$,

$$\begin{aligned}
\sqrt{NT} \left(\widehat{\beta}_{GLS} - \beta \right) &= \frac{\frac{1}{\sqrt{N}} \frac{1}{T^{3/2}} \sum_{i=1}^N z' \Sigma^{-1} u_i}{\frac{1}{N} \frac{1}{T^2} \sum_{i=1}^N z' \Sigma^{-1} z} \\
&= \frac{\frac{1}{\sqrt{N}} \frac{1}{T^{3/2}} \sum_{i=1}^N \frac{1}{\sigma_\varepsilon^2} [u_{iT} - \left(\frac{T+1}{2} + \left(\frac{-T+1}{2} \right) \Delta \right) u_{i1}]}{\frac{1}{N} \frac{1}{T^2} \sum_{i=1}^N \frac{1}{\sigma_\varepsilon^2} \left[\frac{T^2-1}{4} - \frac{(T-1)^2}{4} \Delta + \frac{T-1}{2} \right]} \\
&= \frac{\frac{1}{\sqrt{N}} \frac{1}{T} \frac{1}{\sqrt{T}} \sum_{i=1}^N [(\mu_i + v_{iT}) - \left(\frac{T+1}{2} + \left(\frac{-T+1}{2} \right) \Delta \right) (\mu_i + v_{i1})]}{\frac{1}{N} \sum_{i=1}^N \frac{1}{T^2} \left[\frac{T^2-1}{4} - \frac{(T-1)^2}{4} \Delta + \frac{T-1}{2} \right]} \\
&= \left(\frac{1}{4} - \frac{1}{4} \Delta \right)^{-1} \frac{1}{\sqrt{N}} \frac{1}{T} \frac{1}{\sqrt{T}} \sum_{i=1}^N \left[(\mu_i + v_{iT}) - \left(\frac{T+1}{2} + \left(\frac{-T+1}{2} \right) \Delta \right) (\mu_i + v_{i1}) \right] + o_p(1) \\
&= \left(\frac{1}{4} - \frac{1}{4} \Delta \right)^{-1} \frac{1}{\sqrt{N}} \frac{1}{T} \left[-\frac{T+1}{2} - \left(\frac{-T+1}{2} \right) \Delta \right] \left[\frac{1}{\sqrt{T}} \sum_{i=1}^N (\mu_i + v_{i1}) \right] + o_p(1) \\
&= \left(\frac{1}{4} - \frac{1}{4} \Delta \right)^{-1} \frac{1}{\sqrt{N}} \left(-\frac{1}{2} + \frac{1}{2} \Delta \right) \frac{1}{\sqrt{T}} \sum_{i=1}^N (\mu_i + v_{i1}) + o_p(1) \\
&= \left(\frac{1}{4} - \frac{1}{4} \Delta \right)^{-1} \frac{1}{\sqrt{N}} \left(-\frac{1}{2} + \frac{1}{2} \Delta \right) \frac{1}{\sqrt{T}} \sum_{i=1}^N v_{i1} + o_p(1).
\end{aligned}$$

It follows that for a fixed N we have

$$\left(\frac{1}{4} - \frac{1}{4} \Delta \right)^{-1} \frac{1}{\sqrt{N}} \left(-\frac{1}{2} + \frac{1}{2} \Delta \right) \sum_{i=1}^N \frac{1}{\sqrt{T}} v_{i1} \xrightarrow{d} -2 \frac{1}{\sqrt{N}} \sum_{i=1}^N \sigma_\varepsilon \widetilde{W}(\kappa)$$

and hence

$$\sqrt{NT} \left(\widehat{\beta}_{GLS} - \beta \right) \xrightarrow{d} N(0, 4\kappa\sigma_\varepsilon^2)$$

proving (c). Using the fact that $v_{i1} = O_p(1)$ if $\kappa = 0$ from Assumption 1. Then it is easy to see that

$$\begin{aligned} \sqrt{NT} \left(\widehat{\beta}_{GLS} - \beta \right) &= \frac{\frac{1}{\sqrt{N}} \frac{1}{T} \sum_{i=1}^N z' \Sigma^{-1} u_i}{\frac{1}{N} \frac{1}{T^2} \sum_{i=1}^N z' \Sigma^{-1} z} \\ &= \frac{\frac{1}{\sqrt{N}} \frac{1}{T} \sum_{i=1}^N \frac{1}{\sigma_\varepsilon^2} \left[u_{iT} - \left(\frac{T+1}{2} + \left(\frac{-T+1}{2} \right) \Delta \right) u_{i1} \right]}{\frac{1}{N} \frac{1}{T^2} \sum_{i=1}^N \frac{1}{\sigma_\varepsilon^2} \left[\frac{T^2-1}{4} - \frac{(T-1)^2}{4} \Delta + \frac{T-1}{2} \right]} \\ &= \frac{\frac{1}{\sqrt{N}} \frac{1}{T} \sum_{i=1}^N \left[(\mu_i + v_{iT}) - \left(\frac{T+1}{2} + \left(\frac{-T+1}{2} \right) \Delta \right) (\mu_i + v_{i1}) \right]}{\frac{1}{N} \sum_{i=1}^N \frac{1}{T^2} \left[\frac{T^2-1}{4} - \frac{(T-1)^2}{4} \Delta + \frac{T-1}{2} \right]} \\ &= \left(\frac{1}{4} - \frac{1}{4} \Delta \right)^{-1} \frac{1}{\sqrt{N}} \frac{1}{T} \sum_{i=1}^N \left[(\mu_i + v_{iT}) - \left(\frac{T+1}{2} + \left(\frac{-T+1}{2} \right) \Delta \right) (\mu_i + v_{i1}) \right] + o_p(1) \\ &= \left(\frac{1}{4} - \frac{1}{4} \Delta \right)^{-1} \frac{1}{\sqrt{N}} \frac{1}{T} \left[-\frac{T+1}{2} - \left(\frac{-T+1}{2} \right) \Delta \right] \left[\sum_{i=1}^N (\mu_i + v_{i1}) \right] + o_p(1) \\ &= \left(\frac{1}{4} - \frac{1}{4} \Delta \right)^{-1} \left(-\frac{1}{2} + \frac{1}{2} \Delta \right) \frac{1}{\sqrt{N}} \sum_{i=1}^N (\mu_i + v_{i1}) + o_p(1). \end{aligned}$$

It follows that

$$\left(\frac{1}{4} - \frac{1}{4} \Delta \right)^{-1} \left(-\frac{1}{2} + \frac{1}{2} \Delta \right) \frac{1}{\sqrt{N}} \sum_{i=1}^N (\mu_i + v_{i1}) \xrightarrow{d} N(0, 4(\sigma_\mu^2 + \sigma_\varepsilon^2))$$

proving (d). ■

E The Limiting Distribution of FE with $\rho = 1 + \frac{c}{T}$

We write the FE as

$$\begin{aligned} \sqrt{NT} \left(\widehat{\beta}_{FE} - \beta \right) &= \frac{\frac{1}{\sqrt{N}} \sum_{i=1}^N \frac{1}{\sqrt{T^5}} \sum_{t=1}^T (t - \bar{t}) u_{it}}{\frac{1}{N} \sum_{i=1}^N \frac{1}{T^3} \sum_{t=1}^T (t - \bar{t})^2} \\ &= \frac{\frac{1}{\sqrt{N}} \sum_{i=1}^N \frac{1}{\sqrt{T^5}} \sum_{t=1}^T (t - \bar{t}) (\mu_i + v_{it})}{\frac{1}{N} \sum_{i=1}^N \frac{1}{T^3} \sum_{t=1}^T (t - \bar{t})^2} \\ &= \frac{\frac{1}{\sqrt{N}} \sum_{i=1}^N \frac{1}{\sqrt{T^5}} \sum_{t=1}^T (t - \bar{t}) v_{it}}{\frac{1}{N} \sum_{i=1}^N \frac{1}{T^3} \sum_{t=1}^T (t - \bar{t})^2} \\ &= 12 \left\{ \frac{1}{\sqrt{N}} \sum_{i=1}^N \frac{1}{\sqrt{T^5}} \sum_{t=1}^T (t - \bar{t}) v_{it} \right\} + o_p(1). \end{aligned}$$

Now, for a fixed N we know from Canjels and Watson (1997)

$$\frac{1}{\sqrt{T^5}} \sum_{t=1}^T (t - \bar{t}) v_{it} \xrightarrow{d} \int_0^1 (s - \frac{1}{2}) [W_c(s) + e^{sc} \tilde{W}_c(\kappa)] ds$$

as $T \rightarrow \infty$, where $W_c(r)$ and $\tilde{W}_c(\kappa)$ solve $dW_c(r) = cW_c(r)dr + dW(r)$ and $d\tilde{W}_c(\kappa) = c\tilde{W}_c(\kappa)d\kappa + d\tilde{W}(\kappa)$, respectively, where $W(r)$ is a standard Brownian Motion and $\tilde{W}(r)$ is another standard Brownian Motion independent of $W(r)$.

Thus,

$$\sqrt{NT} (\hat{\beta}_{FE} - \beta) \xrightarrow{d} \frac{1}{\sqrt{N}} \sum_{i=1}^N \left\{ 12\sigma_\varepsilon \int_0^1 (s - \frac{1}{2}) [W_c(s) + e^{sc} \tilde{W}_c(\kappa)] ds \right\}.$$

However,

$$\left\{ 12\sigma_\varepsilon \int_0^1 (s - \frac{1}{2}) [W_c(s) + e^{sc} \tilde{W}_c(\kappa)] ds \right\} \sim N(0, R_1),$$

where $R_1 = A_1 + A_2$ with

$$A_1 = \text{var} \left\{ 12\sigma_\varepsilon \int_0^1 (s - \frac{1}{2}) W_c(s) ds \right\}$$

and

$$A_2 = \text{var} \left\{ 12\sigma_\varepsilon \tilde{W}_c(\kappa) \int_0^1 (s - \frac{1}{2}) e^{sc} \tilde{W}_c(\kappa) ds \right\}.$$

To compute A_1 , notice that

$$\begin{aligned} \int_0^1 (s - \frac{1}{2}) [W_c(s)] ds &= \int_0^1 (s - \frac{1}{2}) \int_0^S e^{c(s-\tau)} dW(\tau) ds \\ &= \int_0^1 \left[\int_\tau^1 (s - \frac{1}{2}) e^{cs} ds \right] e^{-c\tau} dW(\tau) \\ &= \int_0^1 b(\tau) dW(\tau). \end{aligned}$$

Thus

$$A_1 = 144\sigma_\varepsilon^2 \int_0^1 b(s)^2 ds = \sigma_\varepsilon^2 c^{-5} [18(c-2)^2 e^{2c} + 72c(c-2)e^c + 12c^3 + 54c^2 + 72c - 72]$$

$$A_2 = 144\sigma_\varepsilon^2 S_c(\kappa) \left[\int_0^1 (s - \frac{1}{2}) e^{cs} ds \right]^2 = 144\sigma_\varepsilon^2 S_c(\kappa) \left[\frac{ce^c + c - 2(e^c - 1)}{2c^2} \right].$$

Thus for a fixed N as $T \rightarrow \infty$,

$$\sqrt{NT} (\hat{\beta}_{FE} - \beta) \xrightarrow{d} \frac{1}{\sqrt{N}} \sum_{i=1}^N N(0, R_1).$$

Now let $N \rightarrow \infty$,

$$\sqrt{NT} (\hat{\beta}_{FE} - \beta) \xrightarrow{d} N(0, R_1).$$

■

References

- [1] Baltagi, B. H. (1981), "An Experimental Study of Alternative Testing and Estimation Procedures in a Two-Way Error Component Model," *Journal of Econometrics*, 17, 21-49.
- [2] Baltagi, B. H. (1989), "Applications of Necessary and Sufficient Condition for OLS to be BLUE," *Statistics & Probability Letters*, 8, 457-461.
- [3] Baltagi, B. H. (1995), *Econometric Analysis of Panel Data*, Wiley, New York.
- [4] Baltagi, B. H., and Chang, Y.-J. (1992), "Monte Carlo Evidence on Panel Data Regressions with AR(1) Disturbances and An Arbitrary Variance on the Initial Observations," *Journal of Econometrics*, 52, 371-380.
- [5] Baltagi, B. H., and Li, Q. (1991), "A Transformation That Will Circumvent the Problem of Autocorrelation in an Error Component Model," *Journal of Econometrics*, 52, 371-380.
- [6] Beach, B. M., and MacKinnon, J. G. (1978), "A Maximum Likelihood Procedure for Regression with Autocorrelation Errors," *Econometrica*, 46, 51-58.
- [7] Canjels, E., and Watson, M. M. (1997), "Estimating Deterministic Trends in the Presence of Serially Correlated Errors," *The Review of Economics and Statistics*, 79, 184-200.
- [8] Chipman, J. S. (1979), "Efficiency of Least-Squares Estimation of Linear Trend When Residuals are Autocorrelated," *Econometrica*, 47, 115-128.
- [9] Grenander, U. and Rosenblatt, M. (1957), *Statistical Analysis of Stationary Time Series*, Wiley, New York.
- [10] Hamilton, J. (1994), *Time Series Analysis*, Princeton University Press, Princeton.
- [11] Kao, C. (1997), "Spurious Regression and Residual-Based Tests for Cointegration in Panel Data," *Journal of Econometrics*, forthcoming.
- [12] Kao, C., and Chiang, M-H. (1997), "On the Estimation and Inference of a Cointegrated Regression in Panel Data," *Econometric Theory*, under revision.
- [13] Krämer, W. (1982), "Note on Estimating Linear Trend When Residuals are Autocorrelated," *Econometrica*, 50, 1065-1067.

- [14] Kruskal, W. (1968), "When are Gauss-Markov and Least Squares Estimators Identical? A Coordinate Free Approach," *Annals of Mathematical Statistics*, 39, 70-75.
- [15] Maeshiro, A. (1976), "Autoregressive Transformation, Trended Independent Variables and Autocorrelated Disturbance Terms," *The Review of Economics and Statistics*, 58, 497-500.
- [16] McCoskey, S., and Kao, C. (1998), "A Residual-Based Test of the Null of Cointegration in Panel Data," *Econometric Reviews*, 17, 57-84.
- [17] Park, S. J., and Mitchell, B. M. (1980), "Estimating the Autocorrelated Error Model with Trended Data," *Journal of Econometrics*, 13, 185-201.
- [18] Phillips, P. C. B., and Park, J. Y. (1988), "Asymptotic Equivalence of Ordinary Least Squares and Generalized Least Squares in Regressions with Integrated Regressors," *Journal of the American Statistical Association*, 83, 111-115.
- [19] Summers, R., and Heston, A. (1991), "The Penn World Table; An Expanded Set of International Comparisons 1950-1988," *Quarterly Journal of Economics*, 106, 327-368.
- [20] Vogelsang, T. J. (1998), "Trend Function Hypothesis Testing in the Presence of Serial Correlation," *Econometrica*, 66, 123-148.

Table 1: Relative Efficiencies of Estimators

		$\kappa = 0$					$\kappa = 0.1$				
Δ	ρ	OLS	FE	FD	GLS-CO	GLS-PW	OLS	FE	FD	GLS-CO	GLS-PW
0	0	1	.235	.054	.919	.956	1	.235	.054	.919	.956
	0.2	.997	.241	.079	.924	.961	.997	.239	.078	.923	.960
	0.4	.982	.250	.125	.932	.969	.982	.245	.116	.928	.966
	0.6	.950	.269	.217	.944	.982	.949	.255	.184	.935	.974
	0.8	.876	.343	.450	.957	.995	.871	.308	.355	.949	.988
	0.9	.813	.493	.709	.961	.999	.791	.442	.592	.955	.994
	0.95	.799	.654	.875	.969	1.007	.749	.617	.806	.953	.992
	1.0	.853	.831	.957			.803	.896	1.031		
0.2	0	.518	.683	.157	.917	.956	.518	.683	.157	.917	.956
	0.2	.623	.608	.201	.925	.964	.625	.606	.198	.926	.965
	0.4	.732	.516	.257	.934	.973	.738	.509	.241	.937	.977
	0.6	.818	.425	.343	.941	.979	.826	.407	.293	.946	.985
	0.8	.828	.405	.532	.946	.984	.825	.364	.420	.949	.988
	0.9	.785	.519	.746	.948	.985	.764	.465	.622	.949	.988
	0.95	.776	.663	.887	.955	.993	.729	.626	.818	.946	.984
	1.0	.845	.838	.964			.776	.879	1.013		
0.4	0	.268	.839	.193	.747	.779	.268	.839	.193	.747	.779
	0.2	.354	.779	.257	.779	.813	.355	.778	.254	.781	.815
	0.4	.469	.684	.341	.824	.860	.477	.680	.322	.834	.870
	0.6	.611	.559	.450	.872	.909	.629	.546	.393	.894	.932
	0.8	.731	.478	.627	.914	.952	.738	.434	.501	.934	.972
	0.9	.737	.552	.799	.929	.965	.721	.498	.667	.939	.998
	0.95	.741	.678	.907	.937	.973	.699	.639	.836	.936	.973
	1.0	.828	.845	.972			.747	.869	1.000		
0.6	0	.136	.919	.211	.465	.485	.136	.919	.211	.465	.485
	0.2	.188	.878	.289	.526	.549	.189	.878	.287	.528	.551
	0.4	.269	.799	.398	.614	.641	.276	.801	.379	.627	.654
	0.6	.394	.674	.543	.723	.754	.416	.673	.485	.759	.792
	0.8	.570	.559	.734	.841	.876	.592	.522	.603	.883	.920
	0.9	.647	.607	.873	.895	.931	.639	.548	.734	.919	.956
	0.95	.678	.703	.941	.910	.945	.643	.664	.868	.920	.957
	1.0	.790	.852	.981			.706	.863	.994		
0.8	0	.055	.967	.222	.191	.199	.055	.967	.222	.191	.199
	0.2	.078	.943	.311	.251	.261	.079	.943	.308	.252	.263
	0.4	.118	.883	.440	.344	.359	.122	.891	.422	.354	.370
	0.6	.189	.773	.622	.487	.509	.204	.790	.569	.525	.549
	0.8	.328	.645	.847	.690	.719	.359	.645	.734	.759	.793
	0.9	.454	.678	.975	.822	.856	.463	.627	.839	.864	.900
	0.95	.535	.749	1.003	.869	.904	.515	.711	.929	.893	.929
	1.0	.686	.859	.990			.616	.863	.993		

Note:

(a) $N = T = 25$.

(b) Relative efficiency is the ratio of the mean square error of the infeasible GLS estimator to the mean square error of the estimator given in row 2.

Table 2: Relative Efficiencies of Estimators

		$\kappa = 0.25$					$\kappa = 1$				
Δ	ρ	OLS	FE	FD	GLS-CO	GLS-PW	OLS	FE	FD	GLS-CO	GLS-PW
0	0	1	.235	.054	.919	.956	1	.235	.054	.919	.956
	0.2	.997	.240	.078	.923	.960	.997	.240	.078	.923	.960
	0.4	.982	.245	.116	.928	.966	.982	.245	.116	.928	.966
	0.6	.949	.253	.179	.935	.973	.949	.253	.179	.935	.973
	0.8	.869	.291	.316	.944	.983	.868	.288	.310	.943	.982
	0.9	.769	.398	.501	.953	.992	.752	.369	.451	.947	.986
	0.95	.689	.569	.721	.952	.991	.702	.499	.612	.950	.989
	1.0	.735	1.021	1.175			.599	1.657	1.907		
0.2	0	.518	.683	.157	.917	.956	.518	.683	.157	.917	.956
	0.2	.625	.606	.198	.926	.965	.625	.606	.198	.926	.965
	0.4	.738	.508	.240	.938	.977	.738	.508	.240	.938	.977
	0.6	.828	.404	.286	.947	.986	.827	.404	.286	.947	.986
	0.8	.825	.345	.374	.949	.988	.824	.341	.367	.949	.987
	0.9	.745	.418	.527	.951	.989	.728	.388	.474	.948	.987
	0.95	.674	.577	.731	.948	.987	.588	.506	.619	.950	.989
	1.0	.683	.959	1.105			.492	1.375	1.583		
0.4	0	.268	.839	.123	.747	.779	.268	.839	.193	.747	.779
	0.2	.355	.778	.254	.781	.815	.355	.778	.254	.781	.815
	0.4	.477	.680	.322	.834	.870	.477	.680	.322	.834	.870
	0.6	.633	.545	.386	.897	.935	.633	.545	.386	.897	.935
	0.8	.745	.415	.450	.942	.981	.744	.411	.442	.943	.982
	0.9	.707	.449	.566	.949	.987	.692	.417	.510	.949	.988
	0.95	.649	.590	.748	.944	.982	.569	.518	.634	.951	.989
	1.0	.637	.914	1.052			.407	1.151	1.325		
0.6	0	.136	.919	.211	.465	.485	.136	.919	.211	.465	.485
	0.2	.189	.878	.287	.528	.551	.189	.878	.287	.528	.551
	0.4	.276	.801	.379	.627	.655	.276	.801	.379	.627	.655
	0.6	.420	.674	.477	.764	.797	.419	.674	.477	.764	.797
	0.8	.607	.507	.551	.905	.943	.608	.503	.541	.908	.946
	0.9	.636	.499	.628	.939	.977	.626	.465	.567	.947	.985
	0.95	.604	.613	.777	.937	.975	.536	.538	.659	.951	.989
	1.0	.592	.883	1.016			.342	.989	1.139		
0.8	0	.079	.943	.308	.191	.263	.055	.967	.222	.191	.199
	0.2	.078	.943	.311	.252	.261	.078	.943	.308	.252	.263
	0.4	.122	.891	.421	.355	.371	.122	.891	.421	.355	.371
	0.6	.207	.794	.563	.531	.554	.207	.794	.562	.531	.554
	0.8	.379	.634	.689	.797	.832	.382	.632	.679	.804	.838
	0.9	.476	.586	.738	.903	.940	.479	.556	.677	.926	.964
	0.95	.496	.661	.838	.922	.959	.454	.585	.717	.949	.987
	1.0	.520	.867	.998			.289	.894	1.029		

Note:

(a) $N = T = 25$.

(b) Relative efficiency is the ratio of the mean square error of the infeasible GLS estimator to the mean square error of the estimator given in row 2.