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
Syracuse University. Center for Policy Research, cdkao@maxwell.syr.edu

Jamie Emerson

Center for Policy Research
Working Paper No. 1

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

Center for Policy Research
Maxwell School of Citizenship and Public Affairs
Syracuse University
Syracuse, New York 13244-1020

March 1999

$5.00

CENTER FOR POLICY RESEARCH - SPRING 1999

Timothy M. Smeeding, Director
Professor of Economics & Public Administration

Associate Directors

Margaret M. Austin
Associate Director,
Budget and Administration

Douglas Holtz-Eakin
Chair, Professor of Economics
Associate Director, Center for Policy Research

Douglas Wolf
Professor of Public Administration
Associate Director, Aging Studies Program

John Yinger
Professor of Economics and Public Administration
Associate Director, Metropolitan Studies Program

SENIOR RESEARCH ASSOCIATES

Stacy Dickert-Conlin.................................... Economics
William Duncombe.......................Public Administration
Thomas Dunn .... Economics
Vernon Greene ....Public Administration
David Greytak... Economics
Leah Gutierrez.............................Public Administration
Christine Himes.. Sociology
Jacqueline Johnson...................................... Sociology
Bernard Jump...............................Public Administration
Duke Kao..Economics
Jeff Kubik...Economics
Madonna Harrington Meyer..........................Sociology
Jerry Miner..Economics
Hisahiro Naito...Economics
Jan Ondrich...Economics
John Palmer................................Public Administration
Grant Reeher...Political Science
Jodi Sandfort ....Public Administration
Michael Wasylenko......................................Economics
Douglas Wolf...............................Public Administration
Assata Zerai ....Sociology

RESEARCH ASSOCIATES

Andrew Houtenville..Economics
Diane Macunovich..Economics

GRADUATE ASSOCIATES

Matthew Andrews........................Public Administration
Hakan Aykan................................Public Administration
Reagan Baughman ....Economics
Robert Bifulco...............................Public Administration
Don Bruce ....Economics
Christine Caffrey...Sociology
Noelle Chesley.............................Public Administration
Mohua Das...Economics
Deb Dougherty...Sociology
Seth Giertz..Economics
Tess Heintze ....Public Administration
Tim Hunter...................................Public Administration
Pam Herd ....Sociology
Garrett Ingoglia.............................Public Administration
Kwangho Jung ....Public Administration
James Laditka..............................Public Administration
Donald Marples..Economics
Christine Muller...............................Economics/Biology
Katherin Ross...Social Science
Shalini Sharma..Economics
Hava Tillipman...Sociology
Mehmet Serkan Tosun..............................Economics
Robert Weathers..Economics
Kristina Wolff..Sociology

STAFF

Martha Bonney.....................................Publications and Events Coordinator
Karen Cimilluca...................Librarian/Office Coordinator
Kati Foley..........................Administrative Assistant, LIS
Esther Gray ....Administrative Secretary
Detlef Jurkat .... Translator
Anne Metzger-Wormuth.............................Receptionist
Denise Paul.....................Administrative Assistant, NTJ
Annie Pennella............Editorial Assistant, Gerontologist
Mary Santy... Secretary
Debbie Tafel............................Secretary to the Director
Ann Wicks...................... Administrative Secretary
Lobrenzo Wingo................. Computer Consultant

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: February 16, 1999

Abstract

In this paper, we study the limiting distributions for ordinary least squares (OLS), fixed effects (FE), first difference (FD), and 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 the GLS. However, when the error term is I(1), the GLS could be less efficient than the FD or FE estimators and the 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. Monte Carlo experiments are employed to compare the performance of these estimators in finite samples. The main findings are: (1) the two-step 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 well for all values of ρ and Δ.

1 Introduction

In this paper we study the limiting distributions for the ordinary least squares (OLS), fixed effects (FE), first difference (FD), and 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, the FE model and the random-effect model. The 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 when the error terms are serially correlated (e.g., Baltagi and Li, 1991). However, as we show in Theorems 1-6, the GLS is the

*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 Hall, Syracuse University, Syracuse, NY 13244-1020; e-mail: cdkao@maxwell.syr.edu.

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 under which the OLS estimator is 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 the GLS and the OLS is less efficient than the GLS. However, when the error term is $I(1)$, the GLS could be less efficient than the FE and FD, and the 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.

Section 2 develops the asymptotic theory for OLS, FE, FD, and GLS estimators with an $I(0)$ error term. Section 3 gives the limiting distributions of OLS, FE, FD, and GLS estimators with an $I(1)$ error term. Section 4 discusses the effects of a fitted intercept. In Section 5, we discuss the asymptotics of the estimators when the error is nearly $I(1)$. Section 6 discusses the feasible GLS estimators. Section 7 presents Monte Carlo results to evaluate the finite sample properties of the proposed estimators. . In Section 8 we summarize the findings. All proofs are in the Appendix.

A word on notation. We use $\stackrel{d}{\to}$ to denote convergence in distribution, $\stackrel{p}{\to}$ to denote convergence in probability, $X \sim F$ to denote random variable X has distribution F, $[x]$ 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

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

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

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

where $\varepsilon_{it} \sim iid\left(0, \sigma_\varepsilon^2\right)$. The μ_i are assumed to be independent of v_{it} and $v_{it} \sim \left(0, \sigma_v^2\right)$, 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 $v_{i1} = \sum_{j=0}^{[\kappa T]} \rho^j \varepsilon_{i1-j.}$, *where κ is a parameter that governs the variance of the initial condition.*

Remark 1 *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 $I(1)$.*

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

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

$$\widehat{\beta}_{OLS} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T} t y_{it}}{\sum_{i=1}^{N}\sum_{t=1}^{T} t^2}, \tag{3}$$

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

and

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

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

$$y = \beta X + u \tag{6}$$

where y is $NT \times 1$, X is a vector of $x' = (1, 2, ..., T)$ 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

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

where

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

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

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

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

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 $1-6$ are taken as $T \to \infty$ followed by $N \to \infty$ 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 $\varepsilon_{it} \sim iid\left(0, \sigma_{\varepsilon}^2\right)$ and initial condition in Assumption 1. Then

(a) $\sqrt{N}T\left(\widehat{\beta}_{OLS} - \beta\right) \xrightarrow{d} \mathrm{N}\left(0, \frac{9}{4}\sigma_{\mu}^2\right)$,

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

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

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

Remark 2 *1. 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).*

2. 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, $\widehat{\beta}_B$:

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

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.

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

4. We expect that the equivalence results in Theorems 1-6 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).

5. Note that

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

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

$$A^{-1}x = \frac{1}{1-\rho^2}\begin{bmatrix} 1 & -\rho & 0 & 0 & 0 & \cdots & \cdots & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & 0 & \cdots & \cdots & 0 \\ 0 & -\rho & 1+\rho^2 & -\rho & 0 & \cdots & \cdots & 0 \\ & & & \vdots & & & & \\ 0 & 0 & 0 & 0 & \cdots & -\rho & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & 0 & \cdots & 0 & -\rho & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 3 \\ \vdots \\ T-1 \\ T \end{bmatrix}$$

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

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

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

Clearly $R\left(\Sigma^{-1}x\right) \neq R(x)$ *and hence* $R\left(\Omega^{-1}X\right) \neq R(X)$, *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 $\varepsilon_{it} = d(L)\epsilon_{it}$,*with* $d(L) = \sum_{j=0}^{\infty} d_j L^j$, $\sum_{j=0}^{\infty} j\,|d_j| < \infty$ *and* $\epsilon_{it} \sim iid\left(0, \sigma_\epsilon^2\right)$.

Then we have the 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) $\sqrt{NT}\left(\widehat{\beta}_{OLS} - \beta\right) \xrightarrow{d} \mathrm{N}\left(0, \frac{9}{4}\sigma_{\mu}^{2}\right)$,

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

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

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

Remark 3 *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 $\Sigma = \left(\sigma_{\varepsilon}^2 A + \sigma_{\mu}^2 \iota_T \iota_T^{'}\right)$ 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

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

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

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

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

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

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

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

where

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

Next,

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

where

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

$$= \frac{1}{\sigma_\varepsilon^2}\begin{bmatrix} -\Delta & 0 & 0 & \cdots & 1 \end{bmatrix}\begin{bmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \\ \vdots \\ y_{iT} \end{bmatrix} = \frac{1}{\sigma_\varepsilon^2}\left[y_{iT} - \Delta y_{i1}\right].$$

It follows that the GLS can be written as

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

Remark 4 *1. If* $\Delta = \frac{\sigma_\mu^2}{\sigma_\varepsilon^2+\sigma_\mu^2} = 1$*, then the GLS is reduced to the FD estimator.*

2. Since v_{it} *is* $I(1)$*, then for all* i*,*

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

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

The limiting distributions of the 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 $\varepsilon_{it} \sim iid\left(0, \sigma_\varepsilon^2\right)$ *and* $v_{i1} = \sum_{j=0}^{[\kappa T]}\varepsilon_{1-j}$*. Then*

(a) $\sqrt{NT}\left(\widehat{\beta}_{OLS} - \beta\right) \xrightarrow{d} \mathrm{N}\left(0, \left(\frac{6}{5} + \frac{9}{4}\kappa\right)\sigma_\varepsilon^2\right),$

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

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

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

Remark 5 *1. 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.*

2. *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 the FD or FE estimators. For example, the GLS is less efficient than the FD unless $\kappa=0$ or $\Delta=1$. Also the GLS is more efficient than the FE only if $(1-\Delta)^{2}\kappa<\frac{1}{5}$.*

3. *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.*

4. *The FD estimator is the most efficient estimator with $\rho=1$.*

5. *It can be shown that*

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

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

6. *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 (1998) 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 the 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

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

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 a stationary error term and a 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 $\varepsilon_{it} \sim iid\left(0, \sigma_\varepsilon^2\right)$ and initial condition in Assumption 1. Then

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

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

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

Remark 6 *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 $\varepsilon_{it}\sim iid\left(0,\sigma_\varepsilon^2\right)$ and initial condition in Assumption 1. Then

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

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

(c) $\sqrt{NT}\left(\widehat{\beta}_{GLS}-\beta\right)\xrightarrow{d}\mathrm{N}\left(0,\frac{4}{3}\sigma_\varepsilon^2\right).$

Remark 7 *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. However, the GLS is the least efficient estimator.*

5 Nearly I(1) Errors

In recent years, there has been considerable interest in the asymptotic properties of the estimation and inference of β in (1) when ρ is close to one in the time-series (i.e., when $N=1$) econometrics literature. In this section we assume $\rho=1+c/T$ in (2), i.e., the v_{it} follows a local-to-unit or a nearly $I(1)$ process. The asymptotics of the OLS, FE, FD, and GLS estimators are summarized in the following theorem:

Theorem 5 *Let y_{it} be generated from a simple time trend model in (1) where*

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

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

(a) $\sqrt{NT}\left(\widehat{\beta}_{OLS}-\beta\right)\xrightarrow{d}\mathrm{N}\left(0,\sigma_\varepsilon^2R_0\right),$

(b) $\sqrt{NT}\left(\widehat{\beta}_{FE}-\beta\right)\xrightarrow{d}\mathrm{N}(0,\sigma_\varepsilon^2R_1),$

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

(d) $\sqrt{NT}\left(\widehat{\beta}_{GLS}-\beta\right) \xrightarrow{d} \mathrm{N}\left(0, \frac{9\sigma_\varepsilon^2}{(c^2-3c+3)^2} R_2\right)$,

where

$$R_0 = \frac{\frac{9}{2}(c-1)^2 e^{2c} + 3c^3 + \frac{9}{2}c^2 - \frac{9}{2}}{c^5} + \frac{9}{2}\left(-1+e^{2c\kappa}\right)\frac{\left(ce^c - e^c + 1\right)^2}{c^5},$$

$R_2 = Var\left\{\left[W_c(1) - (1-e^c)\tilde{W}_c(\kappa)\right] + c^2\int_0^1 s\left[W_c(s) + e^{sc}\tilde{W}_c(\kappa)\right]ds - c\left[W_c(1) + e^c\tilde{W}_c(\kappa)\right]\right\}$, *and* R_1 *is given in the Appendix E.*

It is clear that in the limiting distributions the FE and FD estimators are the same with and without intercept. The asymptotics of the FE, FD, and GLS estimators when there is an intercept are stated in the following theorem:

Theorem 6 *Let* y_{it} *be generated from a simple time trend model in (9) where*

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

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

(a) $\sqrt{NT}\left(\widehat{\beta}_{FE}-\beta\right) \xrightarrow{d} \mathrm{N}(0, \sigma_\varepsilon^2 R_1)$,

(b) $\sqrt{NT}\left(\widehat{\beta}_{FD}-\beta\right) \xrightarrow{d} \mathrm{N}\left(0, \sigma_\varepsilon^2\left[S_c(1) + (1-e^c)S_c(\kappa)\right]\right)$,

(c) $\sqrt{NT}\left(\widehat{\beta}_{GLS}-\beta\right) \xrightarrow{d} N\left(0, \sigma_\varepsilon^2\left(\frac{12}{c^2+12-6c}\right)^2 R_3\right)$,

where

$$R_3 = Var\left\{\left(\frac{2-c}{2}\right)\left(W_c(1) - (1-e^c)\tilde{W}_c(\kappa)\right) + c^2\int_0^1\left(s-\frac{1}{2}\right)\left[W_c(s) + e^{sc}\tilde{W}_c(\kappa)\right]ds\right\}.$$

Remark 8 *1. Note*

$$\begin{aligned}\lim_{c\to 0} R_0 &= \frac{9}{4}\frac{4c^3e^{2c} - 14c^2e^{2c} + 22ce^{2c} - 11e^{2c} - 5 + 16e^c - 16ce^c}{c^5} + \frac{9}{2}\left(-1+e^{2c\kappa}\right)\frac{\left(ce^c - e^c + 1\right)^2}{c^5} \\ &= \frac{6}{5} + \frac{9}{4}\kappa,\end{aligned}$$

$$\begin{aligned}\lim_{c\to 0} R_1 &= \lim_{c\to 0}(A_1 + A_2) = \left(\begin{array}{c} c^{-5}\left[18(c-2)^2e^{2c} + 72c(c-2)e^c + 12c^3 + 54c^2 + 72c - 72\right] \\ +144 S_c(\kappa)\left[\dfrac{ce^c + c - 2(e^c-1)}{2c^2}\right]\end{array}\right) \\ &= \frac{6}{5},\end{aligned}$$

and

$$\lim_{c\to 0} R_3 = 1. \tag{10}$$

2. Note from (d) of Theorem 5 and (c) of Theorem 6 that as $c \to 0$ we have

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

which is different from (d) in Theorem 2 and (c) in Theorem 4. It is because that the limit of letting c go to zero and then follows by a large T may not be the same by increasing T first and follows by letting c go to zero.

6 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.,

$$\widehat{\rho} = \frac{\sum_{i=1}^{N}\sum_{t=2}^{T}\widehat{u}_{it}\widehat{u}_{it-1}}{\sum_{i=1}^{N}\sum_{t=2}^{T}\left(\widehat{u}_{it-1}\right)^2}, \tag{11}$$

where $\widehat{u}_{it}$ is the estimated residual, taken from the FE estimation of the model in (1). It can be shown that $\widehat{\rho}$ in (11) 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 the 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 a 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 $I(1)$.

7 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 provides a useful guide for choosing among the estimators in small samples. The model is set as follows:

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

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

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

$\varepsilon_{it} \overset{iid}{\sim} N\left(0, \sigma_\varepsilon^2\right)$ 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 an 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. We only report the results for the model without intercept since the results for the model with intercept are similar to the model without intercept.

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 then 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 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 $\kappa > 0.1$, though the better performance of the FD estimator decreases as Δ increases. Overall, the FE estimator performs relatively well when compared to other methods.

8 Conclusion

This paper considers a linear time trend model with a one-way error component with a 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 well for all values of ρ and Δ.

Appendix

A Proof of Theorem 1

Proof. Equation (3) can be expressed as

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

We multiply (12) by $\sqrt{NT^3}$, resulting in

$$\sqrt{NT^3}\left(\widehat{\beta}_{FE} - \beta\right) = \frac{\frac{1}{\sqrt{NT^3}}\sum_{i=1}^{N}\sum_{t=1}^{T}(t-\bar{t})u_{it}}{\frac{1}{NT^3}\sum_{i=1}^{N}\sum_{t=1}^{T}(t-\bar{t})^2}. \tag{13}$$

It is straightforward to show that

$$\begin{aligned}\sum_{i=1}^{N}\sum_{t=1}^{T}(t-\bar{t})^2 &= \frac{NT\left(T^2-1\right)}{12} \\ &= O\left(N^{-1}T^3\right).\end{aligned} \tag{14}$$

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

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

We turn next to the numerator in (13). 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

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

and

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

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

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

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

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

and

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

Obviously

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

as $N\to\infty$ (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

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

Then for a fixed N, we obtain

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

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

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

Hence,

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

as $N \to \infty$. Finally,

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

proving (a). Note that

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

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

$$
\begin{aligned}
x^{'}A^{-1}y_i &= \frac{1}{1-\rho^2}x^{'}\begin{bmatrix} 1 & -\rho & 0 & 0 & 0 & \cdots & \cdots & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & 0 & \cdots & \cdots & 0 \\ 0 & -\rho & 1+\rho^2 & -\rho & 0 & \cdots & \cdots & 0 \\ & & & \vdots & & & & \\ 0 & 0 & 0 & 0 & \cdots & -\rho & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & 0 & \cdots & 0 & -\rho & 1 \end{bmatrix}\begin{bmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \\ \vdots \\ y_{iT-1} \\ y_{iT} \end{bmatrix} \\
&= \frac{1}{1-\rho^2}\begin{bmatrix} 1-2\rho & 2\left(1-\rho\right)^2 & 3\left(1-\rho\right)^2 & \cdots & \left(T-1\right)\left(1-\rho\right)^2 & T-\left(T-1\right)\rho \end{bmatrix}\begin{bmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \\ \vdots \\ y_{iT-1} \\ y_{iT} \end{bmatrix} \\
&= \frac{1}{1-\rho^2}\left\{y_{i1}\left(1-2\rho\right)+y_{i2}2\left(1-\rho\right)^2+y_{i3}3\left(1-\rho\right)^2+\cdots+y_{iT}\left[T-\left(T-1\right)\rho\right]\right\} \\
&= \frac{1}{1-\rho^2}\left\{y_{i1}\left(1-2\rho\right)+\left(1-\rho\right)^2\sum_{t=2}^{T-1}ty_{it}+y_{iT}\left[T-\left(T-1\right)\rho\right]\right\},
\end{aligned}
$$

and

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

Then

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

$$-\frac{1}{\sigma_v^2}\frac{\sigma_\mu^2}{\sigma_v^2+\theta\sigma_\mu^2}\frac{1}{(1+\rho)^2}\left[1+(1-\rho)\left(\sum_{t=2}^{T-1}t\right)+T\right]^2,$$

$$\begin{aligned}
\frac{1}{T^3}x'\Sigma^{-1}x &= \frac{1}{3}\frac{(1-\rho)^2}{\sigma_\varepsilon^2}-\frac{1}{4}\frac{(1-\rho)^2}{\sigma_\varepsilon^2}\frac{1-\rho^2}{(1+\rho)^2}\frac{1+\rho}{1-\rho}+o(1)\\
&= \frac{1}{3}\frac{(1-\rho)^2}{\sigma_\varepsilon^2}-\frac{1}{4}\frac{(1-\rho)^2}{\sigma_\varepsilon^2}+o(1)\\
&= \frac{1}{12}\frac{(1-\rho)^2}{\sigma_\varepsilon^2}+o(1),
\end{aligned}$$

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

$$\begin{aligned}
x'\Sigma^{-1}u_i &= \frac{1}{\sigma_v^2}x'\left(A^{-1}-\frac{\sigma_\mu^2}{\sigma_v^2+\theta\sigma_\mu^2}a_Ta_T'\right)u_i\\
&= \frac{1}{\sigma_\varepsilon^2}\left\{u_{i1}(1-2\rho)+(1-\rho)^2\sum_{t=2}^{T-1}tu_{it}+u_{iT}\left[T-(T-1)\rho\right]\right\}\\
&\quad -\frac{1}{\sigma_v^2}\frac{\sigma_\mu^2}{\sigma_v^2+\theta\sigma_\mu^2}\left(\frac{1}{1+\rho}\right)^2\left[1+T+(1-\rho)\sum_{t=2}^{T-1}t\right]\left[u_{i1}+(1-\rho)\sum_{t=2}^{T-1}u_{it}+u_{iT}\right]\\
&= \frac{1}{\sigma_\varepsilon^2}\left\{(\mu_i+v_{i1})(1-2\rho)+(1-\rho)^2\sum_{t=2}^{T-1}t(\mu_i+v_{it})+(\mu_i+v_{iT})\left[T-(T-1)\rho\right]\right\}\\
&\quad -\frac{1}{\sigma_v^2}\frac{\sigma_\mu^2}{\sigma_v^2+\theta\sigma_\mu^2}\left(\frac{1}{1+\rho}\right)^2\\
&\quad \left[1+T+(1-\rho)\sum_{t=2}^{T-1}t\right]\left[(\mu_i+v_{i1})+(1-\rho)\sum_{t=2}^{T-1}(\mu_i+v_{it})+(\mu_i+v_{iT})\right], \qquad (15)
\end{aligned}$$

and

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

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

since

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

$$
\lim_{T\to\infty}\frac{\sigma_\mu^2}{\sigma_v^2+\theta\sigma_\mu^2}(T-2)=\frac{1+\rho}{1-\rho},
$$

and

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

It is known that

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

and then

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

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

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

and

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

Then

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

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

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

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

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

Then

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

and hence

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

proving (c). ■

B Proof of Theorem 2

Proof. 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 motions. We write the FE as

$$\begin{aligned} \sqrt{NT}\left(\widehat{\beta}_{FE} - \beta\right) &= \frac{\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{1}{\sqrt{T^5}} \sum_{t=1}^{T} (t - \bar{t}) u_{it}}{\frac{1}{N} \sum_{i=1}^{N} \frac{1}{T^3} \sum_{t=1}^{T} (t - \bar{t})^2} \\ &= \frac{\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{1}{\sqrt{T^5}} \sum_{t=1}^{T} (t - \bar{t}) v_{it}}{\frac{1}{N} \sum_{i=1}^{N} \frac{1}{T^3} \sum_{t=1}^{T} (t - \bar{t})^2}. \end{aligned} \tag{16}$$

Let us begin with the numerator in (16). Using Proposition 17.1 in Hamilton (1994), we see that for a fixed N as $T \to \infty$

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

Then

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

as $N \to \infty$. Thus

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

proving (b). Similarly,

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

It follows that

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

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

$$\begin{aligned}\sqrt{NT}\left(\widehat{\beta}_{OLS}-\beta\right) &= 3\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{1}{\sqrt{T^5}}\sum_{t=1}^{T}tv_{it}+o_p(1)\\ &\xrightarrow{d} \mathrm{N}\left(0,\left(\frac{6}{5}+\frac{9}{4}\kappa\right)\sigma_{\varepsilon}^2\right)\end{aligned}$$

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

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

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

$$\frac{1}{T}\left(T-\Delta\right)\to 1$$

and

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

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

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

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

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

proving (c). ■

C Proof of Theorem 3

Proof. It will be convenient to center the observation so that $z = t - \overline{t}$ and

$$y_i = \begin{bmatrix} y_{i1} - \overline{y}_i \\ y_{i2} - \overline{y}_i \\ y_{i3} - \overline{y}_i \\ \vdots \\ y_{iT} - \overline{y}_i \end{bmatrix}.$$

From the proof of Theorem 1 we know

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

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

$$= \left(1-\bar{t}, 2-\bar{t}, \cdots, T-1-\bar{t}, T-\bar{t}\right) \begin{bmatrix} 1 \\ 1-\rho \\ \vdots \\ 1-\rho \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1-\rho & \cdots & 1-\rho & 1 \end{bmatrix} \begin{bmatrix} 1-\bar{t} \\ 2-\bar{t} \\ \vdots \\ T-1-\bar{t} \\ T-\bar{t} \end{bmatrix}$$

$$= \left[1-\bar{t}+T-\bar{t}+(1-\rho)\sum_{t=2}^{T-1}\left(t-\bar{t}\right)\right]^2 = 0,$$

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

$$= \left[1-\bar{t}+T-\bar{t}+(1-\rho)\sum_{t=2}^{T-1}\left(t-\bar{t}\right)\right]\left[y_{i1}-\overline{y}_i+y_{iT}-\overline{y}_i+(1-\rho)\sum_{t=2}^{T-1}\left(y_{it}-\overline{y}_i\right)\right]$$

$$= 0,$$

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

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

and

$$\frac{(1-\rho)^2}{1-\rho^2}\frac{1}{T^3}\sum_{t=1}^{T}\left(t-\bar{t}\right)^2 \to \frac{1}{12}\frac{(1-\rho)^2}{1-\rho^2}.$$

It follows that

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

$$= \frac{1}{\sigma_v^2}z^{'}A^{-1}z-\frac{1}{\sigma_v^2}\frac{\sigma_\mu^2}{\sigma_v^2+\theta\sigma_\mu^2}z a_T a_T^{'} z$$

$$= \frac{1}{\sigma_v^2}\left\{\frac{(1-\rho)^2}{1-\rho^2}\sum_{t=1}^{T}\left(t-\bar{t}\right)^2+\frac{\rho}{2\left(1-\rho^2\right)}\left[(T+1)-(T-1)\rho\right]\left[T-1\right]\right\}$$

$$= \frac{(1-\rho)^2}{\sigma_\varepsilon^2}\sum_{t=1}^{T}\left(t-\bar{t}\right)^2+\frac{\rho}{2\sigma_\varepsilon^2}\left[(T+1)-(T-1)\rho\right]\left[T-1\right],$$

$$\frac{1}{T^3} z^{'} \Sigma^{-1} z = \frac{1}{12} \frac{(1-\rho)^2}{\sigma_\varepsilon^2} + o(1),$$

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

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

and

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

Hence,

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

since

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

proving (c). ■

D Proof of Theorem 4

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

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

where

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

and

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

$$= \frac{1}{\sigma_\varepsilon^2}\left[(y_{iT}-\overline{y}_i)-\left(\overline{t}+\left(1-\overline{t}\right)\Delta\right)(y_{i1}-\overline{y}_i)\right]$$

$$= \frac{1}{\sigma_\varepsilon^2}\left[(y_{iT}-\overline{y}_i)-\left(\frac{T+1}{2}+\left(\frac{-T+1}{2}\right)\Delta\right)(y_{i1}-\overline{y}_i)\right].$$

Then

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

It follows that for a fixed N we have

$$\left(\frac{1}{4}-\frac{1}{4}\Delta\right)^{-1}\frac{1}{\sqrt{N}}\left(-\frac{1}{2}+\frac{1}{2}\Delta\right)\sum_{i=1}^N\frac{1}{\sqrt{T}}(v_{i1}-\overline{v}_i)\xrightarrow{d}-2\frac{1}{\sqrt{N}}\sum_{i=1}^N\sigma_\varepsilon\left(\widetilde{W}(\kappa)-\int_0^1\left[W(s)+\widetilde{W}(\kappa)\right]ds\right)$$

and hence

$$\sqrt{NT}\left(\widehat{\beta}_{GLS}-\beta\right)\xrightarrow{d}\mathrm{N}\left(0,\frac{4}{3}\sigma_\varepsilon^2\right)$$

proving (c). ■

E Proof of Theorem 5

Proof. The approach we follow is based on Canjels and Watson (1997). To prove (a) we first note

$$\begin{aligned}
\sqrt{NT}\left(\widehat{\beta}_{OLS}-\beta\right) &= \frac{\frac{1}{\sqrt{N}}\sum_{i=1}^N\frac{1}{T^{5/2}}\sum_{t=1}^T tu_{it}}{\frac{1}{N}\sum_{i=1}^N\frac{1}{T^3}\sum_{t=1}^T t^2} \\
&= \frac{\frac{1}{\sqrt{N}}\sum_{i=1}^N\frac{1}{T^{5/2}}\sum_{t=1}^T t\left(\mu_i+v_{it}\right)}{\frac{1}{N}\sum_{i=1}^N\frac{1}{T^3}\sum_{t=1}^T t^2} \\
&= \frac{\frac{1}{\sqrt{N}}\sum_{i=1}^N\frac{1}{T^{5/2}}\left[\mu_i\sum_{t=1}^T t+\sum_{t=1}^T tv_{it}\right]}{\frac{1}{N}\sum_{i=1}^N\frac{1}{T^3}\sum_{t=1}^T t^2}
\end{aligned}$$

$$= \quad 3\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{1}{T^{5/2}}\left[\mu_i\sum_{t=1}^{T}t+\sum_{t=1}^{T}tv_{it}\right]+o_p(1).$$

Then, for a fixed N,

$$\begin{aligned}
& 3\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{1}{T^{5/2}}\left[\mu_i\sum_{t=1}^{T}t+\sum_{t=1}^{T}tv_{it}\right] \\
= & \quad 3\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{1}{T^{3/2}}\left[\mu_i\frac{\sum_{t=1}^{T}t}{T}+\sum_{t=1}^{T}\left(\frac{t}{T}\right)v_{it}\right] \\
= & \quad 3\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{1}{T^{3/2}}\left[\sum_{t=1}^{T}\left(\frac{t}{T}\right)v_{it}\right]+o_p(1) \\
& \xrightarrow{d} \frac{1}{\sqrt{N}}\sigma_\varepsilon\sum_{i=1}^{N}3\int_0^1 s\left[W_c(s)+e^{sc}\tilde{W}_c(\kappa)\right]ds
\end{aligned}$$

as $T\to\infty$. Note

$$3\int_0^1 s\left[W_c(s)+e^{sc}\tilde{W}_c(\kappa)\right]ds\sim N\left(0,R_0\right),$$

where

$$R_0=Var\left[3\int_0^1 sW_c(s)ds\right]+Var\left[\tilde{W}_c(\kappa)3\int_0^1 se^{sc}ds\right].$$

It is easy to see that

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

Thus

$$\begin{aligned}
Var\left[3\int_0^1 sW_c(s)ds\right] &= 9\int_0^1\left(\int_\tau^1 se^{cs}ds\right)^2 d\tau \\
&= \frac{9}{4}\frac{4c^3e^{2c}-14c^2e^{2c}+22ce^{2c}-11e^{2c}-5+16e^c-16ce^c}{c^5}
\end{aligned}$$

and

$$\begin{aligned}
Var\left[\tilde{W}_c(\kappa)3\int_0^1 se^{sc}ds\right] &= 9S_c\left(\kappa\right)\left[\int_0^1 se^{sc}ds\right]^2 \\
&= \frac{9}{2}\left(-1+e^{2c\kappa}\right)\frac{\left(ce^c-e^c+1\right)^2}{c^5}.
\end{aligned}$$

It follows that

$$R_0=\frac{9}{4}\frac{4c^3e^{2c}-14c^2e^{2c}+22ce^{2c}-11e^{2c}-5+16e^c-16ce^c}{c^5}+\frac{9}{2}\left(-1+e^{2c\kappa}\right)\frac{\left(ce^c-e^c+1\right)^2}{c^5}.$$

By the Lindeberg-Levy central limit theorem, as $N \to \infty$

$$\sqrt{NT}\left(\widehat{\beta}_{OLS} - \beta\right) \xrightarrow{d} \mathrm{N}\left(0, \sigma_{\varepsilon}^{2} R_0\right).$$

This proves (a). We write the FE as

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

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

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

as $T \to \infty$, 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,

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

However,

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

where $R_1 = A_1 + A_2$ with

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

and

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

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

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

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

where

$$S_c(\kappa) = \frac{1}{-2c}\left(1 - e^{2c\kappa}\right).$$

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

$$\sqrt{NT}\left(\widehat{\beta}_{FE} - \beta\right) \xrightarrow{d} \frac{1}{\sqrt{N}}\sum_{i=1}^N \mathrm{N}(0, \sigma_\varepsilon^2 R_1).$$

Now let $N \to \infty$,

$$\sqrt{NT}\left(\widehat{\beta}_{FE} - \beta\right) \xrightarrow{d} \mathrm{N}(0, \sigma_\varepsilon^2 R_1)$$

proving (b). To prove (c) note for a fixed N

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

as $T \to \infty$. Then

$$\frac{1}{\sqrt{N}}\sum_{i=1}^N \mathrm{N}\left(0, S_c(1) + (1-e^c)S_c(\kappa)\right) \xrightarrow{d} \mathrm{N}\left(0, \sigma_\varepsilon^2\left[S_c(1) + (1-e^c)\,S_c(\kappa)\right]\right)$$

as $N \to \infty$. Hence

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

proving (c). To prove (d) we first write the GLS as

$$\sqrt{NT}\left(\widehat{\beta}_{GLS} - \beta\right) = \frac{\frac{1}{\sqrt{N}}\frac{1}{T^{3/2}}\sum_{i=1}^N x'\Sigma^{-1}u_i}{\frac{1}{N}\frac{1}{T^2}\sum_{i=1}^N x'\Sigma^{-1}x}$$

where

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

and

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

We first find the limiting distributions of $x^{'}\Sigma^{-1}x$ and $x^{'}\Sigma^{-1}u_i$. For a fixed N as $T\to\infty$, we have

$$
\lim_{T\to\infty}\frac{1}{T}\left\{\frac{(1-\rho)^2}{\sigma_\varepsilon^2}\sum_{t=1}^{T}t^2+\frac{1}{\sigma_\varepsilon^2}\left[-\rho^2+T\rho(T-T\rho+1)\right]\right\}=\frac{1}{3}\frac{c^2-3c+3}{\sigma_\varepsilon^2},
$$

and

$$
\lim_{T\to\infty}\frac{1}{T}\frac{1}{\sigma_v^2}\frac{\sigma_\mu^2}{\sigma_v^2+\theta\sigma_\mu^2}\frac{1}{(1+\rho)^2}\left[1+(1-\rho)\left(\sum_{t=2}^{T-1}t\right)+T\right]^2=0.
$$

It follows that

$$
\frac{1}{T}x^{'}\Sigma^{-1}x\to\frac{1}{3}\frac{c^2-3c+3}{\sigma_\varepsilon^2}.
$$

Similarly,

$$
\begin{aligned}
\frac{1}{T^{1/2}}x^{'}\Sigma^{-1}u_i &= \frac{1}{T^{1/2}}(\mu_i+v_{i1})\frac{1}{\sigma_\varepsilon^2}(1-2\rho) \\
&\quad +T^2\frac{1}{\sigma_\varepsilon^2}(1-\rho)^2\frac{1}{T^{5/2}}\sum_{t=2}^{T-1}t(\mu_i+v_{it}) \\
&\quad +\frac{1}{T^{1/2}}(\mu_i+v_{iT})\frac{1}{\sigma_\varepsilon^2}[T-(T-1)\rho] \\
&\quad -\frac{1-\rho^2}{\sigma_\varepsilon^2}\frac{\sigma_\mu^2}{\frac{\sigma_\varepsilon^2}{1-\rho^2}+\theta\sigma_\mu^2}\left(\frac{1}{1+\rho}\right)^2 \\
&\quad \frac{1}{T^{1/2}}\left[1+T+(1-\rho)\sum_{t=2}^{T-1}t\right]\left[(\mu_i+v_{i1})+(1-\rho)\sum_{t=2}^{T-1}(\mu_i+v_{it})+(\mu_i+v_{iT})\right] \\
&= -\frac{1}{\sigma_\varepsilon^2}\frac{1}{T^{1/2}}(\mu_i+v_{i1})+\frac{c^2}{\sigma_\varepsilon^2}\frac{1}{T^{5/2}}\sum_{t=2}^{T-1}t(\mu_i+v_{it})+\frac{(1-c)}{\sigma_\varepsilon^2}\frac{1}{T^{1/2}}(\mu_i+v_{iT})
\end{aligned}
$$

$$
\begin{aligned}
& -\frac{1}{\sqrt{T}} v_{i1} \frac{1-\rho^2}{\sigma_\varepsilon^2} \frac{\sigma_\mu^2}{\frac{\sigma_\varepsilon^2}{1-\rho^2}+\theta\sigma_\mu^2} \left(\frac{1}{1+\rho}\right)^2 \left[1+T+(1-\rho)\sum_{t=2}^{T-1} t\right] \\
& -\frac{1-\rho^2}{\sigma_\varepsilon^2} \frac{\sigma_\mu^2}{\frac{\sigma_\varepsilon^2}{1-\rho^2}+\theta\sigma_\mu^2} \left(\frac{1}{1+\rho}\right)^2 \left[1+T+(1-\rho)\sum_{t=2}^{T-1} t\right] \left((1-\rho)\frac{1}{T^{1/2}}\sum_{t=2}^{T-1} v_{it}\right) \\
& -\frac{1}{\sqrt{T}} v_{iT} \frac{1-\rho^2}{\sigma_\varepsilon^2} \frac{\sigma_\mu^2}{\frac{\sigma_\varepsilon^2}{1-\rho^2}+\theta\sigma_\mu^2} \left(\frac{1}{1+\rho}\right)^2 \left[1+T+(1-\rho)\sum_{t=2}^{T-1} t\right] + o_p(1) \\
= & -\frac{1}{\sigma_\varepsilon^2}\frac{1}{T^{1/2}} v_{i1} + \frac{c^2}{\sigma_\varepsilon^2}\frac{1}{T^{5/2}}\sum_{t=2}^{T-1} t v_{it} + \frac{(1-c)}{\sigma_\varepsilon^2}\frac{1}{T^{1/2}} v_{iT} + o_p(1) \\
= & \frac{1}{\sigma_\varepsilon^2}\left[\frac{(1-c)}{T^{1/2}} v_{iT} - \frac{1}{T^{1/2}} v_{i1} + \frac{c^2}{T^{5/2}}\sum_{t=2}^{T-1} t v_{it}\right] + o_p(1),
\end{aligned}
$$

as $T \to \infty$ using

$$
\lim_{T\to\infty} \frac{1-\rho^2}{\sigma_\varepsilon^2} \frac{\sigma_\mu^2}{\frac{\sigma_\varepsilon^2}{1-\rho^2}+\theta\sigma_\mu^2} \left(\frac{1}{1+\rho}\right)^2 \left[1+T+(1-\rho)\sum_{t=2}^{T-1} t\right] = 0.
$$

It follows that

$$
\begin{aligned}
& \frac{1}{T^{1/2}} x^{'}\Sigma^{-1} u_i \\
& \xrightarrow{d} \frac{1}{\sigma_\varepsilon}\left\{\left[W_c(1) - (1-e^c)\tilde{W}_c(\kappa)\right] + c^2 \int_0^1 s\left[W_c(s) + e^{sc}\tilde{W}_c(\kappa)\right] ds - c\left[W_c(1) + e^c \tilde{W}_c(\kappa)\right]\right\}
\end{aligned}
$$

using

$$
\begin{aligned}
\frac{1}{\sqrt{T}}(v_{iT} - v_{i1}) &= \frac{1}{\sqrt{T}}\left(\widetilde{v}_{iT} + \rho^{T-1} v_{i1} - v_{i1}\right) \\
&= \frac{1}{\sqrt{T}}\left(\widetilde{v}_{iT} - \left(1-\rho^{T-1}\right) v_{i1}\right) \\
&\xrightarrow{d} \sigma_\varepsilon \left[W_c(1) - (1-e^c) W_c(\kappa)\right]
\end{aligned}
$$

and

$$
\frac{1}{T^{5/2}}\sum_{t=2}^{T-1} t v_{it} \to^d \sigma_\varepsilon \int_0^1 s\left[W_c(s) + e^{sc}\tilde{W}_c(\kappa)\right] ds.
$$

It follows that

$$
\begin{aligned}
\sqrt{NT}\left(\widehat{\beta}_{GLS} - \beta\right) &= \frac{\frac{1}{\sqrt{N}}\frac{1}{\sqrt{T}}\sum_{i=1}^N x^{'}\Sigma^{-1} u_i}{\frac{1}{N}\frac{1}{T}\sum_{i=1}^N x^{'}\Sigma^{-1} x} \\
&\xrightarrow{d} \frac{3\sigma_\varepsilon}{(c^2-3c+3)}\frac{1}{\sqrt{N}}\sum_{i=1}^N \left\{ \begin{array}{c} \left[W_c(1) - (1-e^c)\tilde{W}_c(\kappa)\right] \\ +c^2\int_0^1 s\left[W_c(s) + e^{sc}\tilde{W}_c(\kappa)\right] ds - c\left[W_c(1) + e^c\tilde{W}_c(\kappa)\right] \end{array} \right\} + o_p(1) \\
&\sim \mathrm{N}\left(0, \frac{9\sigma_\varepsilon^2}{\left(c^2-3c+3\right)^2} R_2\right),
\end{aligned}
$$

where

$$R_2 = Var\left\{\left[W_c(1) - (1-e^c)\,\tilde{W}_c(\kappa)\right] + c^2\int_0^1 s\left[W_c(s) + e^{sc}\tilde{W}_c(\kappa)\right]ds - c\left[W_c(1) + e^c\tilde{W}_c(\kappa)\right]\right\}.$$

proving (d). ■

F Proof of Theorem 6

Proof. To prove (c) we first note

$$\sqrt{NT}\left(\widehat{\beta}_{GLS} - \beta\right) = \frac{\frac{1}{\sqrt{N}}\frac{1}{\sqrt{T}}\sum_{i=1}^{N} z^{'}\Sigma^{-1}u_i}{\frac{1}{N}\frac{1}{T}\sum_{i=1}^{N} z^{'}\Sigma^{-1}z}$$

where

$$\begin{aligned}
z^{'}\Sigma^{-1}z &= \frac{(1-\rho)^2}{\sigma_\varepsilon^2}\sum_{t=1}^{T}\left(t-\overline{t}\right)^2 + \frac{\rho}{2\sigma_\varepsilon^2}\left[(T+1)-(T-1)\,\rho\right][T-1] \\
&= \frac{c^2}{T^2\sigma_\varepsilon^2}\sum_{t=1}^{T}\left(t-\overline{t}\right)^2 + \frac{1+\frac{c}{T}}{2\sigma_\varepsilon^2}\left[(T+1)-(T-1)\left(1+\frac{c}{T}\right)\right][T-1] \\
&= \frac{c^2}{T^2\sigma_\varepsilon^2}\sum_{t=1}^{T}\left(t-\overline{t}\right)^2 + \frac{T+c}{2T\sigma_\varepsilon^2}\left(\frac{2T^2 - cT^2 + 2cT - 2T - c}{T}\right)
\end{aligned}$$

and

$$\begin{aligned}
& z^{'}\Sigma^{-1}u_i \\
=\; & \frac{1}{\sigma_\varepsilon^2}\left\{(v_{i1}-\overline{v}_i)\left(\frac{-T+1}{2} - \frac{-T+3}{2}\rho\right) + (1-\rho)^2\sum_{t=2}^{T-1}\left(t-\overline{t}\right)(v_{it}-\overline{v}_i) + (v_{iT}-\overline{v}_i)\left[\frac{T-1}{2} - \frac{T-3}{2}\rho\right]\right\} \\
=\; & \frac{1}{\sigma_\varepsilon^2}\left\{(v_{i1}-\overline{v}_i)\left(\frac{cT-2T-3c}{2T}\right) + \frac{c^2}{T^2}\sum_{t=2}^{T-1}\left(t-\overline{t}\right)(v_{it}-\overline{v}_i) + (v_{iT}-\overline{v}_{it})\left(\frac{2T-cT+3c}{2T}\right)\right\}.
\end{aligned}$$

For a fixed N as $T\to\infty$, we have

$$\lim_{T\to\infty}\frac{1}{T}z^{'}\Sigma^{-1}z = \frac{1}{12}\frac{c^2}{\sigma_\varepsilon^2} + \frac{1}{2}\frac{2-c}{\sigma_\varepsilon^2}.$$

We next find the limiting distributions of $z^{'}\Sigma^{-1}u_i$.

$$\frac{1}{\sqrt{T}}z^{'}\Sigma^{-1}u_i$$

$$\xrightarrow{d} \frac{1}{\sigma_\varepsilon}\left\{\begin{array}{c}\left(\tilde{W}_c(\kappa)-\int_0^1\left[W_c(s)+e^{sc}\tilde{W}_c(\kappa)\right]ds\right)\left(\frac{c-2}{2}\right)+c^2\int_0^1\left(s-\frac{1}{2}\right)\left[W_c(s)+e^{sc}\tilde{W}_c(\kappa)\right]ds \\ +\left(W_c(1)+e^c\tilde{W}_c(\kappa)-\int_0^1\left[W_c(s)+e^{sc}\tilde{W}_c(\kappa)\right]ds\right)\left(\frac{2-c}{2}\right)\end{array}\right\}$$

$$= \frac{1}{\sigma_\varepsilon}\left\{\left(\frac{2-c}{2}\right)\left(W_c(1)+\left(e^c-1\right)\tilde{W}_c(\kappa)\right)+c^2\int_0^1\left(s-\frac{1}{2}\right)\left[W_c(s)+e^{sc}\tilde{W}_c(\kappa)\right]ds\right\}.$$

Thus

$$\begin{aligned}
&\sqrt{NT}\left(\widehat{\beta}_{GLS}-\beta\right) \\
=\; & \frac{\frac{1}{\sqrt{N}}\frac{1}{\sqrt{T}}\sum_{i=1}^N z'\Sigma^{-1}u_i}{\frac{1}{N}\frac{1}{T}\sum_{i=1}^N z'\Sigma^{-1}z} \\
& \xrightarrow{d} \frac{\frac{1}{\sqrt{N}}\sum_{i=1}^N\frac{1}{\sigma_\varepsilon}\left\{\left(\frac{2-c}{2}\right)\left(W_c(1)-\left(1-e^c\right)\tilde{W}_c(\kappa)\right)+c^2\int_0^1\left(s-\frac{1}{2}\right)\left[W_c(s)+e^{sc}\tilde{W}_c(\kappa)\right]ds\right\}}{\frac{1}{N}\sum_{i=1}^N\left(\frac{1}{12}\frac{c^2}{\sigma_\varepsilon^2}+\frac{1}{2}\frac{2-c}{\sigma_\varepsilon^2}\right)} \\
=\; & \frac{12\sigma_\varepsilon}{c^2+12-6c}\frac{1}{\sqrt{N}}\sum_{i=1}^N\left\{\left(\frac{2-c}{2}\right)\left(W_c(1)-\left(1-e^c\right)\tilde{W}_c(\kappa)\right)+c^2\int_0^1\left(s-\frac{1}{2}\right)\left[W_c(s)+e^{sc}\tilde{W}_c(\kappa)\right]ds\right\}+o(1) \\
\sim\; & \mathrm{N}\left(0,\sigma_\varepsilon^2\left(\frac{12}{c^2+12-6c}\right)^2R_3\right),
\end{aligned}$$

where

$$R_3 = Var\left\{\left(\frac{2-c}{2}\right)\left(W_c(1)-\left(1-e^c\right)\tilde{W}_c(\kappa)\right)+c^2\int_0^1\left(s-\frac{1}{2}\right)\left[W_c(s)+e^{sc}\tilde{W}_c(\kappa)\right]ds\right\}.$$

■

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*, John Wiley & Sons, 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*, John Wiley & Sons, 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," Manuscript.

[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	.372	.422	1	.235	.054	.373	.422
	0.2	.997	.241	.079	.417	.461	.997	.239	.078	.417	.460
	0.4	.982	.250	.125	.489	.529	.982	.245	.116	.489	.523
	0.6	.950	.269	.217	.616	.654	.949	.255	.184	.617	.639
	0.8	.876	.343	.450	.839	.868	.871	.308	.355	.841	.853
	0.9	.813	.493	.709	.949	.961	.791	.442	.592	.954	.963
	0.95	.799	.654	.875	.971	.974	.749	.617	.806	.963	.968
	1.0	.853	.831	.957			.803	.896	1.031		
0.2	0	.518	.683	.157	.763	.845	.518	.683	.157	.763	.845
	0.2	.623	.608	.201	.828	.914	.625	.606	.198	.829	.913
	0.4	.732	.516	.257	.849	.929	.738	.509	.241	.855	.924
	0.6	.818	.425	.343	.852	.918	.826	.407	.293	.859	.903
	0.8	.828	.405	.532	.913	.951	.825	.364	.420	.912	.935
	0.9	.785	.519	.746	.958	.974	.764	.465	.622	.961	.974
	0.95	.776	.663	.887	.963	.969	.729	.626	.818	.954	.962
	1.0	.845	.838	.964			.776	.879	1.013		
0.4	0	.268	.839	.193	.422	.449	.268	.839	.193	.422	.449
	0.2	.354	.779	.257	.571	.614	.355	.778	.254	.573	.616
	0.4	.469	.684	.341	.734	.797	.477	.680	.322	.746	.804
	0.6	.611	.559	.450	.861	.934	.629	.546	.393	.883	.939
	0.8	.731	.478	.627	.938	.986	.738	.434	.501	.944	.978
	0.9	.737	.552	.799	.956	.977	.721	.498	.667	.957	.975
	0.95	.741	.678	.907	.949	.957	.699	.639	.836	.938	.948
	1.0	.828	.845	.972			.747	.869	1.000		
0.6	0	.136	.919	.211	.188	.195	.136	.919	.211	.188	.195
	0.2	.188	.878	.289	.281	.294	.189	.878	.287	.282	.295
	0.4	.269	.799	.398	.429	.457	.276	.801	.379	.440	.466
	0.6	.394	.674	.543	.637	.687	.416	.673	.485	.671	.714
	0.8	.570	.559	.734	.849	.899	.592	.522	.603	.875	.916
	0.9	.647	.607	.873	.913	.940	.639	.548	.734	.917	.941
	0.95	.678	.703	.941	.914	.926	.643	.664	.868	.902	.916
	1.0	.790	.852	.981			.706	.863	.994		
0.8	0	.055	.967	.222	.065	.066	.055	.967	.222	.065	.066
	0.2	.078	.943	.311	.093	.101	.079	.943	.308	.099	.101
	0.4	.118	.883	.440	.157	.157	.122	.891	.422	.166	.174
	0.6	.189	.773	.622	.288	.356	.204	.790	.569	.314	.329
	0.8	.328	.645	.847	.539	.571	.359	.645	.734	.587	.618
	0.9	.454	.678	.975	.718	.748	.463	.627	.839	.736	.764
	0.95	.535	.749	1.003	.793	.811	.515	.711	.929	.784	.803
	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

		κ = 0.25					κ = 1				
Δ	ρ	OLS	FE	FD	GLS-CO	GLS-PW	OLS	FE	FD	GLS-CO	GLS-PW
0	0	1	.235	.054	.372	.422	1	.235	.054	.372	.422
	0.2	.997	.240	.078	.417	.460	.997	.240	.078	.417	.460
	0.4	.982	.245	.116	.489	.523	.982	.245	.116	.489	.523
	0.6	.949	.253	.179	.617	.635	.949	.253	.179	.617	.635
	0.8	.878	.352	.447	.846	.871	.873	.317	.348	.844	.852
	0.9	.769	.398	.501	.958	.964	.752	.369	.451	.960	.966
	0.95	.689	.569	.721	.948	.957	.702	.499	.612	.914	.928
	1.0	.735	1.021	1.175			.599	1.657	1.907		
0.2	0	.518	.683	.157	.763	.845	.518	.683	.157	.763	.845
	0.2	.625	.606	.198	.829	.913	.625	.606	.198	.829	.913
	0.4	.738	.508	.240	.856	.924	.738	.508	.240	.856	.924
	0.6	.828	.404	.286	.859	.900	.827	.404	.286	.859	.901
	0.8	.831	.418	.531	.923	.958	.829	.377	.414	.919	.936
	0.9	.745	.418	.527	.962	.973	.728	.388	.474	.963	.973
	0.95	.674	.577	.731	.939	.949	.588	.506	.619	.902	.918
	1.0	.683	.959	1.105			.492	1.375	1.583		
0.4	0	.268	.839	.193	.423	.449	.268	.839	.193	.423	.449
	0.2	.355	.778	.254	.573	.616	.355	.778	.254	.573	.616
	0.4	.477	.680	.322	.747	.805	.477	.680	.322	.747	.805
	0.6	.633	.545	.386	.887	.941	.633	.545	.386	.887	.941
	0.8	.732	.494	.628	.949	.994	.741	.452	.496	.953	.982
	0.9	.707	.449	.566	.958	.974	.692	.417	.510	.958	.972
	0.95	.649	.590	.748	.923	.936	.569	.518	.634	.885	.902
	1.0	.637	.914	1.052			.407	1.151	1.325		
0.6	0	.136	.919	.211	.188	.195	.136	.919	.211	.188	.195
	0.2	.189	.878	.287	.282	.295	.189	.878	.287	.282	.295
	0.4	.276	.801	.379	.441	.467	.276	.801	.379	.441	.467
	0.6	.420	.674	.477	.677	.719	.419	.674	.477	.677	.719
	0.8	.568	.581	.736	.856	.905	.591	.545	.598	.882	.919
	0.9	.636	.499	.628	.923	.945	.626	.465	.567	.923	.945
	0.95	.604	.613	.777	.888	.903	.536	.538	.659	.850	.869
	1.0	.592	.883	1.016			.342	.989	1.139		
0.8	0	.055	.967	.222	.065	.066	.055	.967	.222	.065	.066
	0.2	.078	.943	.308	.099	.101	.078	.943	.308	.099	.101
	0.4	.122	.891	.421	.165	.174	.122	.891	.421	.165	.174
	0.6	.207	.794	.563	.164	.332	.207	.794	.562	.160	.332
	0.8	.323	.668	.848	.537	.568	.356	.662	.727	.586	.616
	0.9	.476	.586	.738	.761	.783	.479	.556	.677	.772	.799
	0.95	.496	.661	.838	.776	.796	.454	.585	.717	.751	.772
	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.

Center for Policy Research Working Paper Series

Number	Author(s)	Title
1	Kao and Emerson	On the Estimation of a Linear Time Trend Regression with a One-Way Error Component Model in the Presence of Serially Correlated Errors