Syracuse University
SURFACE

Center for Policy Research | Maxwell School of Citizenship and Public Affairs

1-1-1999

On the Estimation and Inference of a Cointegrated Regression in Panel Data

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

Follow this and additional works at: http://surface.syr.edu/cpr

Part of the Mathematics Commons

Recommended Citation
Kao, Chihwa, "On the Estimation and Inference of a Cointegrated Regression in Panel Data" (1999). *Center for Policy Research.* Paper 145.
http://surface.syr.edu/cpr/145

This Working Paper is brought to you for free and open access by the Maxwell School of Citizenship and Public Affairs at SURFACE. It has been accepted for inclusion in Center for Policy Research by an authorized administrator of SURFACE. For more information, please contact surface@syr.edu.

Center for Policy Research
Working Paper No. 2

ON THE ESTIMATION AND INFERENCE OF A COINTEGRATED REGRESSION IN PANEL DATA

Chihwa Kao and Min-Hsien Chiang

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 and Inference of a Cointegrated Regression in Panel Data

Chihwa Kao*
Syracuse University

Min-Hsien Chiang
National Cheng-Kon University

First Draft: May 1995
First Revised Draft January 1997
Second Revised Draft: September 1997
Third Revised Draft: August 1998
This Draft: September 16, 1998

Abstract

The main contribution of this paper is to add to the literature by suggesting a dynamic OLS (DOLS) estimator and providing a serious comparison of the finite sample properties of the OLS, fully modified OLS (FMOLS), and DOLS estimators in panel cointegrated regression models. Monte Carlo results illustrate the sampling behavior of the proposed estimators and show that (1) the OLS estimator has a non-negligible bias in finite samples, (2) the FMOLS estimator does not improve over the OLS estimator in general, and (3) the DOLS outperforms both the OLS and FMOLS estimators.

Key Words: *Panel Data, OLS, FMOLS, DOLS, Homogeneous Panel, Heterogeneous Panel.*

1 Introduction

Evaluating the statistical properties of data along the time dimension has proven to be very different from analysis of the cross-section dimension. As economists have gained access to better data with more observations across time, understanding these properties has grown increasingly important. An area of particular

*The first part of this paper was previously circulated under the title, On the Estimation and Inference for Cointegration in Panel Data When the Cross-Section and Time-Series Dimensions are Comparable, by Kao and Chen. We are thankful to four referees and Peter Phillips for pointing out several technical errors in an earlier version and providing comments that led to improvements of this paper. We also thank Suzanne McCoskey, Peter Pedroni, Andrew Levin and participants of the 1998 North American Winter Meetings of the Econometric Society for helpful comments and Bangtian Chen for his research assistance on an earlier draft of this paper. Thanks also go to Martha Bonney for correcting our English and carefully checking the manuscript to enhance its readability. 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.

concern in time-series econometrics has been the use of non-stationary data. With the desire to study the behavior of cross-sectional data over time and the increasing use of panel data, e.g., Summers and Heston (1991) data, one new research area is examining the properties of non-stationary time-series data in panel form. It is an intriguing question to ask: how exactly does this hybrid style of data combine the statistical elements of traditional cross-sectional analysis and time-series analysis? In particular, what is the correct way to analyze non-stationarity, the spurious regression problem, and cointegration in panel data?

Given the immense interest in testing for unit roots and cointegration in time-series data, not much attention has been paid to testing the unit roots in panel data. The only theoretical studies we know of in this area are Breitung and Meyer (1994); Quah (1994); Levin and Lin (1993); Im, Pesaran, and Shin (1995); and Maddala and Wu (1996). Breitung and Meyer (1994) derived the asymptotic normality of the Dickey-Fuller test statistic for panel data with a large cross-section dimension and a small time-series dimension. Quah (1994) studied a unit root test for panel data that simultaneously have extensive cross-section and time-series variation. He showed that the asymptotic distribution for the proposed test is a mixture of the standard normal and Dickey-Fuller-Phillips asymptotics. Levin and Lin (1993) derived the asymptotic distributions for unit roots on panel data and showed that the power of these tests increases dramatically as the cross-section dimension increases. Im et al. (1995) critiqued the Levin and Lin panel unit root statistics and proposed alternatives. Maddala and Wu (1996) provided a comparison of the tests of Im et al. (1995) and Levin and Lin (1993). They suggested a new test based on the Fisher test.

Recently, some attention has been given to the cointegration tests and estimation with regression models in panel data, e.g., Kao (1997), McCoskey and Kao (1998), Pedroni (1996, 1997) and Phillips and Moon (1998). Kao (1997) studied a spurious regression in panel data, along with asymptotic properties of the ordinary least squares (OLS) estimator and other conventional statistics. Kao (1997) showed that the OLS estimator is consistent for its true value, but the t-statistic diverges so that inferences about the regression coefficient, β, are wrong with a probability that goes to one. Furthermore, Kao (1997) examined the Dickey-Fuller (DF) and the augmented Dickey-Fuller (ADF) tests to test the null hypothesis of no cointegration in panel data. McCoskey and Kao (1998) proposed further tests for the null hypothesis of cointegration in panel data. Pedroni (1997) derived asymptotic distributions for residual-based tests of cointegration for both homogeneous and heterogeneous panels. Pedroni (1996) proposed a fully modified estimator for heterogeneous panels. Phillips and Moon (1998) developed both sequential limit and joint limit theories for non-stationary panel data. Pesaran and Smith (1995) are not directly concerned with cointegration but do touch on a number of related issues, including the potential problems of homogeneity misspecification for cointegrated panels.

This paper makes two main contributions. First, it adds to the literature by suggesting a computationally simpler dynamic OLS (DOLS) estimator in panel cointegrated regression models. Second, it provides a serious study of the finite sample properties of the OLS, fully modified OLS (FMOLS), and DOLS estimators.

Section 2 introduces the model and assumptions. Section 3 develops the asymptotic theory for the OLS, FMOLS and DOLS estimators. Section 4 gives the limiting distributions of the FMOLS and DOLS estimators for heterogeneous panels. Section 5 presents Monte Carlo results to illustrate the finite sample properties of the OLS, FMOLS, and DOLS estimators. Section 6 summarizes the findings. The proofs of Theorems 1, 2, and 4 are not presented since the proofs can be found in Phillips and Moon (1998) and Pedroni (1997). Appendix contains the proofs of the Theorems 5 and 6.

A word on notation. We write the integral $\int_0^1 W(s)ds$ as $\int W$ when there is no ambiguity over limits. We define $\Omega^{1/2}$ to be any matrix such that $\Omega = \left(\Omega^{1/2}\right)\left(\Omega^{1/2}\right)^{'}$. We use $\|A\|$ to denote $\left\{tr\left(A^{'}A\right)\right\}^{1/2}$, $|A|$ to denote the determinant of A, $\Rightarrow$ to denote weak convergence, $\xrightarrow{p}$ to denote convergence in probability, $[x]$ to denote the largest integer $\leq x$, $I(0)$ and $I(1)$ to signify a time-series that is integrated of order zero and one, respectively, and $BM(\Omega)$ to denote Brownian motion with the covariance matrix Ω.

2 The Model and Assumptions

Consider the following fixed effect panel regression:

$$y_{it} \quad = \quad \alpha_i \quad + \quad x_{it}^{'}\beta \quad + \quad u_{it},\ i = 1, ..., N,\ t = 1, ..., T, \tag{1}$$

where $\{y_{it}\}$ are 1×1, β is a $k \times 1$ vector of the slope parameters, $\{\alpha_i\}$ are the intercepts, and $\{u_{it}\}$ are the stationary disturbance terms. We assume that $\{x_{it}\}$ are $k \times 1$ integrated processes of order one for all i, where

$$x_{it} = x_{it-1} + \varepsilon_{it}.$$

Under these specifications, (1) describes a system of cointegrated regressions, i.e., y_{it} is cointegrated with x_{it}. The initialization of this system is $y_{i0} = x_{i0} = O_p(1)$ as $T \to \infty$ for all i. The individual constant term α_i can be extended into general deterministic time trends such as $\alpha_{0i} + \alpha_{1i}t+, ..., +\alpha_{pi}t^p$.

Assumption 1 *The asymptotic theory employed in this paper is a sequential limit theory established by Phillips and Moon (1997) in which $T \to \infty$ followed by $N \to \infty$.*

Next, we characterize the innovation vector $w_{it} = \left(u_{it}, \varepsilon_{it}^{'}\right)^{'}$. We assume that w_{it} is a linear process that satisfies the following assumption.

Assumption 2 *For each i, we assume (e.g., Phillips, 1995) :*

(a) $w_{it} = \Pi(L)\epsilon_{it} = \sum_{j=0}^{\infty} \Pi_j \epsilon_{it-j}, \sum_{j=0}^{\infty} j^a \|\Pi_j\| < \infty, |\Pi(1)| \neq 0$ *for some* $a > 1$.

(b) ϵ_{it} *is i.i.d. with zero mean, variance matrix* Σ_ϵ, *and finite fourth order cumulants.*

Assumption 2 implies that (e.g., Phillips and Solo, 1992) the partial sum process $\frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} w_{it}$ satisfies the following multivariate invariance principle:

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} w_{it} \Rightarrow B_i(r) = BM_i(\Omega) \text{ as } T \to \infty \text{ for all } i, \tag{2}$$

where

$$B_i = \begin{bmatrix} B_{ui} \\ B_{\varepsilon i} \end{bmatrix}.$$

The long-run covariance matrix of $\{w_{it}\}$ is given by

$$\begin{aligned} \Omega &= \sum_{j=-\infty}^{\infty} E\left(w_{ij} w_{i0}^{'}\right) \\ &= \Pi(1)\Sigma_\epsilon \Pi(1)^{'} \\ &= \Sigma + \Gamma + \Gamma^{'} \\ &= \begin{bmatrix} \Omega_u & \Omega_{u\varepsilon} \\ \Omega_{\varepsilon u} & \Omega_\varepsilon \end{bmatrix}, \end{aligned}$$

where

$$\Gamma = \sum_{j=1}^{\infty} E\left(w_{ij} w_{i0}^{'}\right) = \begin{bmatrix} \Gamma_u & \Gamma_{u\varepsilon} \\ \Gamma_{\varepsilon u} & \Gamma_\varepsilon \end{bmatrix} \tag{3}$$

and

$$\Sigma = E\left(w_{i0} w_{i0}^{'}\right) = \begin{bmatrix} \Sigma_u & \Sigma_{u\varepsilon} \\ \Sigma_{\varepsilon u} & \Sigma_\varepsilon \end{bmatrix} \tag{4}$$

are partitioned conformably with w_{it}.

Assumption 3 Ω_ε *is non-singular, i.e.,* $\{x_{it}\}$ *are not cointegrated.*

Define

$$\Omega_{u.\varepsilon} = \Omega_u - \Omega_{u\varepsilon}\Omega_\varepsilon^{-1}\Omega_{\varepsilon u}. \tag{5}$$

Then, B_i can be rewritten as

$$B_i = \begin{bmatrix} B_{ui} \\ B_{\varepsilon i} \end{bmatrix} = \begin{bmatrix} \Omega_{u.\varepsilon}^{1/2} & \Omega_{u\varepsilon}\Omega_\varepsilon^{-1/2} \\ 0 & \Omega_\varepsilon^{1/2} \end{bmatrix} \begin{bmatrix} V_i \\ W_i \end{bmatrix}, \tag{6}$$

where $\begin{bmatrix} V_i \\ W_i \end{bmatrix} = BM(I)$ is a standardized Brownian motion. Define the one-sided long-run covariance

$$\begin{aligned} \Delta &= \Sigma + \Gamma \\ &= \sum_{j=0}^{\infty} E\left(w_{ij} w_{i0}^{'}\right) \end{aligned}$$

with

$$\Delta = \begin{bmatrix} \Delta_u & \Delta_{u\varepsilon} \\ \Delta_{\varepsilon u} & \Delta_{\varepsilon} \end{bmatrix}.$$

Here we assume that panels are homogeneous, i.e., the variances are constant across the cross-section units. We will relax this assumption in Section 4 to allow for different variances for different i.

Remark 1 *The benefits of using panel data model have been discussed extensively by Hsiao (1986) and Baltagi (1995), though Hsiao and Baltagi assume the time dimension is small while the cross-section dimension is large. However, in international trade, open macroeconomics, urban regional, public finance, and finance, panel data usually have long time-series and cross-section dimensions. The data of Summers and Heston (1991) data are a notable example.*

Remark 2 *The advantage of using the sequential limit theory is that it offers a quick and easy way to derive the asymptotics as demonstrated by Phillips and Moon (1997). Phillips and Moon (1997) also provide detailed treatments of the connections between the sequential limit theory and the joint limit theory.*

Remark 3 *If one wants to obtain a consistent estimate of β in (1) or wants to test some restrictions on β, then an individual time-series regression or a multiple time-series regression is probably enough. So what are the advantages of using the (N, T) asymptotics, e.g., sequential asymptotics in Assumption 1, instead of T asymptotics? One of the advantages is that we can get a normal approximation of the limit distributions of the estimators and test statistics with the convergence rate $\sqrt{N}T$. More importantly, the biases of the estimators and test statistics can be reduced when N and T are large. For example, later in this paper we will show that the biases of the OLS, FMOLS, and DOLS estimators in Table 2 were reduced by half when the sample size was changed from $(N = 1, T = 20)$ to $(N = 20, T = 20)$. However, in order to obtain an asymptotic normality using the (N, T) asymptotics we need to make some strong assumptions; for example, in this paper we assume that the error terms are independent across i.*

Remark 4 *The results in this paper require that regressors are not cointegrated. Assuming that $I(1)$ regressors are not cointegrated with each other is indeed restrictive. The authors are currently investigating this issue.*

3 OLS, FMOLS, and DOLS Estimators

Let us first study the limiting distribution of the OLS estimator for equation (1). The OLS estimator of β is

$$\widehat{\beta}_{OLS} = \left[\sum_{i=1}^{N}\sum_{t=1}^{T}(x_{it}-\overline{x}_i)(x_{it}-\overline{x}_i)'\right]^{-1}\left[\sum_{i=1}^{N}\sum_{t=1}^{T}(x_{it}-\overline{x}_i)(y_{it}-\overline{y}_i)\right]. \tag{7}$$

All the limits in Theorems $1-6$ are taken as $T\to\infty$ followed by $N\to\infty$ sequentially from Assumption 1. First, we present the following theorem:

Theorem 1 *If Assumptions* $1-3$ *hold, then*

(a) $T\left(\widehat{\beta}_{OLS}-\beta\right)\xrightarrow{p}-3\Omega_{\varepsilon}^{-1}\Omega_{\varepsilon u}+6\Omega_{\varepsilon}^{-1}\Delta_{\varepsilon u}$,

(b) $\sqrt{N}T\left(\widehat{\beta}_{OLS}-\beta\right)-\sqrt{N}\delta_{NT}\Rightarrow \mathrm{N}\left(0,6\Omega_{\varepsilon}^{-1}\Omega_{u.\varepsilon}\right)$,

where

$$\delta_{NT}=\left[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=1}^{T}(x_{it}-x_{it})(x_{it}-\overline{x}_i)'\right]^{-1}\left[\frac{1}{N}\sum_{i=1}^{N}\Omega_{\varepsilon}^{1/2}\left(\int\widetilde{W}_i dW_i'\right)\Omega_{\varepsilon}^{-1/2}\Omega_{\varepsilon u}+\Delta_{\varepsilon u}\right]$$

and $\widetilde{W}_i=W_i-\int W_i$.

The normality of the OLS estimator in Theorem 1 comes naturally. When summing across i, the non-standard asymptotic distribution due to the unit root in the time dimension is smoothed out. From Theorem 1 we note that there is an interesting interpretation of the asymptotic covariance matrix, $6\Omega_{\varepsilon}^{-1}\Omega_{u.\varepsilon}$, i.e., $\Omega_{\varepsilon}^{-1}\Omega_{u.\varepsilon}$ can be seen as the long-run noise-to-signal ratio. We also note that $-\frac{1}{2}\Omega_{\varepsilon u}$ is due to the endogeneity of the regressor x_{it}, and $\Delta_{\varepsilon u}$ is due to the serial correlation. It can be shown easily that

$$\delta_{NT}\xrightarrow{p}-3\Omega_{\varepsilon}^{-1}\Omega_{\varepsilon u}+6\Omega_{\varepsilon}^{-1}\Delta_{\varepsilon u}.$$

If $w_{it}=\left(u_{it},\varepsilon_{it}'\right)'$ are i.i.d., then

$$\delta_{NT}\xrightarrow{p}3\Sigma_{\varepsilon}^{-1}\Sigma_{\varepsilon u},$$

which was examined by Kao and Chen (1995). Let $\widehat{\Omega}_{\varepsilon}$, $\widehat{\Omega}_{\varepsilon u},\widehat{\Omega}_{\varepsilon}$, and $\widehat{\Delta}_{\varepsilon u}$ be consistent estimates of Ω_{ε}, $\Omega_{\varepsilon u},\Omega_{\varepsilon}$, and $\Delta_{\varepsilon u}$ respectively. Then from (b) in Theorem 1, we can define a bias-corrected OLS, $\widehat{\beta}_{OLS}^{+}$,

$$\widehat{\beta}_{OLS}^{+}=\widehat{\beta}_{OLS}-\frac{\widehat{\delta}_{NT}}{T}$$

such that

$$\sqrt{N}T\left(\widehat{\beta}_{OLS}^{+}-\beta\right)\Rightarrow \mathrm{N}\left(0,6\Omega_{\varepsilon}^{-1}\Omega_{u.\varepsilon}\right),$$

where

$$\widehat{\delta}_{NT} = -3\widehat{\Omega}_{\varepsilon}^{-1}\widehat{\Omega}_{\varepsilon u} + 6\widehat{\Omega}_{\varepsilon}^{-1}\widehat{\Delta}_{\varepsilon u}.$$

Chen, McCoskey, and Kao (1996) investigated the finite sample proprieties of the OLS estimator in (7), the t-statistic, the bias-corrected OLS estimator, and the bias-corrected t-statistic. They found that the bias-corrected OLS estimator does not improve over the OLS estimator in general. The results of Chen, McCoskey, and Kao (1996) suggest that alternatives, such as the FMOLS estimator or the DOLS estimator (e.g., Saikkonen, 1991; Stock and Watson, 1993) may be more promising in cointegrated panel regressions. Thus, we begin our study by examining the limiting distribution of the FMOLS estimator, $\widehat{\beta}_{FM}$. The FMOLS estimator is constructed by making corrections for endogeneity and serial correlation to the OLS estimator $\widehat{\beta}_{OLS}$ in (7). Define

$$u_{it}^{+} = u_{it} - \Omega_{u\varepsilon}\Omega_{\varepsilon}^{-1}\varepsilon_{it}, \tag{8}$$

$$\widehat{u}_{it}^{+} = u_{it} - \widehat{\Omega}_{u\varepsilon}\widehat{\Omega}_{\varepsilon}^{-1}\varepsilon_{it}, \tag{9}$$

$$y_{it}^{+} = y_{it} - \Omega_{u\varepsilon}\Omega_{\varepsilon}^{-1}\Delta x_{it}, \tag{10}$$

and

$$\widehat{y}_{it}^{+} = y_{it} - \widehat{\Omega}_{u\varepsilon}\widehat{\Omega}_{\varepsilon}^{-1}\Delta x_{it}. \tag{11}$$

Note that

$$\begin{bmatrix} u_{it}^{+} \\ \varepsilon_{it} \end{bmatrix} = \begin{bmatrix} 1 & -\Omega_{u\varepsilon}\Omega_{\varepsilon}^{-1} \\ 0 & \mathbf{I}_k \end{bmatrix} \begin{bmatrix} u_{it} \\ \varepsilon_{it} \end{bmatrix},$$

which has the long-run covariance matrix

$$\begin{bmatrix} \Omega_{u.\varepsilon} & 0 \\ 0 & \Omega_{\varepsilon} \end{bmatrix},$$

where $\mathbf{I}_k$ is a $k \times k$ identity matrix. The endogeneity correction is achieved by modifying the variable y_{it} in (1) with the transformation

$$\begin{aligned} \widehat{y}_{it}^{+} &= y_{it} - \widehat{\Omega}_{u\varepsilon}\widehat{\Omega}_{\varepsilon}^{-1}\Delta x_{it} \\ &= \alpha_i + x_{it}^{'}\beta + u_{it} - \widehat{\Omega}_{u\varepsilon}\widehat{\Omega}_{\varepsilon}^{-1}\Delta x_{it}. \end{aligned}$$

The serial correlation correction term has the form

$$\begin{aligned} \widehat{\Delta}_{\varepsilon u}^{+} &= \begin{pmatrix} \widehat{\Delta}_{\varepsilon u} & \widehat{\Delta}_{\varepsilon} \end{pmatrix} \begin{pmatrix} 1 \\ -\widehat{\Omega}_{\varepsilon}^{-1}\widehat{\Omega}_{\varepsilon u} \end{pmatrix} \\ &= \widehat{\Delta}_{\varepsilon u} - \widehat{\Delta}_{\varepsilon}\widehat{\Omega}_{\varepsilon}^{-1}\widehat{\Omega}_{\varepsilon u}, \end{aligned}$$

where $\widehat{\Delta}_{\varepsilon u}$ and $\widehat{\Delta}_{\varepsilon}$ are kernel estimates of $\Delta_{\varepsilon u}$ and Δ_{ε}. Therefore, the FMOLS estimator is

$$\widehat{\beta}_{FM} = \left[\sum_{i=1}^{N}\sum_{t=1}^{T}\left(x_{it}-\overline{x}_i\right)\left(x_{it}-\overline{x}_i\right)'\right]^{-1}\left[\sum_{i=1}^{N}\left(\sum_{t=1}^{T}\left(x_{it}-\overline{x}_i\right)\widehat{y}_{it}^{+}-T\widehat{\Delta}_{\varepsilon u}^{+}\right)\right]. \tag{12}$$

Now, we state the limiting distribution of $\widehat{\beta}_{FM}$.

Theorem 2 *If Assumptions* $1-3$ *hold, then* $\sqrt{N}T\left(\widehat{\beta}_{FM}-\beta\right)\Rightarrow \mathrm{N}\left(0,6\Omega_{\varepsilon}^{-1}\Omega_{u.\varepsilon}\right)$.

It can be shown easily that the limiting distribution of $\widehat{\beta}_{FM}$ becomes

$$\sqrt{N}T\left(\widehat{\beta}_{FM}-\beta\right)\Rightarrow \mathrm{N}\left(0,2\Omega_{\varepsilon}^{-1}\Omega_{u.\varepsilon}\right) \tag{13}$$

by the exclusion of the individual-specific intercept, α_i.

Remark 5 *Once the estimates of* w_{it}, $\widehat{w}_{it}$ *were estimated, we used*

$$\widehat{\Sigma} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\widehat{w}_{it}\widehat{w}_{it}' \tag{14}$$

to estimate Σ. Ω *was estimated by*

$$\widehat{\Omega} = \frac{1}{N}\sum_{i=1}^{N}\left\{\frac{1}{T}\sum_{t=1}^{T}\widehat{w}_{it}\widehat{w}_{it}'+\frac{1}{T}\sum_{\tau=1}^{l}\varpi_{\tau l}\sum_{t=\tau+1}^{T}\left(\widehat{w}_{it}\widehat{w}_{it-\tau}'+\widehat{w}_{it-\tau}\widehat{w}_{it}'\right)\right\}, \tag{15}$$

where $\varpi_{\tau l}$ *is a weight function or a kernel. Using Phillips and Durlauf (1986) and sequential limit theory,* $\widehat{\Sigma}$ *and* $\widehat{\Omega}$ *can be shown to be consistent for* Σ *and* Ω.

Next, we propose a DOLS estimator, $\widehat{\beta}_D$, which uses the past and future values of $\triangle x_{it}$ as additional regressors. We then show that the limiting distribution of $\widehat{\beta}_D$ is the same as the FMOLS estimator, $\widehat{\beta}_{FM}$. But first, we need the following additional assumption:

Assumption 4 *The spectral density matrix* $f_{ww}(\lambda)$ *is bounded away from zero and full rank for all* i*, i.e.,*

$$f_{ww}(\lambda)\geq \delta I_T, \lambda\in[0,\pi], \delta>0.$$

When Assumptions 2 and 4 hold, the process $\{u_{it}\}$ can be written as (see Saillonen, 1991):

$$u_{it} = \sum_{j=-\infty}^{\infty} c_{ij}\varepsilon_{it+j}+v_{it}, \tag{16}$$

for all i, where

$$\sum_{j=-\infty}^{\infty}\|c_{ij}\|<\infty,$$

$\{v_{it}\}$ is stationary with zero mean, and $\{v_{it}\}$ and $\{\varepsilon_{it}\}$ are uncorrelated not only contemporaneously but also in all lags and leads. In practice, the leads and lags may be truncated while retaining (16) approximately, so that

$$u_{it} = \sum_{j=-q}^{q} c_{ij}\varepsilon_{it+j} + \dot{v}_{it} \ .$$

for all i. This is because $\{c_{ij}\}$ are assumed to be absolutely summable, i.e., $\sum_{j=-\infty}^{\infty} \|c_{ij}\| < \infty$.

We also need to require that q tends to infinity with T at a suitable rate:

Assumption 5 *$q \to \infty$ as $T \to \infty$ such that $\frac{q^3}{T} \to 0$, and*

$$T^{1/2} \sum_{|j|>q} \|c_{ij}\| \to 0 \tag{17}$$

for all i.

We then substitute (16) into (1) to get

$$y_{it} \quad = \quad \alpha_i \quad + \quad x_{it}^{'}\beta \ + \sum_{j=-q}^{q} c_{ij}\varepsilon_{it+j} + \dot{v}_{it},$$

where

$$\dot{v}_{it} = v_{it} + \sum_{|j|>q} c_{ij}\varepsilon_{it+j}. \tag{18}$$

Therefore, we obtain the DOLS of β, $\widehat{\beta}_D$, by running the following regression:

$$y_{it} \quad = \quad \alpha_i \quad + \quad x_{it}^{'}\beta \ + \sum_{j=-q}^{q} c_{ij}\triangle x_{it+j} + \dot{v}_{it} \ . \tag{19}$$

Next, we show that $\widehat{\beta}_D$ has the same limiting distribution $\widehat{\beta}_{FM}$ as in Theorem 2.

Theorem 3 *If Assumptions $1-5$ hold, then $\sqrt{N}T\left(\widehat{\beta}_D - \beta\right) \Rightarrow \mathrm{N}\left(0, 6\Omega_{\varepsilon}^{-1}\Omega_{u.\varepsilon}\right)$.*

4 Heterogeneous Panels

The paper so far assumes that the panel data are homogeneous. The substantial heterogeneity exhibited by actual data in the cross-sectional dimension may restrict the practical applicability of the FMOLS and DOLS estimators. Also, the estimators in Sections 2 and 3 are not easily extended to cases of broader cross-sectional heterogeneity since the variances and biases are specified in terms of the asymptotic covariance parameters that are assumed to be shared cross-sectionally.

In this section, we propose an alternative representation of the panel FMOLS estimator for heterogeneous panels. Before we discuss the FMOLS estimator we need the following assumptions:

Assumption 6 *We assume the panels are heterogeneous, i.e.,* Ω_i, Γ_i *and* Σ_i *are varied for different* i*. We also assume the invariance principle in (2), (16), and (17) in Assumption 5 still holds.*

Let

$$x_{it}^* = \widehat{\Omega}_{i\varepsilon}^{-1/2} x_{it}, \tag{20}$$

$$u_{it}^* = \widehat{\Omega}_{iu.\varepsilon}^{-1/2} \widehat{u}_{it}^+, \tag{21}$$

$$\begin{aligned} \widehat{u}_{it}^+ &= u_{it} - \widehat{\Omega}_{iu\varepsilon} \widehat{\Omega}_{i\varepsilon}^{-1} \varepsilon_{it}, \\ \widehat{y}_{it}^+ &= y_{it} - \widehat{\Omega}_{iu\varepsilon} \widehat{\Omega}_{i\varepsilon}^{-1} \Delta x_{it} - \widehat{\Omega}_{iu.\varepsilon}^{1/2} \left(\widehat{\Omega}_{iu.\varepsilon}^{-1/2} - \widehat{\Omega}_{i\varepsilon}^{-1/2} \right) x_{it} \beta, \end{aligned} \tag{22}$$

and

$$y_{it}^* = \widehat{\Omega}_{iu.\varepsilon}^{-1/2} \widehat{y}_{it}^+, \tag{23}$$

where $\widehat{\Omega}_{i\varepsilon}$ and $\widehat{\Omega}_{iu.\varepsilon}$ are consistent estimators of $\Omega_{i\varepsilon}$ and

$$\Omega_{iu.\varepsilon} = \Omega_{iu} - \Omega_{iu\varepsilon} \Omega_{i\varepsilon}^{-1} \Omega_{i\varepsilon u},$$

respectively. Similar to Pedroni (1996) the correction term, $\widehat{\Omega}_{iu.\varepsilon}^{1/2} \left(\widehat{\Omega}_{iu.\varepsilon}^{-1/2} - \widehat{\Omega}_{i\varepsilon}^{-1/2} \right) x_{it} \beta$, is needed in (22) in the heterogeneous panel. We note that (22) will be the same as (11) only if $\widehat{\Omega}_{iu.\varepsilon} = \widehat{\Omega}_{i\varepsilon}$ in the heterogeneous panel. Also (22) requires knowing something about the true β. In practice, β in (22) can be replaced by a preliminary OLS, $\widehat{\beta}_{OLS}$. Therefore, let

$$\widehat{y}_{it}^{++} = y_{it} - \widehat{\Omega}_{iu\varepsilon} \widehat{\Omega}_{i\varepsilon}^{-1} \Delta x_{it} - \widehat{\Omega}_{iu.\varepsilon}^{1/2} \left(\widehat{\Omega}_{iu.\varepsilon}^{-1/2} - \widehat{\Omega}_{i\varepsilon}^{-1/2} \right) x_{it} \widehat{\beta}_{OLS},$$

and

$$y_{it}^* = \widehat{\Omega}_{iu.\varepsilon}^{-1/2} \widehat{y}_{it}^{++}.$$

Assumption 7 $\Omega_{i\varepsilon}$ *is not singular for all* i*.*

Then, we define the FMOLS estimator for heterogeneous panels as

$$\widehat{\beta}_{FM}^* = \left[\sum_{i=1}^{N} \sum_{t=1}^{T} \left(x_{it}^* - \overline{x}_i^* \right) \left(x_{it}^* - \overline{x}_i^* \right)' \right]^{-1} \left[\sum_{i=1}^{N} \left(\sum_{t=1}^{T} \left(x_{it}^* - \overline{x}_i^* \right) y_{it}^* - T \widehat{\Delta}_{i\varepsilon u}^* \right) \right], \tag{24}$$

where

$$\widehat{\Delta}_{i\varepsilon u}^* = \widehat{\Omega}_{i\varepsilon}^{-1/2} \widehat{\Delta}_{i\varepsilon u}^+ \widehat{\Omega}_{iu.\varepsilon}^{-1/2}$$

and

$$\begin{aligned} \widehat{\Delta}_{i\varepsilon u}^+ &= \begin{pmatrix} \widehat{\Delta}_{i\varepsilon u} & \widehat{\Delta}_{i\varepsilon} \end{pmatrix} \begin{pmatrix} 1 \\ -\widehat{\Omega}_{i\varepsilon}^{-1} \widehat{\Omega}_{i\varepsilon u} \end{pmatrix} \\ &= \widehat{\Delta}_{i\varepsilon u} - \widehat{\Delta}_{i\varepsilon} \widehat{\Omega}_{i\varepsilon}^{-1} \widehat{\Omega}_{i\varepsilon u}. \end{aligned}$$

Theorem 4 *If Assumptions* $1-2$ *and* $6-7$ *hold, then* $\sqrt{N}T\left(\widehat{\beta}_{FM}^{*}-\beta\right)\Rightarrow \mathrm{N}\left(0,6\mathbf{I}_k\right).$

The DOLS estimator for heterogeneous panels, $\widehat{\beta}_{D}^{*}$, can be obtained by running the following regression:

$$y_{it}^{*} \;=\; \alpha_i \;+\; x_{it}^{*'}\beta \;+\; \sum_{j=-q_i}^{q_i} c_{ij}\triangle x_{it+j}^{*} + \dot{v}_{it}^{*}, \tag{25}$$

where $\dot{v}_{it}^{*}$ is defined similarly as in (18). Note that in (25) different lag truncations, q_i, may have to be used because the error terms are heterogeneous across i. Therefore, we need to assume that q_i tends to infinity with T at a suitable rate for all i:

Assumption 8 $q_i \to \infty$ *as* $T\to\infty$ *such that* $\frac{q_i^3}{T}\to 0$, *and*

$$T^{1/2}\sum_{|j|>q_i}\|c_{ij}\|\to 0 \tag{26}$$

for all i.

In the following theorem we show that $\widehat{\beta}_{D}^{*}$ also has the same limiting distribution as $\widehat{\beta}_{FM}^{*}$.

Theorem 5 *If Assumptions* $1-2$ *and* $6-8$ *hold, then* $\sqrt{N}T\left(\widehat{\beta}_{D}^{*}-\beta\right)\Rightarrow \mathrm{N}\left(0,6\mathbf{I}_k\right).$

Remark 6 *Theorems 4 and 5 show that the limiting distributions of* $\widehat{\beta}_{FM}^{*}$ *and* $\widehat{\beta}_{D}^{*}$ *are free of nuisance parameters.*

Remark 7 *We now consider a linear hypothesis that involves the elements of the coefficient vector* β. *We show that hypothesis tests constructed using the FMOLS and DOLS estimators have asymptotic chi-squared distributions. The null hypothesis has the form:*

$$H_0 : R\beta = r, \tag{27}$$

where r *is an* $m\times 1$ *known vector and* R *is a known* $m\times k$ *matrix describing the restrictions. A natural test statistic of the Wald test using* $\widehat{\beta}_{FM}$ *or* $\widehat{\beta}_{D}$ *for homogeneous panels is*

$$W=\frac{1}{6}NT^2\left(R\widehat{\beta}_D-r\right)'\left[R\widehat{\Omega}_{\varepsilon}^{-1}\widehat{\Omega}_{u.\varepsilon}R'\right]^{-1}\left(R\widehat{\beta}_D-r\right). \tag{28}$$

Remark 8 *For the heterogeneous panels, a natural statistic using* $\widehat{\beta}_{FM}^{*}$ *or* $\widehat{\beta}_{D}^{*}$ *to test the null hypothesis is*

$$W^{*}=\frac{1}{6}NT^2\left(R\widehat{\beta}_D^{*}-r\right)'\left[RR'\right]^{-1}\left(R\widehat{\beta}_D^{*}-r\right). \tag{29}$$

It is clear that W and W^* converge in distribution to a chi-squared random variable with m degrees of freedom, χ^2_m, as $T \to \infty$ and followed by $N \to \infty$ sequentially under the null hypothesis. Hence, we establish the following results:

$$W \Rightarrow \chi^2_m,$$

and

$$W^* \Rightarrow \chi^2_m.$$

Because the FMOLS and the DOLS estimators have the same asymptotic distributions, it is easy to verify that the Wald statistics based on the FMOLS estimator share the same limiting distributions as those based on the DOLS estimator.

5 Monte Carlo Simulations

The ultimate goal of this Monte Carlo study is to compare the sample properties of OLS, FMOLS, and DOLS for two models: a homogeneous panel and a heterogeneous panel. The simulations were performed by a Sun SparcServer 1000 and an Ultra Enterprise 3000. GAUSS 3.2.31 and COINT 2.0 were used to perform the simulations. Random numbers for error terms, $(u^*_{it}, \varepsilon^*_{it})$, for Sections 5.1, 5.2, and 5.4 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. $\{u^*_{it}\}$ and $\{\varepsilon^*_{it}\}$ were constructed with $u_{i0} = 0$ and $\varepsilon_{i0} = 0$.

In order to compare the performance of the OLS, FMOLS, and DOLS estimators, the following data generating process (DGP) was used:

$$y_{it} \quad = \quad \alpha_i \quad + \quad \beta x_{it} \quad + \quad u_{it} \tag{30}$$

and

$$x_{it} = x_{it-1} + \varepsilon_{it}.$$

where $(u_{it}, \varepsilon_{it})$ follows an ARMA(1,1) process:

$$\begin{pmatrix} u_{it} \\ \varepsilon_{it} \end{pmatrix} = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix} \begin{pmatrix} u_{it-1} \\ \varepsilon_{it-1} \end{pmatrix} + \begin{pmatrix} u^*_{it} \\ \varepsilon^*_{it} \end{pmatrix} + \begin{pmatrix} 0.3 & -0.4 \\ \theta_{21} & 0.6 \end{pmatrix} \begin{pmatrix} u^*_{it-1} \\ \varepsilon^*_{it-1} \end{pmatrix}$$

with

$$\begin{pmatrix} u^*_{it} \\ \varepsilon^*_{it} \end{pmatrix} \overset{iid}{\sim} \mathrm{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \sigma_{21} \\ \sigma_{21} & 1 \end{bmatrix} \right).$$

The design in (30) nests several important special cases. First, when $\begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix}$ is replaced by $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ and σ_{21} is constant across i, then the DGP becomes the homogeneous panel in Section 5.1. Second, when $\begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix}$ is replaced by $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, and θ_{21} and σ_{21} are random variable across i, then the DGP is the heterogeneous panel in Section 5.4.

5.1 Homogeneous Panel

To compare the performance of the OLS, FMOLS, and DOLS estimators for the homogeneous panel we conducted Monte Carlo experiments based on a design similar to that of Phillips and Hansen (1990) and Phillips and Loretan (1991).

$$y_{it} \quad = \quad \alpha_i \quad + \quad \beta x_{it} \quad + \quad u_{it}$$

and

$$x_{it} = x_{it-1} + \varepsilon_{it}$$

for $i = 1, ..., N$, $t = 1, ...T$, where

$$\begin{pmatrix} u_{it} \\ \varepsilon_{it} \end{pmatrix} = \begin{pmatrix} u^*_{it} \\ \varepsilon^*_{it} \end{pmatrix} + \begin{pmatrix} 0.3 & -0.4 \\ \theta_{21} & 0.6 \end{pmatrix} \begin{pmatrix} u^*_{it-1} \\ \varepsilon^*_{it-1} \end{pmatrix} \tag{31}$$

with

$$\begin{pmatrix} u^*_{it} \\ \varepsilon^*_{it} \end{pmatrix} \overset{iid}{\sim} \mathrm{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \sigma_{21} \\ \sigma_{21} & 1 \end{bmatrix} \right).$$

We generated α_i from a uniform distribution, $U[0, 10]$, and set $\beta = 2$. From Theorems 1-3 we know that the asymptotic results depend upon variances and covariances of the errors u_{it} and ε_{it}. The design in (31) is a good one since the endogeneity of the system is controlled by only two parameters, θ_{21} and σ_{21}. We allowed θ_{21} and σ_{21} to vary and considered values of $\{0.8, 0.4, 0.0, -0.8\}$ for θ_{21} and $\{-0.8, -0.4, 0.4\}$ for σ_{21}.

The estimate of the long-run covariance matrix in (15) was obtained by using the procedure KERNEL in COINT 2.0 with a Bartlett window. The lag truncation number was set arbitrarily at five. Results with other kernels, such as Parzen and quadratic spectral kernels, are not reported, because no essential differences were found for most cases.

Next, we recorded the results from our Monte Carlo experiments that examined the finite-sample properties of the OLS estimator, $\widehat{\beta}_{OLS}$; the FMOLS estimator, $\widehat{\beta}_{FM}$; and the DOLS estimator, $\widehat{\beta}_{D}$. The results we

report are based on $10,000$ replications and are summarized in Tables 1 - 4 and Figures 1 - 8. The FMOLS estimator was obtained by using a Bartlett window of lag length five as in (15). Four lags and two leads were used for the DOLS estimator.

Table 1 reports the Monte Carlo means and standard deviations (in parentheses) of $\left(\widehat{\beta}_{OLS}-\beta\right)$, $\left(\widehat{\beta}_{FM}-\beta\right)$, and $\left(\widehat{\beta}_{D}-\beta\right)$ for sample sizes $T=N=(20,40,60)$. The biases of the OLS estimator, $\widehat{\beta}_{OLS}$, decrease at a rate of T. For example, with $\sigma_{21}=-0.8$ and $\theta_{21}=0.8$, the bias at $T=20$ is -0.201 and at $T=40$ is -0.104. Also, the biases increase in θ_{21} (with $\theta_{21}>0$) and decrease in σ_{21}.

While we expected the OLS estimator to be biased, we expected the FMOLS estimator would produce much better estimates. However, it is noticeable that the FMOLS estimator has a downward bias when $\theta_{21}\geq 0$ and an upward bias when $\theta_{21}<0$. In general, the FMOLS estimator, $\widehat{\beta}_{FM}$, presents the same degree of difficulty with bias as does the OLS estimator, $\widehat{\beta}_{OLS}$. For example, while the FMOLS estimator, $\widehat{\beta}_{FM}$, reduces the bias substantially and outperforms $\widehat{\beta}_{OLS}$ when $\theta_{21}>0$ and $\sigma_{21}<0,$ the opposite is true when $\theta_{21}>0$ and $\sigma_{21}>0$. Likewise, when $\theta_{21}=-0.8$, $\widehat{\beta}_{FM}$ is less biased than $\widehat{\beta}_{OLS}$ for values of $\sigma_{21}=-0.8$. Yet, for values of $\sigma_{21}=-0.4$, the bias in $\widehat{\beta}_{OLS}$ is less than the bias in $\widehat{\beta}_{FM}$. There seems to be little to choose between $\widehat{\beta}_{OLS}$ and $\widehat{\beta}_{FM}$ when $\theta_{21}<0$. This is probably due to the failure of the nonparametric correction procedure in the presence of a negative serial correlation of the errors, i.e., a negative MA value, $\theta_{21}<0$. Finally, for the cases where $\theta_{21}=0.0$, $\widehat{\beta}_{FM}$ outperforms $\widehat{\beta}_{OLS}$ when $\sigma_{21}<0$. On the other hand, $\widehat{\beta}_{FM}$ is more biased than $\widehat{\beta}_{OLS}$ when $\sigma_{21}>0$.

In contrast, the results in Table 1 show that the DOLS, $\widehat{\beta}_{D}$, is distinctly superior to the OLS and FMOLS estimators for all cases in terms of the mean biases. It was noticeable that the FMOLS leads to a significant bias. Clearly, the DOLS outperformed both the OLS and FMOLS estimators. The FMOLS estimator is also complicated by the dependence of the correction in (11) and (12) upon the preliminary estimator (here we use OLS), which may be biased in finite samples. The DOLS differs from the FMOLS estimator in that the DOLS requires no initial estimation and no nonparametric correction.

It is important to know the effects of the variations in panel dimensions on the results, since the actual panel data have a wide variety of cross-section and time-series dimensions. Table 2 considers 20 different combinations for N and T, each ranging from 20 to 120 with $\sigma_{21}=-0.4$ and $\theta_{21}=0.4$. First, we notice that the cross-section dimension has a significant effect on the biases of $\widehat{\beta}_{OLS}$, $\widehat{\beta}_{FM}$, and $\widehat{\beta}_{D}$ when N is increased from 1 to 20. However, when N is increased from 20 to 40 and beyond, there is little effect on the biases of $\widehat{\beta}_{OLS}$, $\widehat{\beta}_{FM}$, and $\widehat{\beta}_{D}$. From this it seems that in practice the T dimension must exceed the N dimension, especially for the OLS and FMOLS estimators, in order to get a good approximation of the limiting distributions of the estimators. For example, for each of the estimators in Table 2, the reported

bias is substantially less for $(T = 120, N = 40)$ than it is for either $(T = 40, N = 40)$ or $(T = 40, N = 120)$. The results in Table 2 again confirm the superiority of the DOLS. The largest bias in the DOLS with four lags and two leads, DOLS(4,2), is less than or equal to 0.02 for all cases except at $N = 1$ and $T = 20$, which can be compared with a simulation standard error (in parentheses) that is less than .007 when $N \geq 20$ and $T \geq 60$, confirming the accuracy of the DOLS(4,2). The biases in DOLS with two lags and one lead, DOLS(2,1) start off slightly biased at $N = 1$ and $T = 20$, and converge to an almost unbiased coefficient estimate at $N = 20$ and $T = 40$. The biases of DOLS(2,1) move in the opposite direction to those of DOLS(4,2).

Figures 1, 3, 5 and 7 display estimated pdfs for the estimators for $\sigma_{21} = -0.4$ and $\theta = 0.4$ with $N = 40$ ($T = 20$ in Figure 1, $T = 40$ in Figure 3, $T = 60$ in Figure 5 and $T = 120$ in Figure 7). In Figure 1, $N = 40$, $T = 20$, the DOLS is much better centered than the OLS and FMOLS. In Figures 3, 5 and 7, the biases of the OLS and FMOLS were reduced as T increases, the DOLS still dominates the OLS and FMOLS.

Monte Carlo means and standard deviations of the t-statistic, $t_{\beta=\beta_0}$, are given in Table 3. Here, the OLS t-statistic is the conventional t-statistic as printed by standard statistical packages, and the FMOLS and DOLS t-statistics. With all values of θ_{21} and σ_{21}, the DOLS(4,2) t-statistic is well approximated by a standard N(0,1) suggested from the asymptotic results. The DOLS(4,2) t-statistic is much closer to the standard normal density than the OLS t-statistic and the FMOLS t-statistic. When $\theta_{21} > 0$ and $\sigma_{21} < 0$, the OLS t-statistic is more heavily biased than the FMOLS t-statistic. Again, when $\theta_{21} > 0$ and $\sigma_{21} > 0$, the opposite is true. Even when $\theta_{21} = 0$, the FMOLS t-statistic is not well approximated by a standard N(0,1). The OLS t-statistic performs better than the FMOLS t-statistic when $\sigma_{21} = 0.8$ and $\theta_{21} > 0$ and when $\sigma_{21} \leq -0.4$ and $\theta_{21} = -0.8$, but not in other cases. The FMOLS t-statistic in general does not perform better than the OLS t-statistic.

Table 4 shows that both the OLS t-statistic and the FMOLS t-statistic become more negatively biased as the dimension of cross-section N increases. The heavily negative biases of the FMOLS t-statistic in Tables 3-4 again indicate the poor performance of the FMOLS estimator. For the DOLS(4,2), the biases decrease rapidly and the standard errors converge to 1.0 as T increases. Similar to Table 2, we observe from Table 4 that for the DOLS t-statistic the T dimension is more important than the N dimension in reducing the biases of the t-statistics. However, the improvement of the DOLS t-statistic is rather marginal as T increases.

Figures 2, 4, 6 and 8 display estimated pdfs for the t-statistics for $\sigma_{21} = -0.4$ and $\theta = 0.4$ with $N = 40$ ($T = 20$ in Figure 2, $T = 40$ in Figure 4, $T = 60$ in Figure 6 and $T = 120$ in Figure 8). The figures show clearly that the DOLS t-statistic is well approximated by a standard N(0,1) especially as T increases. From the results in Tables 2 and 4 and Figures 1-8 we note that the sequential limit theory approximates the

limiting distributions of the DOLS and its t-statistic very well.

It is known that when the length of time series is short the estimate $\widehat{\Omega}$ in (15) may be sensitive to the length of the bandwidth. In Tables 2 and 4, we first investigate the sensitivity of the FMOLS estimator with respect to the choice of length of the bandwidth. We extend the experiments by changing the lag length from 5 to 2 for a Barlett window. Overall, the results show that changing the lag length from 5 to 2 does not lead to substantial changes in biases for the FMOLS estimator and its t-statistic. However, the biases of the DOLS estimator and its t-statistic are reduced substantially when the lags and leads are changed from $(2,1)$ to $(4,2)$ as predicted from Theorem 3. The results from Tables 2 and 4 show that the DOLS method gives different estimates of β and the t-statistic depending on the number of lags and leads we choose. This seems to be a drawback of the DOLS estimator. Further research is needed on how to choose the lags and leads for the DOLS estimator in the panel setting.

5.2 ARMA(1,1) Error Terms

In this section, we look at simulations where, instead of the errors being generated by an MA(1) process in (31), the errors are generated by an ARMA(1,1) process in (30). As pointed out by a referee, the MA(1) specification in (31) may be unfair to the FMOLS estimator. One of the reasons why the performance of the DOLS is much better than that of the FMOLS lies in the simulation design in (31), which assumes that the error terms are MA(1) processes. If $(u_{it}\ ,\varepsilon_{it})^{'}$ is an MA(1) process, then u_{it} can be written exactly with three terms, $\varepsilon_{it-1}, \varepsilon_{it}$, and ε_{it+1} and no lag truncation approximation is required for the DOLS.

Tables 5 and 6 report the performance of OLS, FMOLS, and DOLS and their t-statistics when the errors are generated by an ARMA(1,1) process. Tables 5 and 6 show that the FMOLS estimator and its t-statistic are less biased than the OLS estimator for most cases and is outperformed by the DOLS. Again, when $\theta_{21} \geq 0.0$ and $\sigma_{21} = 0.8$ the FMOLS estimator and its t-statistic suffer from severe biases. On the other hand, we observe that DOLS shows less improvement compared with OLS and FMOLS, in contrast to Tables 1 and 3. This observation is correctly pointed out by a referee. However, the good performance of DOLS may disappear for high order ARMA(p.q) error process.

5.3 Non-normal Errors

In this section, we conduct an experiment where the error terms are non-normal. The DGP is similar to that of Gonzalo (1994):

$$\begin{pmatrix} u_{it} \\ \varepsilon_{it} \end{pmatrix} = \begin{pmatrix} u_{it}^* \\ \varepsilon_{it}^* \end{pmatrix} + \begin{pmatrix} 0.3 & -0.4 \\ \theta_{21} & 0.6 \end{pmatrix} \begin{pmatrix} u_{it-1}^* \\ \varepsilon_{it-1}^* \end{pmatrix}, \tag{32}$$

$$u_{it}^* = \left(\frac{1}{\sigma}\right) 0.5\varepsilon_{it}^* + \left(1 - 0.5^2\right)^{1/2} u_{it}^{**},$$

and

$$\varepsilon_{it}^* = \sigma\varepsilon_{it}^{**},$$

where u_{it}^{**} and ε_{it}^{**} are independent exponential random variables with a parameter 1. The results from Tables 7-8 show that while the DOLS estimator performs better in terms of the biases, the distribution of the DOLS t-statistic is far from the asymptotic N(0,1). The standard deviations of the DOLS t-statistic are badly underestimated.

To summarize the results so far, it would appear that the DOLS estimator is the best estimator overall, though the standard error for the DOLS t-statistic shows significant downward bias when the error terms are generated from non-normal distributions.

5.4 Heterogeneous Panel

In Sections 5.1-5.3 we compare the small sample properties of the OLS, FMOLS and DOLS estimators and conclude that the DOLS estimator and its t-statistic generally exhibit the least bias. One of the reasons for the poor performance of the FMOLS estimator in the homogeneous panel is that the FMOLS estimator needs to use a kernel estimator for the asymptotic covariance matrix, while the DOLS does not. By contrast, for the heterogeneous panel both DOLS in (20) and OLS in (33) use kernel estimators. Consequently, one may expect that the much better performance of the DOLS estimator in Sections 5.1-5.3 is limited to only very specialized cases, e.g., in the homogeneous panel. To test this, we now compare the performance of the OLS, FMOLS, and DOLS estimators for a heterogeneous panel using Monte Carlo experiments similar to those in Section 5.1. The DGP is

$$y_{it} \quad = \quad \alpha_i \quad + \quad \beta x_{it} \quad + \quad u_{it}$$

and

$$x_{it} = x_{it-1} + \varepsilon_{it}$$

for $i = 1, ..., N$, $t = 1, ...T$, where

$$\begin{pmatrix} u_{it} \\ \varepsilon_{it} \end{pmatrix} = \begin{pmatrix} u_{it}^* \\ \varepsilon_{it}^* \end{pmatrix} + \begin{pmatrix} 0.3 & -0.4 \\ \theta_{21} & 0.6 \end{pmatrix} \begin{pmatrix} u_{it-1}^* \\ \varepsilon_{it-1}^* \end{pmatrix}$$

with

$$\begin{pmatrix} u_{it}^* \\ \varepsilon_{it}^* \end{pmatrix} \overset{iid}{\sim} \mathrm{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \sigma_{21} \\ \sigma_{21} & 1 \end{bmatrix} \right).$$

As in Section 5.1, we generated α_i from a uniform distribution, $U[0, 10]$, and set $\beta = 2$. In this section, we allowed θ_{21} and σ_{21} to be random in order to generate the heterogeneous panel, i.e., both θ_{21} and σ_{21} are generated from a uniform distribution, $U[-0.8, 0.8]$. An estimate of $\Omega_i = \Sigma_i + \Gamma_i + \Gamma_i^{'}$, $\widehat{\Omega}_i$, was obtained by the COINT 2.0 with a Bartlett window. The lag truncation number was set at 5.

The three estimators considered are the FMOLS, DOLS, and the OLS, where the OLS is defined as

$$\widehat{\beta}_{OLS}^* = \left[\sum_{i=1}^{N} \sum_{t=1}^{T} \left(x_{it}^{**} - \overline{x}_i^{**} \right) \left(x_{it}^{**} - \overline{x}_i^{**} \right)^{'} \right]^{-1} \left[\sum_{i=1}^{N} \sum_{t=1}^{T} \left(x_{it}^{**} - \overline{x}_i^{**} \right) \left(y_{it}^{**} \right) \right] \tag{33}$$

with $x_{it}^{**} = w_i x_{it}$, $y_{it}^{**} = w_i y_{it}$, $\overline{x}_i^{**} = \frac{1}{T} \sum_{t=1}^{T} x_{it}^{**}$, and $w_i = \left[\widehat{\Omega}_i^{-1} \right]_{11}$. Two FMOLS estimators will be considered, one using the lag length of 5 (FMOLS(5)), the second using the lag length of 2 (FMOLS(2)). Two DOLS estimators are also considered: DOLS with four lags and two leads, DOLS(4,2) and DOLS with two lags and one leads, DOLS(2,1). The relatively good performance of the DOLS estimator in a homogeneous panel can also be observed in Table 9. The biases of the OLS and FMOLS estimators are substantial. Again, the DOLS outperforms the OLS and FMOLS. Note from Table 9 that the FMOLS always has more bias than the OLS for all N and T except when $N = 1$. The poor performance of the FMOLS in the heterogenous panels indicates that the FMOLS in Section 4 is not recommended in practice. A possible reason for the poor performance of the FMOLS in heterogenous panels is that it has to go through two nonparametric corrections, as in (22) and (23). Therefore the failure of the nonparametric correction could be very severe for the FMOLS estimator in heterogenous panels. Pedroni (1996) proposed several alternative versions of the FMOLS estimator such as an FMOLS estimator based on the transformation of the estimated residuals and a group-mean based FMOLS estimator. It would be interesting to study further the issues of estimation and inference in heterogenous panels. However, it goes beyond the scope of this paper.

From Table 10, we note that the DOLS t-statistics tend to have heavier tails than predicted by the asymptotic distribution theory, though the bias of the DOLS t-statistic is much lower than those of the OLS and FMOLS t-statistics.

It appears that the DOLS still is the best estimator overall in a heterogeneous panel.

5.5 Summary

The results from the above four Monte Carlo simulations can be summarized as follows: First, for the homogeneous panel, when the serial correlation parameter, θ_{21}, and the endogeneity parameter, σ_{21}, are both negative, the OLS is the most biased estimator. The OLS is biased in almost all cases for the heterogenous panel. Second, the FMOLS is more biased than the OLS when $\theta_{21} \geq 0$ and $\sigma_{21} > 0$ for the homogeneous panel. The FMOLS is severely biased for the heterogenous panel in almost all trials. This indicates the failure of the parametric correction is very serious, especially in the heterogenous panel. Third, DOLS performs very well in all cases for both the homogeneous and heterogenous panels. Adding the number of leads and lags reduces the bias of the DOLS substantially. This was predicted by the asymptotic theory in Theorem 3. Fourth, the sequential limit theory approximates the limit distributions of the DOLS and its t-statistic very well.

6 Conclusion

This paper discusses limiting distributions for the OLS, FMOLS, and DOLS estimators in a cointegrated regression. We also investigate the finite sample proprieties of the OLS, FMOLS, and DOLS estimators. Our findings are summarized as follows:

(i) The OLS estimator has a non-negligible bias in finite samples.

(ii) The FMOLS estimator does not improve over the OLS estimator in general.

(iii) The DOLS estimator may be more promising than the OLS or FMOLS estimators in estimating cointegrated panel regressions.

Appendix

A Proof of Theorem 3

First we write (19) in vector form:

$$\begin{aligned} y_i &= e\alpha_i + x_i\beta + Z_{iq}C + \dot{v}_i \\ &= x_i\beta + Z_iD + \dot{v}_i \quad \text{(say)}, \end{aligned}$$

where y_i is a $T \times 1$ vector of y_{it}; e a is $T \times 1$ unit vector; Z_{iq} is the $T \times 2q$ matrix of observations on the $2 \times q$ regressors $\triangle x_{it-q}, \cdots, \triangle x_{it+q}$; x_i is a vector of $T \times k$ of x_{it}; C is a $(2 \times q) \times 1$ vector of c_{ij}; $\dot{v}_i$ is a $T \times 1$ vector of $\dot{v}_{it}$; Z_i is a $T \times (2 \times q + 1)$ matrix, $Z_i = (e, Z_{iq})$; and D is a $(2 \times q + 1) \times 1$ vector of parameters. Let $Q_i = I - Z_i(Z_i' Z_i)^{-1} Z_i^{'}$. It follows that

$$\left(\widehat{\beta}_D - \beta\right) = \left[\sum_{i=1}^{N} \left(x_i^{'} Q_i x_i\right)\right]^{-1} \left[\sum_{i=1}^{N} \left(x_i' Q_i \ \dot{v}_i\right)\right].$$

We rescale $\left(\widehat{\beta}_D - \beta\right)$ by $\sqrt{N}T$ to get

$$\begin{aligned}
&\sqrt{N}T\left(\widehat{\beta}_D - \beta\right) \\
&= \left[\tfrac{1}{N}\sum_{i=1}^{N} \tfrac{1}{T^2}\left(x_i' Q_i x_i\right)\right]^{-1} \left[\sqrt{N}\tfrac{1}{N}\sum_{i=1}^{N} \tfrac{1}{T}\left(x_i' Q_i \ \dot{v}_i\right)\right] \\
&= \left[\tfrac{1}{N}\sum_{i=1}^{N} \zeta_{6iT}\right]^{-1} \left[\sqrt{N}\tfrac{1}{N}\sum_{i=1}^{N} \zeta_{5iT}\right] \\
&= \left[\xi_{6NT}\right]^{-1} \left[\sqrt{N}\xi_{5NT}\right],
\end{aligned}$$

where $\xi_{5NT} = \frac{1}{N}\sum_{i=1}^{N} \zeta_{5iT}$, $\zeta_{5iT} = \frac{1}{T}\left(x_i' Q_i \ \dot{v}_i\right)$, $\xi_{6NT} = \frac{1}{N}\sum_{i=1}^{N} \zeta_{6iT}$, and $\zeta_{6iT} = \frac{1}{T^2}\left(x_i' Q_i x_i\right)$.

Observe that from Saikkonen (1991)

$$\begin{aligned}
\zeta_{6iT} &= \frac{1}{T^2}\left(x_i' Q_i x_i\right) \\
&= \frac{1}{T^2}\left(x_i^{'} W_T x_i\right) + o_p(1) \\
&= \frac{1}{T^2}\sum_{t=q+1}^{T-q} \left(x_{it} - \overline{x}_i\right)\left(x_{it} - \overline{x}_i\right)^{'} + o_p(1) \\
&\Rightarrow \int \widetilde{B}_{\varepsilon i} \widetilde{B}_{\varepsilon i}^{'},
\end{aligned}$$

and

$$\begin{aligned}
\zeta_{5iT} &= \frac{1}{T}\left(x_i' Q_i \ \dot{v}_i\right) \\
&= \frac{1}{T}\left(x_i^{'} W_T \ \dot{v}_i\right) + o_p(1) \\
&= \frac{1}{T}\sum_{t=q+1}^{T-q} \left(x_{it} - \overline{x}_i\right) \dot{v}_{it} + o_p(1) \\
&\Rightarrow \int \widetilde{B}_{\varepsilon i} dB_{ui}^{+},
\end{aligned}$$

as $T \to \infty$ for all i, where $\widetilde{B}_{\varepsilon i} = B_{\varepsilon i} - \int B_{\varepsilon i}$ and $W_T = I_T - \frac{1}{T} e e^{'}$. Then applying the multivariate Lindeberg-Levy central limit theorem to $\frac{1}{\sqrt{N}} \int \widetilde{B}_{\varepsilon i} dB_{ui}^{+}$ and combining this with the limit of $\frac{1}{N}\sum_{i=1}^{N} \int \widetilde{B}_{\varepsilon i} \widetilde{B}_{\varepsilon i}^{'}$

as in Theorem 2, we have

$$\left[\frac{1}{N}\sum_{i=1}^{N}\int \widetilde{B}_{\varepsilon i}\widetilde{B}_{\varepsilon i}^{'}\right]^{-1}\left[\frac{1}{\sqrt{N}}\int \widetilde{B}_{\varepsilon i}dB_{ui}^{+}\right]\Rightarrow \mathrm{N}\left(0,6\Omega_{\varepsilon}^{-1}\Omega_{u.\varepsilon}\right)$$

as $N\to\infty$. It follows that using the sequential limit theory

$$\sqrt{N}T\left(\widehat{\beta}_{D}-\beta\right)\Rightarrow \mathrm{N}\left(0,6\Omega_{\varepsilon}^{-1}\Omega_{u.\varepsilon}\right)$$

as required. ■

B Proof of Theorem 5

The proof is the same as that of Theorem 3. First, similar to Theorem 3, we write (25) in vector form:

$$\begin{aligned} y_i^* &= e\alpha_i + x_i^*\beta + Z_{iq}^*C + \dot{v}_i^* \\ &= x_i^*\beta + Z_i^*D + \dot{v}_i^* \quad \text{(say)}, \end{aligned}$$

and define y_i^*, e, Z_{iq}^*, x_i^*, C, $\dot{v}_i^*$ Z_i^*, Z_i, D, and Q_i^* as in the proof of Theorem 3. Then we have:

$$\begin{aligned} &\sqrt{N}T\left(\widehat{\beta}_D^* - \beta\right) \\ &= \left[\tfrac{1}{N}\sum_{i=1}^{N}\tfrac{1}{T^2}\left(x_i'^*Q_i^*x_i^*\right)\right]^{-1}\left[\sqrt{N}\tfrac{1}{N}\sum_{i=1}^{N}\tfrac{1}{T}\left(x_i^{*\prime}Q_i^*\ \dot{v}_i^*\right)\right] \\ &= \left[\tfrac{1}{N}\sum_{i=1}^{N}\zeta_{8iT}\right]^{-1}\left[\sqrt{N}\tfrac{1}{N}\sum_{i=1}^{N}\zeta_{7iT}\right] \\ &= \left[\xi_{8NT}\right]^{-1}\left[\sqrt{N}\xi_{7NT}\right], \end{aligned}$$

where $\xi_{7NT} = \frac{1}{N}\sum_{i=1}^{N}\zeta_{7iT}$, $\zeta_{7iT} = \frac{1}{T}\left(x_i^{*\prime}Q_i^*\ \dot{v}_i^*\right)$, $\xi_{8NT} = \frac{1}{N}\sum_{i=1}^{N}\zeta_{8iT}$, and $\zeta_{8iT} = \frac{1}{T^2}\left(x_i^{*\prime}Q_i^*x_i^*\right)$.

Observe that from Assumption 8, we have

$$\begin{aligned} \zeta_{8iT} &= \frac{1}{T^2}\left(x_i^{*\prime}Q_i^*x_i^*\right) \\ &= \frac{1}{T^2}\left(x_i^{*'}W_T^*x_i^*\right) + o_p(1) \\ &= \frac{1}{T^2}\sum_{t=q_i+1}^{T-q_i}\left(x_{it}^* - \overline{x}_i^*\right)\left(x_{it}^* - \overline{x}_i^*\right)^{'} + o_p(1) \\ &\Rightarrow \int \widetilde{W}_i\widetilde{W}_i, \end{aligned}$$

and

$$\zeta_{7iT} = \frac{1}{T}\left(x_i^{*\prime}Q_i^*\ \dot{v}_i^*\right)$$

$$
\begin{aligned}
&= \frac{1}{T}\left(x_i^{*\prime} W_T \, \dot{v}_i^*\right) + o_p(1) \\
&= \frac{1}{T}\sum_{t=q_i+1}^{T-q_i} (x_{it}^* - \overline{x}_i^*) \, \dot{v}_{it}^* + o_p(1) \\
&\Rightarrow \int \widetilde{W}_i dV_i,
\end{aligned}
$$

as $T \to \infty$ for all i. The remainder of the proof follows that of Theorem 3. ■

References

[1] Baltagi, B. (1995), *Econometric Analysis of Panel Data*, New York, John Wiley & Sons.

[2] Billingsley, P. (1986), *Probability and Measure*, John Wiley & Sons, New York.

[3] Breitung, J., and Meyer, W. (1994), "Testing for Unit Roots in Panel Data: Are Wages on Different Bargaining Levels Cointegrated?" *Applied Economics*, 26, 353-361.

[4] Chen, B., McCoskey, S., and Kao, C. (1996), "Estimation and Inference of a Cointegrated Regression in Panel Data: A Monte Carlo Study," Manuscript, Center for Policy Research, Syracuse University.

[5] Gonzalo, J. (1994), "Five Alternative Methods of Estimating Long-Run Equilibrium Relationships," *Journal of Econometrics*, 60, 203-233.

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

[7] Im, K., Pesaran, H., and Shin, Y. (1995), "Testing for Unit Roots in Heterogeneous Panels," Manuscript, University of Cambridge.

[8] Kao, C. (1997), "Spurious Regression and Residual-Based Tests for Cointegration in Panel Data," *Journal of Econometrics*, forthcoming.

[9] Kao, C., and Chen, B. (1995), "On the Estimation and Inference for Cointegration in Panel Data When the Cross-Section and Time-Series Dimensions are Comparable," Manuscript, Center for Policy Research, Syracuse University.

[10] Levin, A., and Lin, C.-F. (1993), "Unit Root Tests in Panel Data: New Results," Discussion Paper, Department of Economics, UC-San Diego.

[11] Maddala, G. S., and Wu, S. (1996), "A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test: Evidence From Simulations and the Bootstrap," Manuscript, Department of Economics, Ohio State University.

[12] McCoskey, S., and Kao, C. (1998), "A Residual-Based Test of the Null of Cointegration in Panel Data," *Econometric Reviews*, 17, 57-84.

[13] Pesaran, H., and Smith, R. (1995), "Estimating Long-Run Relationships from Dynamic Heterogeneous Panels," *Journal of Econometrics*, 68, 79-113.

[14] Pedroni, P. (1997), "Panel Cointegration: Asymptotics and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP Hypothesis," Indiana University Working Papers in Economics, No. 95-013.

[15] Pedroni, P. (1996), "Fully Modified OLS for Heterogeneous Cointegrated Panels and the Case of Purchasing Power Parity," Indiana University Working Papers in Economics, No. 96-20.

[16] Phillips, P. C. B. (1995), "Fully Modified Least Squares and Vector Autoregression," *Econometrica*, 63, 1023-1078.

[17] Phillips, P. C. B., and Durlauf, S. N. (1986), "Multiple Time Series Regression with Integrated Processes," *Review of Economic Studies*, 53, 473-495.

[18] Phillips, P. C. B., and Hansen, B. E. (1990), "Statistical Inference in Instrumental Variables Regression with I(1) Processes," *Review of Economic Studies*, 57, 99-125.

[19] Phillips, P. C. B., and Loretan, M. (1991), "Estimating Long-Run Economic Equilibria," *Review of Economic Studies*, 58, 407-436.

[20] Phillips, P. C. B., and Moon, H. (1998), "Linear Regression Limit Theory for Nonstationary Panel Data," Cowles Foundation Discussion Paper.

[21] Phillips, P. C. B., and Solo, V. (1992), "Asymptotics for Linear Processes," *Annals of Statistics*, 20, 971-1001.

[22] Quah, D. (1994), "Exploiting Cross Section Variation for Unit Root Inference in Dynamic Data," *Economics Letters*, 44, 9-19.

[23] Saikkonen, P. (1991), "Asymptotically Efficient Estimation of Cointegrating Regressions," *Econometric Theory*, 58, 1-21.

[24] 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.

[25] Stock, J., and Watson, M. (1993), "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems," *Econometrica*, 61, 783-820.

Table 1: Means Biases and Standard Deviations of OLS, FMOLS, and DOLS Estimators

	$\sigma_{21} = -0.8$			$\sigma_{21} = -0.4$			$\sigma_{21} = 0.8$		
	$\widehat{\beta}_{OLS}-\beta$	$\widehat{\beta}_{FM}-\beta$	$\widehat{\beta}_{D}-\beta$	$\widehat{\beta}_{OLS}-\beta$	$\widehat{\beta}_{FM}-\beta$	$\widehat{\beta}_{D}-\beta$	$\widehat{\beta}_{OLS}-\beta$	$\widehat{\beta}_{FM}-\beta$	$\widehat{\beta}_{D}-\beta$
θ_{21}= 0.8									
T = 20	-.201	-.176	-.001	-.097	-.113	-.002	-.022	-.069	-.009
	(.049)	(.044)	(.040)	(.032)	(.035)	(.033)	(.011)	(.016)	(.009)
T = 40	-.104	-.099	-.000	-.049	-.062	-.001	-.011	-.036	-.004
	(.019)	(.017)	(.013)	(.012)	(.013)	(.011)	(.004)	(.006)	(.003)
T = 60	-.070	-.069	-.000	-.033	-.042	-.000	-.007	-.024	-.003
	(.010)	(.009)	(.007)	(.007)	(.007)	(.006)	(.002)	(.003)	(.002)
θ_{21}= 0.4									
T = 20	-.132	-.064	-.001	-.082	-.068	-.002	-.014	-.073	-.003
	(.038)	(.025)	(.027)	(.030)	(.029)	(.031)	(.013)	(.018)	(.013)
T = 40	-.066	-.038	-.001	-.041	-.038	-.001	-.007	-.037	-.001
	(.014)	(.009)	(.027)	(.011)	(.011)	(.009)	(.005)	(.006)	(.004)
T = 60	-.044	-.027	-.000	-.027	-.026	-.001	-.005	-.025	-.001
	(.007)	(.005)	(.005)	(.006)	(.006)	(.005)	(.002)	(.003)	(.002)
θ_{21}= 0.0									
T = 20	-.079	-.002	.001	-.059	-.019	.002	.005	-.069	.006
	(.027)	(.015)	(.017)	(.026)	(.022)	(.026)	(.016)	(.021)	(.017)
T = 40	-.039	-.005	.001	-.029	-.012	.001	.002	-.035	.003
	(.009)	(.005)	(.005)	(.009)	(.008)	(.008)	(.006)	(.007)	(.005)
T = 60	-.026	-.004	.000	-.019	-.009	-.001	.001	-.023	.002
	(.005)	(.003)	(.003)	(.005)	(.004)	(.008)	(.003)	(.004)	(.003)
θ_{21}= −0.8									
T = 20	-.029	.038	.007	-.019	.036	.007	.114	.012	.000
	(.016)	(.012)	(.008)	(.017)	(.015)	(.014)	(.034)	(.028)	(.031)
T = 40	-.015	.018	.003	-.009	.018	.003	.057	.011	-.000
	(.006)	(.004)	(.002)	(.006)	(.005)	(.004)	(.012)	(.009)	(.009)
T = 60	-.009	.011	.002	-.007	.012	.002	.038	.010	.000
	(.003)	(.002)	(.001)	(.003)	(.002)	(.002)	(.007)	(.005)	(.005)

Note:
(a) N =T.
(b) A lag length 5 of the Bartlett windows is used for the FMOLS estimator.
(c) 4 lags and 2 leads are used for the DOLS estimator.

Table 2: Means Biases and Standard Deviations of OLS, FMOLS, and DOLS Estimators for Different N and T

(N,T)	$\widehat{\beta}_{OLS}-\beta$	$\widehat{\beta}_{FM(5)}-\beta$	$\widehat{\beta}_{FM(2)}-\beta$	$\widehat{\beta}_{D(4,2)}-\beta$	$\widehat{\beta}_{D(2.1)}-\beta$
(1,20)	-.135	-.104	-.122	-.007	.031
	(.184)	(.196)	(.189)	(.297)	(.211)
(1,40)	-.070	-.059	-.065	-.001	.015
	(.093)	(.012)	(.092)	(.106)	(.090)
(1,60)	-.047	-.041	-.043	-.001	.009
	(.063)	(.064)	(.061)	(.064)	(.057)
(1,120)	-.024	-.023	-.022	-.001	.004
	(.032)	(.031)	(.031)	(.029)	(.027)
(20,20)	-.082	-.068	-.075	-.002	.017
	(.030)	(.029)	(.029)	(.031)	(.028)
(20,40)	-.042	-.039	-.039	-.001	.008
	(.016)	(.015)	(.015)	(.015)	(.014)
(20,60)	-.028	-.027	-.026	-.000	.006
	(.010)	(.010)	(.009)	(.009)	(.009)
(20,120)	-.014	-.014	-.013	-.000	.003
	(.005)	(.005)	(.005)	(.005)	(.004)
(40,20)	-.081	-.066	-.073	-.001	.017
	(.022)	(.021)	(.021)	(.022)	(.019)
(40,40)	-.041	-.038	-.038	-.001	.008
	(.011)	(.011)	(.011)	(.009)	(.009)
(40,60)	-.028	-.026	-.025	-.001	.005
	(.007)	(.007)	(.007)	(.007)	(.006)
(40,120)	-.014	-.014	-.013	-.000	.003
	(.004)	(.004)	(.003)	(.003)	(.004)
(60,20)	-.080	-.067	-.073	-.002	.016
	(.017)	(.017)	(.017)	(.018)	(.016)
(60,40)	-.041	-.038	-.038	-.001	.008
	(.009)	(.009)	(.009)	(.008)	(.008)
(60,60)	-.027	-.026	-.025	-.001	.005
	(.006)	(.006)	(.006)	(.005)	(.005)
(60,120)	-.014	-.014	-.012	-.000	.003
	(.003)	(.003)	(.003)	(.003)	(.003)
(120,20)	-.079	-.066	-.072	-.002	.016
	(.012)	(.012)	(.012)	(.012)	(.011)
(120,40)	-.041	-.037	-.037	-.001	.008
	(.006)	(.006)	(.006)	(.006)	(.005)
(120,60)	-.027	-.026	-.025	-.001	.005
	(.004)	(.004)	(.004)	(.004)	(.004)
(120,120)	-.014	-.014	-.013	-.000	.003
	(.002)	(.002)	(.002)	(.002)	(.002)

Note:
(a) A lag length 5 and 2 of the Bartlett windows are used for the FMOLS(5) and FMOLS(2) estimators.
(b) 4 lags and 2 leads and 2 lags and 1 lead are used for the DOLS(4,2) and DOLS(2,1) estimators.
(c) $\sigma_{21} = -0.4$ and $\theta_{21} = 0.4$.

Table 3: Means Biases and Standard Deviations of t-statistics

	$\sigma_{21} = -0.8$			$\sigma_{21} = -0.4$			$\sigma_{21} = 0.8$		
	OLS	FMOLS	DOLS	OLS	FMOLS	DOLS	OLS	FMOLS	DOLS
θ_{21}= 0.8									
T = 20	-7.247	-5.594	-.047	-4.650	-4.823	-.086	-1.758	-7.927	-1.049
	(1.526)	(1.330)	(1.281)	(1.393)	(1.414)	(1.423)	(.859)	(1.719)	(1.122)
T = 40	-10.047	-8.435	-.004	-6.503	-6.833	-.069	-2.491	-11.584	-1.386
	(1.484)	(1.382)	(1.119)	(1.389)	(1.366)	(1.187)	(.847)	(1.826)	(1.006)
T = 60	-12.250	-10.749	-.004	-7.937	-8.429	-.084	-3.030	-14.402	-1.633
	(1.468)	(1.439)	(1.093)	(1.397)	(1.377)	(1.135)	(.847)	(1.840)	(.959)
θ_{21}= 0.4									
T = 20	-5.425	-2.377	-.046	-3.905	-3.017	-.124	-.925	-6.864	-.277
	(1.340)	(1.042)	(1.132)	(1.334)	(1.282)	(1.402)	(.867)	(1.642)	(1.203)
T = 40	-7.507	-4.558	-.017	-5.462	-4.401	-.104	-1.336	-9.744	-.362
	(1.302)	(1.071)	(1.023)	(1.325)	(1.205)	(1.168)	(.856)	(1.665)	(1.054)
T = 60	-9.161	-6.012	-.009	-6.676	-5.489	-.126	-1.626	-11.966	-.408
	(1.287)	(1.109)	(1.009)	(1.329)	(1.197)	(1.118)	(.859)	(1.644)	(.999)
θ_{21}= 0.0									
T = 20	-3.927	-.145	.054	-2.944	-1.006	.096	.277	-5.198	.439
	(1.200)	(.919)	(.993)	(1.241)	(1.180)	(1.342)	(.897)	(1.503)	(1.277)
T = 40	-5.453	-.796	.001	-4.134	-1.684	.168	.334	-7.086	.547
	(1.173)	(.888)	(.926)	(1.229)	(1.086)	(1.134)	(.885)	(1.441)	(1.104)
T = 60	-6.674	-1.294	.147	-5.070	-2.198	.199	.405	-8.556	.663
	(1.161)	(.899)	(.927)	(1.229)	(1.065)	(1.088)	(.891)	(1.395)	(1.047)
θ_{21}= -0.8									
T = 20	-2.067	3.694	.635	-1.229	2.893	.530	4.495	.542	.013
	(1.066)	(1.201)	(.732)	(1.084)	(1.214)	(1.107)	(1.123)	(1.209)	(1.350)
T = 40	-2.898	5.509	.948	-1.758	4.041	.741	6.255	1.349	-.002
	(1.050)	(1.243)	(.712)	(1.067)	(1.161)	(.984)	(1.088)	(1.103)	(1.160)
T = 60	-3.574	7.130	1.236	-2.188	4.983	.913	7.630	1.975	.003
	(1.040)	(1.281)	(.737)	(1.061)	(1.143)	(.964)	(1.092)	(1.087)	(1.109)

Note:
(a) N =T.
(b) A lag length 5 of the Bartlett windows is used for the FMOLS estimator.
(c) 4 lags and 2 leads are used for the DOLS estimator.

Table 4: Means Biases and Standard Deviations of t-statistics for Different N and T

(N,T)	OLS	FMOLS(5)	FMOLS(2)	DOLS(4,2)	DOLS(2,1)
(1,20)	-1.169	-1.264	-1.334	-.304	.232
	(1.497)	(2.326)	(2.031)	(3.224)	(2.109)
(1,40)	-1.116	-1.169	-1.232	-.113	.258
	(1.380)	(1.805)	(1.738)	(2.086)	(1.689)
(1,60)	-1.090	-1.162	-1.195	-.071	.254
	(1.357)	(1.692)	(1.676)	(1.778)	(1.554)
(1,120)	-1.092	-1.239	-1.217	-.056	.234
	(1.333)	(1.165)	(1.652)	(1.531)	(1.448)
(20,20)	-3.905	-3.017	-3.156	-.124	.695
	(1.334)	(1.281)	(1.230)	(1.402)	(1.184)
(20,40)	-3.934	-3.202	-3.169	-.114	.634
	(1.307)	(1.206)	(1.200)	(1.186)	(1.099)
(20,60)	-3.861	-3.202	-3.111	-.053	.677
	(1.306)	(1.150)	(1.191)	(1.122)	(1.079)
(20,120)	-3.893	-3.247	-3.141	-.073	.642
	(1.312)	(1.149)	(1.209)	(1.078)	(1.061)
(40,20)	-5.439	-4.163	-4.342	-.088	1.008
	(1.347)	(1.269)	(1.226)	(1.358)	(1.169)
(40,40)	-5.462	-4.401	-4.344	-.104	.928
	(1.325)	(1.205)	(1.197)	(1.168)	(1.092)
(40,60)	-5.457	-4.506	-4.339	-.098	.913
	(1.328)	(1.199)	(1.192)	(1.121)	(1.081)
(40,120)	-5.469	-4.647	-4.356	-.106	.879
	(1.296)	(1.190)	(1.176)	(1.050)	(1.033)
(60,20)	-6.677	-5.097	-5.314	-.169	1.179
	(1.329)	(1.258)	(1.208)	(1.361)	(1.162)
(60,40)	-6.699	-5.384	-5.309	-.162	1.097
	(1.323)	(1.204)	(1.192)	(1.169)	(1.094)
(60,60)	-6.676	-5.489	-5.289	-.126	1.106
	(1.329)	(1.197)	(1.191)	(1.118)	(1.074)
(60,120)	-6.677	-5.656	-5.299	-.115	1.083
	(1.311)	(1.196)	(1.182)	(1.056)	(1.041)
(120,20)	-9.407	-7.153	-7.446	-.220	1.662
	(1.350)	(1.262)	(1.215)	(1.348)	(1.163)
(120,40)	-9.418	-7.753	-7.753	-.193	1.565
	(1.313)	(1.171)	(1.171)	(1.157)	(1.085)
(120,60)	-9.411	-7.717	-7.429	-.177	1.549
	(1.310)	(1.182)	(1.174)	(1.093)	(1.053)
(120,120)	-9.408	-7.932	-7.432	-.152	1.530
	(1.315)	(1.195)	(1.181)	(1.057)	(1.040)

Note:
(a) A lag length 5 and 2 of the Bartlett windows are used for the FMOLS(5) and FMOLS(2) estimators.
(b) 4 lags and 2 leads and 2 lags and 1 lead are used for the DOLS(4,2) and DOLS(2,1) estimators.
(c) $\sigma_{21} = -0.4$ and $\theta_{21} = 0.4$.

Table 5: Means Biases and Standard Deviations of OLS, FMOLS, and DOLS Estimators

	$\sigma_{21} = -0.8$			$\sigma_{21} = -0.4$			$\sigma_{21} = 0.8$		
	$\widehat{\beta}_{OLS}-\beta$	$\widehat{\beta}_{FM}-\beta$	$\widehat{\beta}_{D}-\beta$	$\widehat{\beta}_{OLS}-\beta$	$\widehat{\beta}_{FM}-\beta$	$\widehat{\beta}_{D}-\beta$	$\widehat{\beta}_{OLS}-\beta$	$\widehat{\beta}_{FM}-\beta$	$\widehat{\beta}_{D}-\beta$
θ_{21}= 0.8									
T = 20	-.110	-.101	.003	-.049	-.062	.000	-.009	-.036	-.003
	(.042)	(.038)	(.037)	(.029)	(.020)	(.030)	(.011)	(.012)	(.009)
T = 40	-.052	-.052	.001	-.024	-.031	.000	-.004	-.017	-.001
	(.015)	(.014)	(.012)	(.010)	(.011)	(.010)	(.004)	(.004)	(.003)
T = 60	-.034	-.035	.000	-.015	-.021	-.000	-.003	-.012	-.001
	(.008)	(.008)	(.007)	(.006)	(.006)	(.005)	(.002)	(.002)	(.002)
θ_{21}= 0.4									
T = 20	-.073	-.039	.001	-.045	-.038	-.000	-.006	-.037	-.001
	(.032)	(.024)	(.024)	(.028)	(.027)	(.028)	(.013)	(.014)	(.012)
T = 40	-.034	-.020	.000	-.021	-.019	-.000	-.002	-.017	-.001
	(.011)	(.008)	(.008)	(.010)	(.009)	(.009)	(.004)	(.004)	(.004)
T = 60	-.022	-.013	.000	-.013	-.012	-.000	-.002	-.012	-.000
	(.006)	(.004)	(.004)	(.005)	(.005)	(.005)	(.002)	(.002)	(.002)
θ_{21}= 0.0									
T = 20	-.046	-.006	.001	-.035	-.013	.001	-.001	-.034	.003
	(.025)	(.015)	(.015)	(.025)	(.022)	(.023)	(.016)	(.016)	(.015)
T = 40	-.021	-.003	.000	-.016	-.006	.001	-.001	-.016	.001
	(.009)	(.005)	(.005)	(.008)	(.007)	(.008)	(.006)	(.005)	(.005)
T = 60	-.014	-.002	.001	-.011	-.004	.001	-.000	-.010	.002
	(.005)	(.003)	(.003)	(.005)	(.004)	(.004)	(.003)	(.003)	(.003)
θ_{21}= −0.8									
T = 20	-.020	.017	.002	-.016	.017	.003	.035	.012	.000
	(.016)	(.009)	(.007)	(.017)	(.013)	(.012)	(.024)	(.024)	(.031)
T = 40	-.008	.008	.002	-.007	.008	.001	.016	.007	-.000
	(.005)	(.003)	(.002)	(.006)	(.004)	(.004)	(.009)	(.009)	(.009)
T = 60	-.006	.005	.001	-.005	.005	.001	.011	.005	.000
	(.003)	(.001)	(.001)	(.003)	(.002)	(.002)	(.005)	(.005)	(.005)

Note:
(a) N =T.
(b) A lag length 5 of the Bartlett windows is used for the FMOLS estimator.
(c) 4 lags and 2 leads are used for the DOLS estimator.
(d) The error terms are generated by an ARMA(1,1) process from equation (27).

Table 6: Means Biases and Standard Deviations of t-statistics

	$\sigma_{21} = -0.8$			$\sigma_{21} = -0.4$			$\sigma_{21} = 0.8$		
	OLS	FMOLS	DOLS	OLS	FMOLS	DOLS	OLS	FMOLS	DOLS
θ_{21}= 0.8									
T = 20	-5.316	-3.569	.119	-3.411	-2.912	.006	-1.158	-4.589	-.347
	(1.929)	(1.323)	(1.290)	(1.924)	(1.390)	(1.417)	(1.426)	(1.420)	(1.139)
T = 40	-7.013	-4.601	.090	-4.583	-3.580	.009	-1.723	-6.144	-.505
	(1.903)	(1.219)	(1.119)	(1.949)	(1.216)	(1.166)	(1.445)	(1.343)	(1.011)
T = 60	-8.437	-5.22	.068	-5.523	-4.206	-.006	-2.097	-7.428	-.603
	(1.899)	(1.195)	(1.077)	(1.969)	(1.178)	(1.111)	(1.435)	(1.294)	(.978)
θ_{21}= 0.4									
T = 20	-4.152	-1.857	.056	-3.064	-1.877	-.025	-.705	-3.858	-.068
	(1.762)	(1.106)	(1.132)	(1.867)	(1.314)	(1.388)	(1.454)	(1.373)	(1.208)
T = 40	-5.424	-2.576	.045	-4.069	-2.346	-.011	-1.099	-5.034	-.134
	(1.733)	(1.044)	(1.027)	(1.880)	(1.149)	(1.152)	(1.479)	(1.268)	(1.053)
T = 60	-6.521	-3.179	.034	-4.899	-2.779	-.027	-1.343	-6.016	-.144
	(1.721)	(1.036)	(1.004)	(1.898)	(1.114)	(1.096)	(1.473)	(1.211)	(1.014)
θ_{21}= 0.0									
T = 20	-3.184	-.353	.034	-2.538	-.732	.038	-.047	-2.825	.230
	(1.644)	(.952)	(.956)	(1.769)	(1.226)	(1.313)	(1.498)	(1.327)	(1.276)
T = 40	-4.120	-.624	.047	-3.327	-.967	.075	-.194	-3.557	.212
	(1.616)	(.897)	(.909)	(1.771)	(1.085)	(1.116)	(1.528)	(1.194)	(1.095)
T = 60	-4.952	-.827	.058	-4.131	-1.141	.206	-.064	-4.005	.693
	(1.599)	(.904)	(.913)	(1.746)	(1.021)	(1.118)	(1.498)	(1.096)	(1.094)
θ_{21}= -0.8									
T = 20	-1.956	1.733	.214	-1.496	1.429	.221	2.315	.564	.002
	(1.529)	(.933)	(.663)	(1.589)	(1.015)	(1.052)	(1.577)	(1.195)	(1.551)
T = 40	-2.471	2.511	.317	-1.888	1.917	.294	3.089	.876	-.005
	(1.507)	(.871)	(.664)	(1.578)	(1.010)	(.956)	(1.644)	(1.088)	(1.239)
T = 60	-2.966	3.270	.428	-2.267	2.237	.363	3.736	1.132	.003
	(1.484)	(.897)	(.694)	(1.571)	(.999)	(.941)	(1.676)	(1.062)	(1.155)

Note:
(a) N =T.
(b) A lag length 5 of the Bartlett windows is used for the FMOLS estimator.
(c) 4 lags and 2 leads are used for the DOLS estimator.
(d) The error terms are generated by an ARMA(1,1) process from equation (27).

Table 7: Means Biases and Standard Deviations of OLS, FMOLS, and DOLS Estimators

	$\sigma = 0.25$			$\sigma = 0.5$			$\sigma = 1$		
	$\widehat{\beta}_{OLS}-\beta$	$\widehat{\beta}_{FM}-\beta$	$\widehat{\beta}_{D}-\beta$	$\widehat{\beta}_{OLS}-\beta$	$\widehat{\beta}_{FM}-\beta$	$\widehat{\beta}_{D}-\beta$	$\widehat{\beta}_{OLS}-\beta$	$\widehat{\beta}_{FM}-\beta$	$\widehat{\beta}_{D}-\beta$
θ_{21}= 0.8									
T = 20	-.005	-.011	-.000	-.002	-.007	-.000	-.001	-.004	-.000
	(.009)	(.009)	(.002)	(.006)	(.006)	(.003)	(.003)	(.003)	(.002)
T = 40	-.001	-.003	-.000	-.001	.002	-.028	-.000	-.001	-.000
	(.002)	(.002)	(.000)	(.001)	(.001)	(.001)	(.001)	(.001)	(.000)
T = 60	-.001	-.001	-.000	-.000	-.001	-.000	-.000	-.001	-.000
	(.001)	(.001)	(.000)	(.001)	(.001)	(.000)	(.000)	(.000)	(.000)
θ_{21}= 0.4									
T = 20	-.002	-.008	-.001	-.002	-.008	-.000	-.001	-.005	-.000
	(.009)	(.009)	(.005)	(.009)	(.009)	(.005)	(.004)	(.004)	(.002)
T = 40	-.002	-.005	-.000	-.000	-.002	-.000	-.000	-.001	-.000
	(.004)	(.004)	(.001)	(.002)	(.002)	(.001)	(.001)	(.001)	(.001)
T = 60	-.001	-.002	-.000	-.000	-.001	-.000	-.000	-.001	-.000
	(.002)	(.002)	(.001)	(.001)	(.001)	(.001)	(.000)	(.000)	(.000)
θ_{21}= 0.0									
T = 20	.012	-.010	.001	.005	-.007	.001	.001	-.005	.000
	(.058)	(.057)	(.054)	(.017)	(.016)	(.014)	(.005)	(.005)	(.003)
T = 40	.003	-.002	.000	.001	-.002	.000	.000	-.001	.000
	(.014)	(.014)	(.013)	(.004)	(.004)	(.003)	(.001)	(.001)	(.001)
T = 60	.001	-.001	.000	.001	-.001	-.000	.000	-.001	.000
	(.007)	(.007)	(.006)	(.002)	(.002)	(.002)	(.001)	(.001)	(.000)
θ_{21}= −0.8									
T = 20	.011	.022	-.000	.034	.049	.001	.039	.008	.000
	(.013)	(.012)	(.002)	(.020)	(.019)	(.013)	(.016)	(.014)	(.013)
T = 40	.003	.006	.000	.009	.014	.000	.012	.003	-.000
	(.003)	(.003)	(.001)	(.005)	(.005)	(.003)	(.004)	(.004)	(.003)
T = 60	.001	.003	.000	.004	.007	.000	.005	.002	-.000
	(.001)	(.001)	(.000)	(.002)	(.002)	(.001)	(.002)	(.002)	(.001)

Note:
(a) N =T.
(b) A lag length 5 of the Bartlett windows is used for the FMOLS estimator.
(c) 4 lags and 2 leads are used for the DOLS estimator.
(d) The error terms are non-normal.

Table 8: Means Biases and Standard Deviations of t-statistics

	$\sigma = 0.25$			$\sigma = 0.5$			$\sigma = 1$		
	OLS	FMOLS	DOLS	OLS	FMOLS	DOLS	OLS	FMOLS	DOLS
θ_{21}= 0.8									
T = 20	-.699	-1.248	-.006	-.472	-1.055	-.039	-.406	-1.265	-.118
	(1.311)	(.940)	(.209)	(1.245)	(.931)	(.421)	(1.040)	(.925)	(.520)
T = 40	-.717	-.892	-.002	-.484	-.752	-.003	-.424	-.918	-.096
	(1.253)	(.599)	(.139)	(1.191)	(.597)	(.276)	(.981)	(.588)	(.336)
T = 60	-.741	-.738	-.002	-.506	-.623	-.028	-.445	-.764	-.088
	(1.267)	(.488)	(.113)	(1.199)	(.483)	(.227)	(.979)	(.472)	(.276)
θ_{21}= 0.4									
T = 20	-.259	-.884	-.071	-.259	-.884	-.071	-.199	-1.152	-.019
	(1.243)	(.932)	(.561)	(1.243)	(.932)	(.561)	(1.040)	(.927)	(.567)
T = 40	-.587	-.787	-.007	-.268	-.626	-.054	-.213	-.831	-.016
	(1.250)	(.599)	(.230)	(1.189)	(.599)	(.363)	(.981)	(.589)	(.368)
T = 60	-.611	-.651	-.008	-.289	-.519	-.052	-.232	-.692	-.020
	(1.264)	(.488)	(.188)	(1.197)	(.485)	(.299)	(.978)	(.474)	(.304)
θ_{21}= 0.0									
T = 20	.275	-.164	.014	.340	-.398	.031	.145	-.961	.066
	(1.271)	(.941)	(.896)	(1.236)	(.941)	(.784)	(1.041)	(.931)	(.619)
T = 40	.282	-.106	.013	.347	-.268	.025	.141	-.685	.053
	(1.231)	(.616)	(.579)	(1.186)	(.611)	(.509)	(.982)	(.594)	(.407)
T = 60	.264	-.093	.002	.332	-.226	.013	.125	-.570	.039
	(1.248)	(.505)	(.477)	(1.193)	(.497)	(.421)	(.978)	(.478)	(.337)
θ_{21}= -0.8									
T = 20	1.104	1.714	-.000	2.286	2.528	.035	2.749	.539	.026
	(1.326)	(.951)	(.189)	(1.278)	(.976)	(.650)	(1.067)	(.984)	(.899)
T = 40	1.134	1.249	.001	2.368	1.947	.035	2.946	.598	.008
	(1.262)	(.605)	(.126)	(1.208)	(.633)	(.446)	(.992)	(.672)	(.624)
T = 60	1.163	1.036	.001	2.416	1.637	.033	3.011	.538	-.002
	(1.274)	(.492)	(.102)	(1.214)	(.513)	(.363)	(.981)	(.554)	(.525)

Note:
(a) N =T.
(b) A lag length 5 of the Bartlett windows is used for the FMOLS estimator.
(c) 4 lags and 2 leads are used for the DOLS estimator.
(d) The error terms are non-normal.

Table 9: Means Biases and Standard Deviations of OLS, FMOLS, and DOLS Estimators for Different N and T in a Heterogeneous Panel

(N,T)	$\widehat{\beta}^*_{OLS} - \beta$	$\widehat{\beta}^*_{FM(5)} - \beta$	$\widehat{\beta}^*_{FM(2)} - \beta$	$\widehat{\beta}^*_{D(4,2)} - \beta$	$\widehat{\beta}^*_{D(2.1)} - \beta$
(1,20)	-.102	.076	-.008	-.011	.004
	(.163)	(.319)	(.212)	(.405)	(.264)
(1,40)	-.052	.006	-.018	.001	.006
	(.079)	(.116)	(.084)	(.121)	(.099)
(1,60)	-.035	-.004	-.014	.001	.005
	(.052)	(.066)	(.050)	(.071)	(.061)
(1,120)	-.018	-.008	-.009	.000	.002
	(.026)	(.027)	(.023)	(.030)	(.029)
(20,20)	-.025	-.069	-.073	-.000	.006
	(.032)	(.054)	(.034)	(.054)	(.040)
(20,40)	-.016	-.041	-.035	-.001	.004
	(.014)	(.019)	(.014)	(.020)	(.017)
(20,60)	-.012	-.028	-.023	-.000	.003
	(.009)	(.011)	(.009)	(.012)	(.011)
(20,120)	-.006	-.014	-.011	-.000	.002
	(.004)	(.005)	(.004)	(.005)	(.005)
(40,20)	-.023	-.089	-.083	.000	.007
	(.024)	(.038)	(.024)	(.038)	(.028)
(40,40)	-.015	-.048	-.039	-.001	.004
	(.009)	(.013)	(.009)	(.014)	(.012)
(40,60)	-.013	-.032	-.026	.000	.003
	(.006)	(.008)	(.006)	(.009)	(.008)
(40,120)	-.014	-.014	-.012	-.000	.002
	(.004)	(.004)	(.003)	(.003)	(.004)
(60,20)	-.023	-.073	-.074	.001	.006
	(.019)	(.031)	(.019)	(.031)	(.023)
(60,40)	-.015	-.042	-.036	-.001	.004
	(.008)	(.011)	(.008)	(.011)	(.009)
(60,60)	-.011	-.029	-.023	-.000	.003
	(.005)	(.006)	(.005)	(.007)	(.006)
(60,120)	-.006	-.014	-.011	-.000	.002
	(.002)	(.003)	(.002)	(.003)	(.003)
(120,20)	-.022	-.075	-.072	.001	.016
	(.014)	(.003)	(.022)	(.022)	(.011)
(120,40)	-.015	-.042	-.036	-.001	.004
	(.006)	(.008)	(.006)	(.008)	(.007)
(120,60)	-.011	-.029	-.024	-.000	.003
	(.004)	(.004)	(.004)	(.005)	(.004)
(120,120)	-.006	-.014	-.011	-.000	.002
	(.002)	(.002)	(.002)	(.002)	(.002)

Note:
(a) A lag length 5 and 2 of the Bartlett windows are used for the FMOLS(5) and FMOLS(2) estimators.
(b) 4 lags and 2 leads and 2 lags and 1 lead are used for the DOLS(4,2) and DOLS(2,1) estimators.
(c) $\sigma_{21} \sim U[-0.8, 0.8]$ and $\theta_{21} \sim U[-0.8, 0.8]$.

Table 10: Means Biases and Standard Deviations of t-statistics for Different N and T in a Heterogeneous Panel

(N,T)	OLS	FMOLS(5)	FMOLS(2)	DOLS(4,2)	DOLS(2,1)
(1,20)	-.893	.588	-.058	-.093	.029
	(1.390)	(2.473)	(1.643)	(3.303)	(2.156)
(1,40)	-.861	.101	.280	.009	.106
	(1.265)	(1.849)	(1.331)	(1.980)	(1.618)
(1,60)	-.844	-.095	-.347	.016	.119
	(1.233)	(1.579)	(1.207)	(1.729)	(1.489)
(1,120)	-.845	-.372	-.459	.016	.101
	(1.212)	(1.336)	(1.139)	(1.510)	(1.405)
(20,20)	-1.221	-2.411	-2.530	.010	.219
	(1.578)	(1.902)	(1.192)	(1.983)	(1.468)
(20,40)	-1.629	-2.899	-2.518	-.059	.271
	(1.344)	(1.345)	(.999)	(1.485)	(1.259)
(20,60)	-1.774	-3.031	-2.508	.004	.347
	(1.282)	(1.195)	(.952)	(1.329)	(1.184)
(20,120)	-1.957	-3.095	-2.466	.046	.393
	(1.239)	(1.047)	(.907)	(1.197)	(1.121)
(40,20)	-1.612	-4.381	-4.079	.039	.365
	(1.640)	(1.882)	(1.191)	(1.987)	(1.466)
(40,40)	-2.194	-4.807	-3.969	-.068	.432
	(1.392)	(1.341)	(1.004)	(1.472)	(1.233)
(40,60)	-2.417	-4.905	-3.932	.007	.515
	(1.306)	(1.199)	(.960)	(1.319)	(1.169)
(40,120)	-2.832	-4.886	-3.839	.099	.608
	(1.234)	(1.059)	(.911)	(1.181)	(1.099)
(60,20)	-1.946	-4.408	-4.474	.041	.408
	(1.697)	(1.884)	(1.182)	(1.932)	(1.449)
(60,40)	-2.715	-5.171	-4.407	-.110	.472
	(1.389)	(1.320)	(.976)	(1.452)	(1.221)
(60,60)	-3.045	-5.361	-4.380	-.027	.572
	(1.328)	(1.170)	(.933)	(1.307)	(1.165)
(60,120)	-3.346	-5.420	-4.281	.105	.697
	(1.250)	(1.033)	(.889)	(1.181)	(1.099)
(120,20)	-2.675	-6.382	-6.383	.073	.580
	(1.720)	(1.878)	(1.169)	(1.939)	(1.439)
(120,40)	-3.802	-7.399	-6.272	-.145	.683
	(1.408)	(1.314)	(.967)	(1.444)	(1.215)
(120,60)	-4.269	-7.633	-6.209	-.047	.803
	(1.336)	(1.162)	(.931)	(1.307)	(1.165)
(120,120)	-4.715	-7.723	-6.084	.136	.977
	(1.250)	(1.045)	(.897)	(1.178)	(1.098)

Note:
(a) A lag length 5 and 2 of the Bartlett windows are used for the FMOLS(5) and FMOLS(2) estimators.
(b) 4 lags and 2 leads and 2 lags and 1 lead are used for the DOLS(4,2) and DOLS(2,1) estimators.
(c) $\sigma_{21} \sim U[-0.8, 0.8]$ and $\theta_{21} \sim U[-0.8, 0.8]$.

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
2	Kao and Chiang	On the Estimation and Inference of a Cointegrated Regression in Panel Data