

2000

Nonstationary Panels, Cointegration in Panels and Dynamic Panels: A Survey

Badi H. Baltagi

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

Chihwa Kao

Syracuse University, Maxwell School, 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

Baltagi, Badi H. and Kao, Chihwa, "Nonstationary Panels, Cointegration in Panels and Dynamic Panels: A Survey" (2000). *Center for Policy Research*. Paper 136.

<http://surface.syr.edu/cpr/136>

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. 16**

**NONSTATIONARY PANELS, COINTEGRATION IN PANELS
AND DYNAMIC PANELS: A SURVEY**

Badi H. Baltagi and Chihwa Kao

**Center for Policy Research
Maxwell School of Citizenship and Public Affairs
Syracuse University
426 Eggers Hall
Syracuse, New York 13244-1020
(315) 443-3114 | Fax (315) 443-1081
e-mail: ctrpol@syr.edu**

March 2000

\$5.00

Up-to-date information about CPR's research projects and other activities is available from our World Wide Web site at www-cpr.maxwell.syr.edu. All recent working papers and Policy Briefs can be read and/or printed from there as well.

CENTER FOR POLICY RESEARCH – Spring 2000

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

Dan Black.....	Economics	Duke Kao	Economics
Stacy Dickert-Conlin	Economics	Thomas Kniesner	Economics
William Duncombe	Public Administration	Jeff Kubik.....	Economics
Thomas Dunn	Economics	Jerry Miner	Economics
Gary Engelhardt.....	Economics	John Moran	Economics
Deborah Freund	Public Administration	Jan Ondrich.....	Economics
Vernon Greene	Public Administration	John Palmer.....	Public Administration
Leah Gutierrez	Public Administration	Grant Reeher	Political Science
Madonna Harrington Meyer.....	Sociology	Stuart Rosenthal.....	Economics
Christine Himes	Sociology	Jodi Sandfort.....	Public Administration
Jacqueline Johnson.....	Sociology	Michael Wasylenko	Economics
Bernard Jump	Public Administration	Assata Zerai.....	Sociology

GRADUATE ASSOCIATES

Matthew Andrews.....	Public Administration	Peter Howe	Economics
Yvonne Arsenault	Public Administration	Alyssa Hundrup	Public Administration
Reagan Baughman.....	Economics	Kwangho Jung.....	Public Administration
Robert Bifulco.....	Public Administration	Young Sun Kwon.....	Economics
Christine Caffrey	Sociology	James Laditka.....	Public Administration
Meghan Collins.....	Public Administration	Xin Li.....	Economics
Julie Dombrowski	Sociology	Donald Marples.....	Economics
Karla English.....	Public Administration	Neddy Matshalaga	Sociology
Amy Ferraro.....	Public Administration	Adriana Sandu	Public Administration
Seth Giertz	Economics	Shalini Sharma.....	Economics
Andrzej Grodner	Economics	Mehmet Serkan Tosun.....	Economics
Tess Heintze.....	Public Administration	James Williamson.....	Economics
Pam Herd.....	Sociology		

STAFF

JoAnna Berger.....	Receptionist	Denise Paul.....	Editorial Assistant, <i>NTJ</i>
Martha Bonney	Publications and Events Coordinator	Annie Pennella.....	Editorial Assistant, <i>Gerontologist</i>
Karen Cimilluca.....	Librarian/Office Coordinator	Mary Santy	Secretary
Kati Foley	Administrative Assistant, LIS	Debbie Tafel.....	Secretary to the Director
Esther Gray.....	Administrative Secretary	Ann Wicks	Administrative Secretary
		Lobrenzo Wingo	Computer Consultant

Nonstationary Panels, Cointegration in Panels and Dynamic Panels: A Survey

Badi H. Baltagi
Texas A&M University
Department of Economics
College Station, TX 77843-4228
(409) 845-7380
badi@econ.tamu.edu

and

Chihwa Kao
Syracuse University
Center for Policy Research
Syracuse, NY 13244-1020
(315) 443-3233
cdkao@maxwell.syr.edu

January 28, 2000

Abstract

This paper provides an overview of topics in nonstationary panels: panel unit root tests, panel cointegration tests, and estimation of panel cointegration models. In addition it surveys recent developments in dynamic panel data models.

1 Introduction

Two important areas in the econometrics of panel data that have received a lot of attention recently are dynamic panel data and nonstationary panel time series models. This special issue focuses on these two topics. With the growing use of cross-country data over time to study purchasing power parity, growth convergence and international R&D spillovers, the focus of panel data econometrics has shifted towards studying the asymptotics of macro panels with large N (number of countries) and large T (length of the time series) rather than the usual asymptotics of micro panels with large N and small T . In fact, the limiting distributions of double indexed integrated processes had to be developed, see Phillips and Moon (1999a). The fact that T is allowed to increase to ∞ in macro panel data, generated two strands of ideas. The first rejected the homogeneity of the regression parameters implicit in the use of a pooled regression model in favor of heterogeneous regressions, i.e., one for each country, see Pesaran and Smith (1995), Im, Pesaran and Shin (1997), Lee, Pesaran and Smith (1997), Pesaran, Shin and Smith (1999) and Pesaran and Zhao (1999) to mention a few. This literature critically relies on T being large to estimate each country's regression separately. Another strand of literature applied time series procedures to panels, worrying about nonstationarity, spurious regressions and cointegration. Adding the cross-section dimension to the time-series dimension offers an advantage in the testing for non-stationarity and cointegration. The hope of the econometrics of non-stationary panel data is to combine the best of both worlds: the method of dealing with non-stationary data from the time series and the increased data and power from the cross-section. The addition of the cross-section dimension, under certain assumptions, can act as repeated draws from the same distribution. Thus as the time and cross-section dimension increase panel test statistics and estimators can be derived which converge in distribution to normally distributed random variables.

However, the use of such panel data methods are not without their critics, see Maddala, Wu and Liu (1999) who argue that panel data unit root tests do not rescue purchasing power parity. In fact, the results on PPP with panels are mixed depending on the group of countries studied, the period of study and the type of unit root test used. More damaging is the argument by Maddala, Wu and Liu (1999) that for PPP, panel data tests are the wrong answer to the low power of unit root tests in single time series. After all, the null hypothesis of a single unit root is different from the null hypothesis of a panel unit root for the PPP hypothesis. Using the same line of criticism, Maddala (1999) argues that panel unit root tests did not help settle the question of growth convergence among countries. However, it was useful in spurring much needed research into dynamic panel data models. Also, Quah (1996) who argued that the basic issues of whether poor countries catch up with the rich can never be answered by the use of traditional panels. Instead, Quah suggests formulating and estimating models of income dynamics.

One can find numerous applications of time series methods applied to panels in recent years. Examples from the purchasing power parity (PPP) literature include Bernard and Jones (1996), Jorion and Sweeney (1996), MacDonald (1996), Oh (1996), Wu (1996), Coakley and Fuertes (1997), Culver and Papell (1997), Papell (1997), O'Connell (1998), Choi (1999a), Andersson and Lyhagen (1999), and Canzoneri, Cumby and Diba (1999) to mention a few. On health care expenditures, see McCoskey and Selden (1998), and Gerdtham and Löthgren (1998). On growth and convergence, see Islam (1995), Evans and Karas (1996), Sala-i-Martin (1996), Lee, Pesaran and Smith (1997), and McCoskey and Kao (1999a). On international R&D spillovers, see Funk (1998) and Kao, Chiang and Chen (1999). On exchange rate models, see Groen and Kleibergen (1999), and Groen (1999).

The first part of this paper surveys some of the developments in nonstationary panel models that have occurred since the middle of 1990s. Two other recent surveys on this subject include Phillips and Moon (1999b) on multi-indexed processes and Banerjee (1999) on panel unit roots and cointegration tests. We will pay attention to the following three topics: (1) panel unit root tests, (2) panel cointegration tests, and (3) estimation and inference in the panel cointegration models. The discussion of each topic will be illustrated by examples taken from the aforementioned list of references. Section 2 reviews panel unit root tests, while Section 3 discusses the panel spurious models. Section 4 considers the panel cointegration tests, while Section 5 discusses panel cointegration models. Section 6 reviews some recent development in dynamic panels and Section 7 gives our conclusion.

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 $\Omega = (\Omega^{1/2}) (\Omega^{1/2})'$. We use \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 covariance matrix Ω .

2 Panel Unit Root Tests

Testing for unit roots in time series studies is now common practice among applied researchers and has become an integral part of econometric courses. However, testing for unit roots in panels is recent, see Levin and Lin (1992), Im, Pesaran and Shin (1997), Harris and Tzavalis (1999), Maddala and Wu (1999), Choi (1999a), and Hadri (1999). Exceptions are Bharagava et al. (1982), Boumahdi and Thomas (1991), Breitung and Meyer (1994), and Quah (1994). Bharagava et al. (1982) proposed a test for random walk residuals in a dynamic model with fixed effects. They suggested a modified Durbin-Watson (DW) statistic

based on fixed effects residuals and two other test statistics based on differenced OLS residuals. In typical micro panels with $N \rightarrow \infty$, they recommended their modified DW statistic. Boumahdi and Thomas (1991) proposed a generalization of the Dickey-Fuller test for unit roots in panel data to assess the efficiency of the French capital market using 140 french stock prices over the period January 1973 to February 1986. Breitung and Meyer (1994) applied various modified Dickey-Fuller test statistics to test for unit roots in a panel of contracted wages negotiated at the firm and industry level for Western Germany over the period 1972-1987. Quah (1994) suggested a test for unit root in a panel data model without fixed effects where both N and T go to infinity at the same rate such that N/T is constant. Levin and Lin (1992) generalized this model to allow for fixed effects, individual deterministic trends and heterogeneous serially correlated errors. They assumed that both N and T tend to infinity. However, T increases at a faster rate than N with $N/T \rightarrow 0$. Even though this literature grew from time series and panel data, the way in which N , the number of cross-section units, and T , the length of the time series, tend to infinity is crucial for determining asymptotic properties of estimators and tests proposed for nonstationary panels, see Phillips and Moon (1999a).

2.1 Levin and Lin (1992) Tests

Consider the model

$$y_{it} = \rho_i y_{it-1} + z'_{it} \gamma + u_{it}, \quad i = 1, \dots, N; t = 1, \dots, T, \quad (1)$$

where z_{it} is the deterministic component and u_{it} is a stationary process. z_{it} could be zero, one, the fixed effects, μ_i , or fixed effect as well as a time trend. The Levin and Lin (LL) tests assume that u_{it} are *iid* $(0, \sigma_u^2)$ and $\rho_i = \rho$ for all i . LL are interested in testing the null hypothesis

$$H_0 : \rho = 1 \quad (2)$$

against the alternative hypothesis

$$H_a : \rho < 1.$$

Let $\hat{\rho}$ be the OLS estimator of ρ in (1) and define

$$\begin{aligned} z_t &= (z_{1t}, \dots, z_{Nt})', \\ h(t, s) &= z'_t \left(\sum_{t=1}^T z_t z'_t \right) z_s, \\ \tilde{u}_{it} &= u_{it} - \sum_{s=1}^T h(t, s) u_{is}, \end{aligned}$$

and

$$\tilde{y}_{it} = y_{it} - \sum_{s=1}^T h(t, s) y_{is}. \quad (3)$$

Then we have

$$\sqrt{NT}(\hat{\rho} - 1) = \frac{\frac{1}{\sqrt{N}} \sum_{i=1}^N \frac{1}{T} \sum_{t=1}^T \tilde{y}_{i,t-1} \tilde{u}_{it}}{\frac{1}{N} \sum_{i=1}^N \frac{1}{T^2} \sum_{t=1}^T \tilde{y}_{i,t-1}^2}$$

and the corresponding t -statistic, under the null hypothesis is given by

$$t_\rho = \frac{(\hat{\rho} - 1) \sqrt{\sum_{i=1}^N \sum_{t=1}^T \tilde{y}_{i,t-1}^2}}{s_e},$$

where

$$s_e^2 = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \tilde{u}_{it}^2.$$

Assume that there exists a scaling matrix D_T and piecewise continuous function $Z(r)$ such that

$$D_T^{-1} z_{[Tr]} \rightarrow Z(r)$$

uniformly for $r \in [0, 1]$. For a fixed N , we have

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N \frac{1}{T} \sum_{t=1}^T \tilde{y}_{i,t-1} \tilde{u}_{it} \Rightarrow \frac{1}{\sqrt{N}} \sum_{i=1}^N \int W_{iZ} dW_{iZ}$$

and

$$\frac{1}{N} \sum_{i=1}^N \frac{1}{T^2} \sum_{t=1}^T \tilde{y}_{i,t-1}^2 \Rightarrow \frac{1}{N} \sum_{i=1}^N \int W_{iZ}^2,$$

as $T \rightarrow \infty$, where

$$W_{iZ}(r) = W_i(r) - \left[\int W_i Z' \right] \left[\int Z Z' \right]^{-1} Z(r), \quad (4)$$

is the L_2 projection residual of $W_i(r)$ on $Z_i(r)$, and $W_i(r)$ is a standard Brownian motion. Next we assume that $\int W_{iZ} dW_{iZ}$ and $\int W_{iZ}^2$ are independent across i and have finite second moments. Then it follows that

$$\begin{aligned} & \frac{1}{N} \sum_{i=1}^N \int W_{iZ}^2 \xrightarrow{p} E \left[\int W_{iZ}^2 \right] \\ & \frac{1}{\sqrt{N}} \sum_{i=1}^N \left(\int W_{iZ} dW_{iZ} - E \left[\int W_{iZ} dW_{iZ} \right] \right) \Rightarrow N \left(0, \text{Var} \left(\int W_{iZ} dW_{iZ} \right) \right) \end{aligned}$$

as $N \rightarrow \infty$ by a law of large numbers and the Lindeberg-Levy central limit theorem. The following moments are taken from Levin and Lin (1992):

z_{it}	$E [\int W_{iZ} dW_{iZ}]$	$Var [\int W_{iZ} dW_{iZ}]$	$E [\int W_{iZ}^2]$	$Var [\int W_{iZ}^2]$	
0	0	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{3}$	
1	0	$\frac{1}{3}$	$\frac{1}{2}$?	(5)
μ_i	$-\frac{1}{2}$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{45}$	
$(\mu_i, t)'$	$-\frac{1}{2}$	$\frac{1}{60}$	$\frac{1}{15}$	$\frac{11}{6300}$	

Using (5), Levin and Lin (1992) obtain the following limiting distributions of $\sqrt{NT}(\hat{\rho} - 1)$ and t_ρ :

z_{it}	$\hat{\rho}$	t_ρ	
0	$\sqrt{NT}(\hat{\rho} - 1) \Rightarrow N(0, 2)$	$t_\rho \Rightarrow N(0, 1)$	
1	$\sqrt{NT}(\hat{\rho} - 1) \Rightarrow N(0, 2)$	$t_\rho \Rightarrow N(0, 1)$	(6)
μ_i	$\sqrt{NT}(\hat{\rho} - 1) + 3\sqrt{N} \Rightarrow N(0, \frac{51}{5})$	$\sqrt{1.25}t_\rho + \sqrt{1.875N} \Rightarrow N(0, 1)$	
$(\mu_i, t)'$	$\sqrt{N}(T(\hat{\rho} - 1) + 7.5) \Rightarrow N(0, \frac{2895}{112})$	$\sqrt{\frac{448}{277}}(t_\rho + \sqrt{3.75N}) \Rightarrow N(0, 1)$	

Sequential limit theory, i.e., $T \rightarrow \infty$ followed by $N \rightarrow \infty$, is used to derive the limiting distributions in (6). In case u_{it} is stationary, the asymptotic distributions of $\hat{\rho}$ and t_ρ need to be modified due to the presence of serial correlation.

Harris and Tzavalis (1999) also derived unit root tests for (1) with $z_{it} = \{0\}, \{\mu_i\}$, or $\{(\mu_i, t)'\}$ when the time dimension of the panel, T , is fixed. This is the typical case for micro panel studies. The main results are:

z_{it}	$\hat{\rho}$	
0	$\sqrt{N}(\hat{\rho} - 1) \Rightarrow N\left(0, \frac{2}{T(T-1)}\right)$	
μ_i	$\sqrt{N}\left(\hat{\rho} - 1 + \frac{3}{T+1}\right) \Rightarrow N\left(0, \frac{3(17T^2 - 20T + 17)}{5(T-1)(T+1)^3}\right)$	
$(\mu_i, t)'$	$\sqrt{N}\left(\hat{\rho} - 1 + \frac{15}{2(T+2)}\right) \Rightarrow N\left(0, \frac{15(193T^2 - 728T + 1147)}{112(T+2)^3(T-2)}\right)$	

Harris and Tzavalis (1999) also showed that the assumption that T tends to infinity at a faster rate than N as in LL rather than T fixed as in the case in micro panels, yields tests which are substantially undersized and have low power especially when T is small.

Recently, Frankel and Rose (1996), Oh (1996), and Lothian (1996) tested the PPP hypothesis using panel data. All of these articles use LL tests and some of them report evidence supporting the PPP hypothesis. O'Connell (1998), however, showed that the LL tests suffered from significant size distortion in the presence of correlation among contemporaneous cross-sectional error terms. O'Connell highlighted the importance of controlling for cross-sectional dependence when testing for a unit root in panels of real exchange rates. He showed that, controlling for cross-sectional dependence, no evidence against the null of a random walk can be found in panels of up to 64 real exchange rates.

Virtually all the existing nonstationary panel literature assume cross-sectional independence. It is true that the assumption of independence across i is rather strong, but it is needed in order to satisfy the requirement of the Lindeberg-Levy central limit theorem. Moreover, as pointed out by Quah (1994), modeling cross-sectional dependence is involved because individual observations in a cross-section have no natural ordering. Driscoll and Kraay (1998) presented a simple extension of common nonparametric covariance matrix estimation techniques which yields standard errors that are robust to very general forms of spatial and temporal dependence as the time dimension becomes large. In a recent paper, Conley (1999) presents a spatial model of dependence among agents using a metric of economic distance that provides cross-sectional data with a structure similar to time-series data. Conley proposes a generalized method of moments (GMM) using such dependent data and a class of nonparametric covariance matrix estimators that allow for a general form of dependence characterized by economic distance.

2.2 Im, Pesaran and Shin (1997) Tests

The LL test is restrictive in the sense that it requires ρ to be homogeneous across i . As Maddala (1999) points out, the null may be fine for testing convergence in growth among countries, but the alternative restricts every country to converge at the same rate. Im, Pesaran and Shin (1997) (IPS) allow for a heterogeneous coefficient of y_{it-1} and propose an alternative testing procedure based on averaging individual unit root test statistics. IPS suggested an average of the augmented Dickey-Fuller (ADF) tests when u_{it} is serially correlated with different serial correlation properties across cross-sectional units, i.e., $u_{it} = \sum_{j=1}^{p_i} \varphi_{ij} u_{it-j} + \varepsilon_{it}$. Substituting this u_{it} in (1) we get:

$$y_{it} = \rho_i y_{it-1} + \sum_{j=1}^{p_i} \varphi_{ij} \Delta y_{it-j} + z'_{it} \gamma + \varepsilon_{it}. \quad (7)$$

The null hypothesis is

$$H_0 : \rho_i = 1$$

for all i and the alternative hypothesis is

$$H_a : \rho_i < 1$$

for at least one i . The IPS t -bar statistic is defined as the average of the individual ADF statistic as

$$\bar{t} = \frac{1}{N} \sum_{i=1}^N t_{\rho_i}, \quad (8)$$

where t_{ρ_i} is the individual t -statistic of testing $H_0 : \rho_i = 1$ in (7). It is known that for a fixed N

$$t_{\rho_i} \Rightarrow \frac{\int_0^1 W_{iZ} dW_{iZ}}{\left[\int_0^1 W_{iZ}^2 \right]^{1/2}} = t_{iT} \quad (9)$$

as $T \rightarrow \infty$. IPS assume that t_{iT} are *iid* and have finite mean and variance. Then

$$\frac{\sqrt{N} \left(\frac{1}{N} \sum_{i=1}^N t_{iT} - E[t_{iT} | \rho_i = 1] \right)}{\sqrt{Var[t_{iT} | \rho_i = 1]}} \Rightarrow N(0, 1) \quad (10)$$

as $N \rightarrow \infty$ by the Lindeberg-Levy central limit theorem. Hence

$$t_{IPS} = \frac{\sqrt{N} (\bar{t} - E[t_{iT} | \rho_i = 1])}{\sqrt{Var[t_{iT} | \rho_i = 1]}} \Rightarrow N(0, 1) \quad (11)$$

as $T \rightarrow \infty$ followed by $N \rightarrow \infty$ sequentially. The values of $E[t_{iT} | \rho_i = 1]$ and $Var[t_{iT} | \rho_i = 1]$ have been computed by IPS via simulations for different values of T and p_i 's.

Breitung (1999) studied the local power of LL and IPS tests statistics against a sequence of local alternatives. Breitung found that the LL and IPS tests suffer from a dramatic loss of power if individual specific trends are included. This is due to the bias correction that also removes the mean under the sequence of local alternatives. The simulation results indicate that the power of LL and IPS tests is very sensitive to the specification of the deterministic terms.

McCoskey and Selden (1998) applied the IPS test for testing unit root for per capita national health care expenditures (HE) and gross domestic product (GDP) for a panel of OECD countries. McCoskey and Selden rejected the null hypothesis that these two series contain unit roots. Gerdtham and Löthgren (1998) claimed that the stationarity found by McCoskey and Selden are driven by the omission of time trends in their ADF regression in (7). Using the IPS test with a time trend, Gerdtham and Löthgren found that both HE and GDP are nonstationary. They concluded that HE and GDP are cointegrated around linear trends following the results of McCoskey and Kao (1999b).

2.3 Combining P-Values Tests

Let G_{iT_i} be a unit root test statistic for the i -th group in (1) and assume that as $T_i \rightarrow \infty$, $G_{iT_i} \Rightarrow G_i$. Let p_i be the p -value of a unit root test for cross-section i , i.e., $p_i = F(G_{iT_i})$, where $F(\cdot)$ is the distribution function of the random variable G_i . Maddala and Wu (1999) and Choi (1999a) proposed a Fisher type test

$$P = -2 \sum_{i=1}^N \ln p_i \quad (12)$$

which combines the p -values from unit root tests for each cross-section i to test for unit root in panel data. P is distributed as χ^2 with $2N$ degrees of freedom as $T_i \rightarrow \infty$ for all N . Maddala, Wu and Liu (1999) argued that the IPS and Fisher tests relax the restrictive assumption of the LL test that ρ_i is the same under the alternative. Both the IPS and Fisher tests combine information based on individual unit root tests. However,

the Fisher test has the advantage over the IPS test in that it does not require a balanced panel. Also, the Fisher test can use different lag lengths in the individual ADF regressions and can be applied to any other unit root tests. The disadvantage is that the p -values have to be derived by Monte Carlo simulations. Choi (1999a) echoes similar advantages of the Fisher test: (1) the cross-sectional dimension, N , can be either finite or infinite, (2) each group can have different types of nonstochastic and stochastic components, (3) the time series dimension, T , can be different for each i , and (4) the alternative hypothesis would allow some groups to have unit roots while others may not.

When N is large, Choi (1999a) proposed a modified P test, P_m ,

$$P_m = \frac{\frac{1}{\sqrt{N}} \sum_{i=1}^N (-2 \ln p_i - 2)}{2}$$

since $E[-2 \ln p_i] = 2$ and $Var[-2 \ln p_i] = 4$. Assume that the p_i 's are *iid* and use the Lindeberg-Levy central limit theorem to get

$$P_m \Rightarrow N(0, 1)$$

as $T_i \rightarrow \infty$ followed by $N \rightarrow \infty$.¹

Choi (1999) applied the Fisher test in (12) and the IPS test in (8) to panel data of real exchange rates and provided evidence in favor of the PPP hypothesis. Choi claimed that this is due to the improved finite sample power of the Fisher test. Maddala and Wu (1999) and Maddala, Wu and Liu (1999) find that the Fisher test is superior to the IPS test, but they argue that these panel unit root tests still do not rescue the PPP hypothesis. When allowance is made for the deficiencies in the panel data unit root tests and panel estimation methods, support for PPP turns out to be weak.

2.4 Residual Based LM Test

Hadri (1999) proposed a residual based Lagrange Multiplier (LM) test for the null that the time series for each i are stationary around a deterministic trend against the alternative of a unit root in panel data. Consider the following model

$$y_{it} = z'_{it} \gamma + r_{it} + \varepsilon_{it} \tag{13}$$

where z_{it} is the deterministic component, r_{it} is a random walk

$$r_{it} = r_{it-1} + u_{it}$$

¹Testing for cointegration in panel data by combining p -values tests is a straightforward extension of the testing procedures in this Section. For cointegration tests, the relevant model is equation (16). We let G_{iT_i} be a test for the null of no-cointegration and apply the same tests and asymptotic theory in this Section.

$u_{it} \sim iid(0, \sigma_u^2)$ and ε_{it} is a stationary process. (13) can be written as

$$y_{it} = \gamma' z_{it} + e_{it} \quad (14)$$

where

$$e_{it} = \sum_{j=1}^t u_{ij} + \varepsilon_{it}.$$

Let \hat{e}_{it} be the residuals from the regression in (14) and $\hat{\sigma}_e^2$ be the estimate of the error variance,

$$\hat{\sigma}_e^2 = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \hat{e}_{it}^2.$$

Also, denote by S_{it} the partial sum process of the residuals,

$$S_{it} = \sum_{j=1}^t \hat{e}_{ij}.$$

Then the LM statistic is

$$LM = \frac{\frac{1}{N} \sum_{i=1}^N \frac{1}{T^2} \sum_{t=1}^T S_{it}^2}{\hat{\sigma}_e^2}.$$

It can be shown that

$$LM \xrightarrow{p} E \left[\int W_{iZ}^2 \right]$$

as $T \rightarrow \infty$ followed by $N \rightarrow \infty$ provided $E \left[\int W_{iZ}^2 \right] < \infty$, where W_{iZ} is defined in (4). Also,

$$\frac{\sqrt{N} (LM - E \left[\int W_{iZ}^2 \right])}{\sqrt{Var \left[\int W_{iZ}^2 \right]}} \Rightarrow N(0, 1)$$

as $T \rightarrow \infty$ followed by $N \rightarrow \infty$.

2.5 Finite Sample Properties of Unit Root Tests

Extensive simulations have been conducted to explore the finite sample performance of panel unit root tests, e.g., Karlsson and Löthgren (1999), Im et. al (1997), Maddala and Wu (1999), and Choi (1999a). Choi (1999a) studied the small sample properties of IPS t -bar test in (8) and Fisher's test in (12). Choi's major findings are the following:

1. The empirical size of the IPS and the Fisher test are reasonably close to their nominal size 0.05 when N is small. But the Fisher test shows mild size distortions at $N = 100$, which is expected from the asymptotic theory. Overall, the IPS t -bar test has the most stable size.
2. In terms of the size-adjusted power, the Fisher test seems to be superior to the IPS t -bar test.
3. When a linear time trend is included in the model, the power of all tests decrease considerably.

3 Spurious Regression in Panel Data

Entorf (1997) studied spurious fixed effects regressions when the true model involves independent random walks with and without drifts. Entorf found that for $T \rightarrow \infty$ and N finite, the nonsense regression phenomenon holds for spurious fixed effects models and inference based on t -values can be highly misleading. Kao (1999) also derived the asymptotic distributions of the least squares dummy variable estimator and various conventional statistics from the spurious regression in panel data. The asymptotic theory employed by Kao (1999) is the sequential limit theory established by Phillips and Moon (1999a) in which $T \rightarrow \infty$ followed by $N \rightarrow \infty$ sequentially.

Suppose that $w_{it} = (u_{it}, \varepsilon'_{it})'$ is a bivariate process with zero mean vector and a long-run covariance matrix given by

$$\Omega_i = \lim_{T \rightarrow \infty} \frac{1}{T} E \left(\sum_{t=1}^T w_{it} \right) \left(\sum_{t=1}^T w_{it} \right)' = \Sigma_i + \Gamma_i + \Gamma_i' = \begin{bmatrix} \Omega_{ui} & \Omega_{u\varepsilon i} \\ \Omega_{\varepsilon ui} & \Omega_{\varepsilon i} \end{bmatrix}, \quad (15)$$

where ε_{it} is a $k \times 1$ vector,

$$\Gamma_i = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{k=1}^{T-1} \sum_{t=k+1}^T E \left(w_{it} w'_{it-k} \right) = \begin{bmatrix} \Gamma_{ui} & \Gamma_{u\varepsilon i} \\ \Gamma_{\varepsilon ui} & \Gamma_{\varepsilon i} \end{bmatrix}$$

and

$$\Sigma_i = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T E \left(w_{it} w'_{it} \right) = \begin{bmatrix} \Sigma_{ui} & \Sigma_{u\varepsilon i} \\ \Sigma_{\varepsilon ui} & \Sigma_{\varepsilon i} \end{bmatrix}.$$

Assume w_{it} satisfies the functional central limit theorem so that

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} w_{it} \Rightarrow \Omega_i^{1/2} \begin{bmatrix} V_i(r) \\ W_i(r) \end{bmatrix} = B(\Omega_i) = \begin{bmatrix} B_{yi}(r) \\ B_{xi}(r) \end{bmatrix}$$

as $T \rightarrow \infty$ for all i , where $[V_i(r), W_i(r)]'$ is a $(k+1) \times 1$ dimensional standard Wiener process. Define

$$\Omega = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \Omega_i,$$

$$\Gamma = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \Gamma_i,$$

and

$$\Sigma = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \Sigma_i.$$

Assume Ω , Γ , and Σ are finite. Let $y_{it} = \sum_{s=1}^t u_{is}$ and $x_{it} = \sum_{s=1}^t \varepsilon_{is}$ in which $u_{i0} = \varepsilon_{i0} = O_p(1)$. Suppose y_{it} and x_{it} , are incorrectly estimated by least squares for all i using panel data; the spurious regression model is

$$y_{it} = x'_{it}\beta + z'_{it}\gamma + e_{it}, \quad (16)$$

where e_{it} is $I(1)$. The OLS estimator of β is

$$\hat{\beta} = \left[\sum_{i=1}^N \sum_{t=1}^T \tilde{x}_{it} \tilde{x}'_{it} \right] \left[\sum_{i=1}^N \sum_{t=1}^T \tilde{x}_{it} \tilde{y}_{it} \right], \quad (17)$$

where \tilde{y}_{it} is defined in (3) and

$$\tilde{x}_{it} = x_{it} - \sum_{s=1}^T h(t, s) x_{is}.$$

Let

$$\begin{aligned} \begin{bmatrix} B_{ui}(r) \\ B_{\varepsilon i}(r) \end{bmatrix} &= \Omega_i^{1/2} \begin{bmatrix} V_i(r) \\ W_i(r) \end{bmatrix} \\ &= \begin{bmatrix} \Omega_{u,\varepsilon i}^{1/2} & \Omega_{u\varepsilon i} \Omega_{\varepsilon i}^{-1/2} \\ 0 & \Omega_{\varepsilon i}^{1/2} \end{bmatrix} \begin{bmatrix} V_i(r) \\ W_i(r) \end{bmatrix}, \end{aligned}$$

where

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

Note that for a fixed N we have

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \frac{1}{T^2} \sum_{t=1}^T \tilde{x}_{it} \tilde{x}'_{it} &\Rightarrow \frac{1}{N} \sum_{i=1}^N \int B_{\varepsilon i Z} B'_{\varepsilon i Z} \\ &= \frac{1}{N} \sum_{i=1}^N \Omega_{\varepsilon i}^{1/2} \left(\int W_{iZ} W'_{iZ} \right) \left(\Omega_{\varepsilon i}^{1/2} \right)' \end{aligned}$$

and

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \frac{1}{T^2} \sum_{t=1}^T \tilde{x}_{it} \tilde{y}_{it} &\Rightarrow \frac{1}{N} \sum_{i=1}^N \int B_{\varepsilon i Z} B_{ui Z} \\ &= \frac{1}{N} \sum_{i=1}^N \int \left[\Omega_{\varepsilon i}^{1/2} W_{iZ} V_{iZ} \Omega_{u,\varepsilon i}^{1/2} + \Omega_{\varepsilon i}^{1/2} W_{iZ} W'_{iZ} \Omega_{\varepsilon i}^{-1/2} \Omega_{u\varepsilon i} \right] \end{aligned}$$

as $T \rightarrow \infty$, where

$$\begin{aligned} W_{iZ}(r) &= W_i(r) - \left[\int W_i Z'_i \right] \left[\int Z Z' \right] Z(r), \\ V_{iZ}(r) &= V_i(r) - \left[\int V_i Z'_i \right] \left[\int Z Z' \right] Z(r), \end{aligned}$$

$$B_{\varepsilon i Z}(r) = B_{x i}(r) - \left[\int B_{x i} Z'_i \right] \left[\int Z Z' \right] Z(r),$$

and

$$B_{u i Z}(r) = B_{y i}(r) - \left[\int B_{y i} Z'_i \right] \left[\int Z Z' \right] Z(r).$$

Assume $\int W_{iZ} W'_{iZ}$ and $\int W_{iZ} V_{iZ}$ are independent across i and then by a law of large numbers we have

$$\begin{aligned} & \frac{1}{N} \sum_{i=1}^N \Omega_{\varepsilon i}^{1/2} \left(\int W_{iZ} W'_{iZ} \right) \left(\Omega_{\varepsilon i}^{1/2} \right)' \\ & \xrightarrow{p} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \Omega_{\varepsilon i}^{1/2} E \left[\int W_{iZ} W'_{iZ} \right] \left(\Omega_{\varepsilon i}^{1/2} \right)' \\ & = E \left[\int W_{iZ} W'_{iZ} \right] \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \Omega_{\varepsilon i} \\ & = E \left[\int W_{iZ} W'_{iZ} \right] \Omega_{\varepsilon}, \end{aligned} \tag{18}$$

$$\begin{aligned} & \frac{1}{N} \sum_{i=1}^N \int \left[\Omega_{\varepsilon i}^{1/2} W_{iZ} V_{iZ} \Omega_{u, \varepsilon i}^{1/2} + \Omega_{\varepsilon i}^{1/2} W_{iZ} W'_{iZ} \Omega_{\varepsilon i}^{-1/2} \Omega_{u, \varepsilon i} \right] \\ & \xrightarrow{p} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \left[\Omega_{\varepsilon i}^{1/2} E(W_{iZ} V_{iZ}) \Omega_{u, \varepsilon i}^{1/2} + \Omega_{\varepsilon i}^{1/2} E(W_{iZ} W'_{iZ}) \Omega_{\varepsilon i}^{-1/2} \Omega_{u, \varepsilon i} \right] \\ & = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \Omega_{\varepsilon i}^{1/2} E(W_{iZ} W'_{iZ}) \Omega_{\varepsilon i}^{-1/2} \Omega_{u, \varepsilon i} \\ & = E(W_{iZ} W'_{iZ}) \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \Omega_{u, \varepsilon i} \\ & = E(W_{iZ} W'_{iZ}) \Omega_{u, \varepsilon}, \end{aligned} \tag{19}$$

using

$$\begin{aligned} z_{it} & E \left[\left(\int W_{iZ} W'_{iZ} \right) \right] \\ 0 & \frac{1}{2} \\ 1 & \frac{1}{2} \\ \mu_i & \frac{1}{6} \mathbf{I}_k \\ (\mu_i, t) & \frac{1}{15} \mathbf{I}_k \end{aligned}$$

and

$$E \left[\int W_{iZ} V_{iZ} \right] = 0$$

as $N \rightarrow \infty$ provided that $E \left[\int W_{iZ} W'_{iZ} \right]$, $Var \left[\int W_{iZ} W'_{iZ} \right]$, $E \left[\int W_{iZ} V_{iZ} \right]$, and $Var \left[\int W_{iZ} V_{iZ} \right]$ are bounded. Using (18) and (19) we have

$$\widehat{\beta} \xrightarrow{p} \left[E \left(\int W_{iZ} W'_{iZ} \right) \Omega_\varepsilon \right]^{-1} \left[E \left(\int W_{iZ} W'_{iZ} \right) \Omega_{u\varepsilon} \right].$$

Then we obtain

$$\widehat{\beta} \xrightarrow{p} \Omega_\varepsilon^{-1} \Omega_{u\varepsilon} \quad (20)$$

as $T \rightarrow \infty$ followed by $N \rightarrow \infty$. (20) shows that the OLS estimator of β , $\widehat{\beta}$, is consistent for its true value, $\Omega_\varepsilon^{-1} \Omega_{u\varepsilon}$. It is known that if a time-series regression for a given i is performed in model (16), the OLS estimator of β is spurious. This is due to the fact that the noise, e_{it} , is as strong as the signal, x_{it} , since both e_{it} and x_{it} are I(1). In the panel regression (16) with a large number of cross-sections, the strong noise of e_{it} is attenuated by pooling the data and a consistent estimate of β can be extracted. The asymptotics of the OLS estimator are very different from those of the spurious regression in pure time series. This has an important consequence for residual-based cointegration tests in panel data, because the null distribution of residual-based cointegration tests depends on the asymptotics of the OLS estimator. This point is explained further in the next section.

4 Panel Cointegration Tests

4.1 Kao Tests

Kao (1999) presented two types of cointegration tests in panel data, the Dickey-Fuller (DF) and augmented Dickey Fuller (ADF) types tests. The DF type tests from Kao can be calculated from the estimated residuals in (16) as:

$$\widehat{e}_{it} = \rho \widehat{e}_{it-1} + v_{it}, \quad (21)$$

where

$$\widehat{e}_{it} = \widetilde{y}_{it} - \widetilde{x}'_{it} \widehat{\beta},$$

$$\widetilde{y}_{it} = x_{it} - \sum_{s=1}^T h(t, s) y_{is},$$

and

$$\widetilde{x}_{it} = x_{it} - \sum_{s=1}^T h(t, s) x_{is}.$$

In order to test the null hypothesis of no cointegration, the null can be written as $H_0 : \rho = 1$. The OLS estimate of ρ and the t -statistic are given as:

$$\hat{\rho} = \frac{\sum_{i=1}^N \sum_{t=2}^T \hat{e}_{it} \hat{e}_{it-1}}{\sum_{i=1}^N \sum_{t=2}^T \hat{e}_{it}^2}$$

and

$$t_\rho = \frac{(\hat{\rho}-1) \sqrt{\sum_{i=1}^N \sum_{t=2}^T \hat{e}_{it-1}^2}}{s_e},$$

where $s_e^2 = \frac{1}{NT} \sum_{i=1}^N \sum_{t=2}^T (\hat{e}_{it} - \hat{\rho} \hat{e}_{it-1})^2$. Kao proposed the following four DF type tests by assuming $z_{it} = \{\mu_i\}$:

$$DF_\rho = \frac{\sqrt{NT}(\hat{\rho} - 1) + 3\sqrt{N}}{\sqrt{10.2}},$$

$$DF_t = \sqrt{1.25}t_\rho + \sqrt{1.875N},$$

$$DF_\rho^* = \frac{\sqrt{NT}(\hat{\rho} - 1) + \frac{3\sqrt{N}\hat{\sigma}_v^2}{\hat{\sigma}_{0v}}}{\sqrt{3 + \frac{36\hat{\sigma}_v^4}{5\hat{\sigma}_{0v}^4}}},$$

and

$$DF_t^* = \frac{t_\rho + \frac{\sqrt{6N}\hat{\sigma}_v}{2\hat{\sigma}_{0v}}}{\sqrt{\frac{\hat{\sigma}_{0v}^2}{2\hat{\sigma}_v^2} + \frac{3\hat{\sigma}_v^2}{10\hat{\sigma}_{0v}^2}}},$$

where $\hat{\sigma}_v^2 = \hat{\Sigma}_u - \hat{\Sigma}_{u\varepsilon} \hat{\Sigma}_\varepsilon^{-1}$ and $\hat{\sigma}_{0v}^2 = \hat{\Omega}_u - \hat{\Omega}_{u\varepsilon} \hat{\Omega}_\varepsilon^{-1}$. While DF_ρ and DF_t are based on the strong exogeneity of the regressors and errors, DF_ρ^* and DF_t^* are for the cointegration with endogenous relationship between regressors and errors. For the *ADF* test, we can run the following regression:

$$\hat{e}_{it} = \rho \hat{e}_{it-1} + \sum_{j=1}^p \vartheta_j \Delta \hat{e}_{it-j} + v_{itp}. \quad (22)$$

With the null hypothesis of no cointegration, the ADF test statistics can be constructed as:

$$ADF = \frac{t_{ADF} + \frac{\sqrt{6N}\hat{\sigma}_v}{2\hat{\sigma}_{0v}}}{\sqrt{\frac{\hat{\sigma}_{0v}^2}{2\hat{\sigma}_v^2} + \frac{3\hat{\sigma}_v^2}{10\hat{\sigma}_{0v}^2}}}$$

where t_{ADF} is the t -statistic of ρ in (22). The asymptotic distributions of DF_ρ , DF_t , DF_ρ^* , DF_t^* , and *ADF* will converge to a standard normal distribution $N(0, 1)$ by the sequential limit theory.

4.2 Residual Based LM Test

McCoskey and Kao (1998) derived a residual-based test for the null of cointegration rather than the null of no cointegration in panels. This test is an extension of the LM test and the locally best invariant (LBI) test for an MA unit root in the time series literature, see Harris and Inder (1994) and Shin (1994). Under the null, the asymptotics no longer depend on the asymptotic properties of the estimating spurious regression, rather the asymptotics of the estimation of a cointegrated relationship are needed. For models which allow the cointegrating vector to change across the cross-sectional observations, the asymptotics depend merely on the time series results as each cross-section is estimated independently. For models with common slopes, the estimation is done jointly and therefore the asymptotic theory is based on the joint estimation of a cointegrated relationship in panel data.

For the residual based test of the null of cointegration, it is necessary to use an efficient estimation technique of cointegrated variables. In the time series literature a variety of methods have been shown to be efficient asymptotically. These include the fully modified (FM) estimator of Phillips and Hansen (1990) and the dynamic least squares (DOLS) estimator as proposed by Saikkonen (1991) and Stock and Watson (1993). For panel data, Kao and Chiang (1999) showed that both the FM and DOLS methods can produce estimators which are asymptotically normally distributed with zero means.

The model presented allows for varying slopes and intercepts:

$$y_{it} = \alpha_i + x'_{it}\beta_i + e_{it}, \quad (23)$$

$$x_{it} = x_{it-1} + \varepsilon_{it} \quad (24)$$

$$e_{it} = \gamma_{it} + u_{it}, \quad (25)$$

and

$$\gamma_{it} = \gamma_{it-1} + \theta u_{it},$$

where u_{it} are *i.i.d* $(0, \sigma_u^2)$. The null of hypothesis of cointegration is equivalent to $\theta = 0$.

The test statistic proposed by McCoskey and Kao (1998) is defined as follows:

$$\overline{LM} = \frac{\frac{1}{N} \sum_{i=1}^N \frac{1}{T^2} \sum_{t=1}^T S_{it}^{+2}}{s^{+2}}, \quad (26)$$

where S_{it} is partial sum process of the residuals,

$$S_{it}^+ = \sum_{j=1}^t \hat{e}_{ij}$$

with

$$s^{+2} = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \hat{e}_{it}^{+2}.$$

The FM estimator non-parametrically corrects for the possible serial correlation and weakly exogenous regressors in a cointegrated regression. The DOLS estimator uses lagged and future differences of x_{it} to correct for these effects.

The asymptotic result for the test is:

$$\sqrt{N}(\overline{LM} - \mu_v) \Rightarrow N(0, \sigma_v^2). \quad (27)$$

The moments, μ_v and σ_v^2 , can be found through Monte Carlo simulation. The limiting distribution of \overline{LM} is then free of nuisance parameters and robust to heteroskedasticity.

Urban economists have long sought to explain the relationship between urbanization levels and output. McCoskey and Kao (1999a) revisited this question and test the long run stability of a production function including urbanization using non-stationary panel data techniques. McCoskey and Kao apply the IPS test and LM in (26) and show that a long run relationship between urbanization, output per worker and capital per worker cannot be rejected for the sample of thirty developing countries or the sample of twenty-two developed countries over the period 1965-1989. They do find, however, that the sign and magnitude of the impact of urbanization varies considerably across the countries. These results offer new insights and potential for dynamic urban models rather than the simple cross-section approach.

4.3 Pedroni Tests

Pedroni (1997) also proposed several tests for the null hypothesis of cointegration in a panel data model that allows for considerable heterogeneity. His tests can be classified into two categories. The first set is similar to the tests discussed above, and involve averaging test statistics for cointegration in the time series across cross-sections. The second set group the statistics such that instead of averaging across statistics, the averaging is done in pieces so that the limiting distributions are based on limits of piecewise numerator and denominator terms.

The first set of statistics as discussed includes a form of the average of the Phillips and Ouliaris (1990) statistic:

$$\tilde{Z}_\rho = \sum_{i=1}^N \frac{\sum_{t=1}^T (\hat{e}_{it-1} \Delta \hat{e}_{it} - \hat{\lambda}_i)}{(\sum_{t=1}^T \hat{e}_{it-1}^2)}, \quad (28)$$

where \hat{e}_{it} is estimated from (16), and $\hat{\lambda}_i = \frac{1}{2} (\hat{\sigma}_i^2 - \hat{s}_i^2)$, for which $\hat{\sigma}_i^2$ and \hat{s}_i^2 are individual long-run and contemporaneous variances respectively of the residual \hat{e}_{it} . For his second set of statistics, Pedroni defines four panel variance ratio statistics. Let $\hat{\Omega}_i$ be a consistent estimate of Ω_i , the long-run variance-covariance matrix given in (15). Define \hat{L}_i to be the lower triangular Cholesky composition of $\hat{\Omega}_i$ such that in the scalar case $\hat{L}_{22i} = \hat{\sigma}_\varepsilon$ and $\hat{L}_{11i} = \hat{\sigma}_u^2 - \frac{\hat{\sigma}_{ue}^2}{\hat{\sigma}_\varepsilon^2}$ is the long-run conditional variance. In this survey we consider only one of these statistics:

$$Z_{t_{\rho_{NT}}} = \frac{\sum_{i=1}^N \sum_{t=2}^T \hat{L}_{11i}^{-2} (\hat{e}_{it-1} \Delta \hat{e}_{it} - \hat{\lambda}_i)}{\sqrt{\hat{\sigma}_{NT}^2 (\sum_{i=1}^N \sum_{t=2}^T \hat{L}_{11i}^{-2} \hat{e}_{it-1}^2)}} , \quad (29)$$

where

$$\tilde{\sigma}_{NT} = \frac{1}{N} \sum_{i=1}^N \frac{\hat{\sigma}_i^2}{\hat{L}_{11i}^2}.$$

It should be noted that Pedroni bases his test on the average of the numerator and denominator terms respectively, rather than the average for the statistic as a whole.

Using results on convergence of functionals of Brownian motion, Pedroni finds the following result:

$$Z_{t_{\rho_{NT}}} + 1.73\sqrt{N} \Rightarrow N(0, 0.93).$$

Note that this distribution applies to the model including an intercept and not including a time trend. Asymptotic results for other model specifications can be found in Pedroni (1997). The intuition on these tests with varying slopes is not straightforward. The convergence in distribution is based on individual convergence of the numerator and denominator terms. What is the intuition of rejection of the null hypothesis? Using the average of the overall test statistic allows more ease in interpretation: rejection of the null hypothesis means that enough of the individual cross-sections have statistics “far away” from the means predicted by theory were they to be generated under the null.

Pedroni (1999) derives asymptotic distributions and critical values for several residual based tests of the null of no cointegration in panels where there are multiple regressors. The model includes regressions with individual specific fixed effects and time trends. Considerable heterogeneity is allowed across individual members of the panel with regards to the associated cointegrating vectors and the dynamics of the underlying error process.

4.4 Likelihood-Based Cointegration Test

Larsson, Lyhagen, and Löthgren (1998) presented a likelihood-based (LR) panel test of cointegrating rank in heterogeneous panel models based on the average of the individual rank trace statistics developed by Johansen (1995). The proposed LR-bar statistic is very similar to the IPS t -bar statistic in (8) - (11). In

Monte Carlo simulation, Larsson et al. (1998) investigated the small sample properties of the standardized LR-bar statistic. They found that the proposed test requires a large time series dimension. Even if the panel has a large cross sectional dimension, the size of the test will be severely distorted.

Groen and Kleibergen (1999) proposed a likelihood-based framework for cointegrating analysis in panels of a fixed number of vector error correction models. Maximum likelihood estimators of the cointegrating vectors are constructed using iterated generalized method of moments (GMM) estimators. Using these estimators Groen and Kleibergen construct likelihood ratio statistics, $LR(\Pi_B|\Pi_A)$, to test for a common cointegration rank across the individual vector error correction models, both with heterogeneous and homogeneous cointegrating vectors. Interestingly, the limiting distribution of $LR(\Pi_B|\Pi_A)$ is invariant to the covariance matrix of the error terms, Ω , which implies that $LR(\Pi_B|\Pi_A)$ is robust with respect to the choices of Ω . Let us define the $LR_s(r|k)$ as the summation of the N individual trace statistics

$$LR_s(r|k) = \sum_{i=1}^N LR_i(r|k) \quad (30)$$

where $LR_i(r|k)$ is the i -th Johansen's likelihood ratio statistic, so that

$$LR_i(r|k) \Rightarrow tr \left(\int dB_{k-r,i} B'_{k-r,i} \left[\int dB_{k-r,i} B'_{k-r,i} \right] \int dB_{k-r,i} B'_{k-r,i} \right)$$

as $T \rightarrow \infty$. Now for a fixed N , it is clear that

$$\begin{aligned} LR_s(r|k) &= \sum_{i=1}^N LR_i(r|k) \\ &\Rightarrow \sum_{i=1}^N tr \left(\int dB_{k-r,i} B'_{k-r,i} \left[\int dB_{k-r,i} B'_{k-r,i} \right] \int dB_{k-r,i} B'_{k-r,i} \right) \end{aligned} \quad (31)$$

as $T \rightarrow \infty$ by a continuous mapping theorem. It follows that $LR_s(r|k)$ is asymptotically equivalent to $LR(\Pi_B|\Pi_A)$ when N is fixed and T is large. This means that nothing is lost by assuming that Ω has zero non-diagonal covariances as far as the asymptotics are concerned for the proposed test statistics in this paper. More importantly, the tests based on the cross-independence like (30) will perform just as well (asymptotically) as the tests based on the cross-dependence such as $LR(\Pi_B|\Pi_A)$. Groen and Kleibergen verified that the likelihood-based cointegration tests proposed by Larsson, Lyhagen and Löthgran (1998) in (30) are robust with respect to the cross-dependence in panel data. The (asymptotic) equivalence of $LR_s(r|k)$ and $LR(\Pi_B|\Pi_A)$ found in Groen and Kleibergen has profound implications to econometricians and applied economists, e.g., there exists tests/estimators based on the cross-independence which are equivalent to tests/estimators based on the cross-dependence in nonstationary panel time series. Define $\overline{LR}(r|k)$ be

the average of $LR_i(r|k)$:

$$\overline{LR}(r|k) = \frac{1}{N} LR_s(r|k) = \frac{1}{N} \sum_{i=1}^N LR_i(r|k).$$

It can be shown that

$$\frac{\overline{LR}(r|k) - E[\overline{LR}(r|k)]}{\text{Var}[\overline{LR}(r|k)]} \Rightarrow N(0, 1)$$

as $T \rightarrow \infty$ followed by $N \rightarrow \infty$ by a continuous mapping theorem and a central limit theorem provided $E[\overline{LR}(r|k)]$ and $\text{Var}[\overline{LR}(r|k)]$ are bounded. Define

$$\overline{LR}(\Pi_B|\Pi_A) = \frac{1}{N} LR(\Pi_B|\Pi_A). \quad (32)$$

For a fixed N , it is easy to show that

$$\begin{aligned} \overline{LR}(\Pi_B|\Pi_A) &= \frac{1}{N} LR(\Pi_B|\Pi_A) \\ &\Rightarrow \frac{1}{N} \sum_{i=1}^N \text{tr} \left(\int dB_{k-r,i} B'_{k-r,i} \left[\int dB_{k-r,i} B'_{k-r,i} \right] \int dB_{k-r,i} B'_{k-r,i} \right) \\ &= \frac{1}{N} \sum_{i=1}^N Z_{ki} \end{aligned}$$

where

$$Z_{ki} = \text{tr} \left(\int dB_{k-r,i} B'_{k-r,i} \left[\int dB_{k-r,i} B'_{k-r,i} \right] \int dB_{k-r,i} B'_{k-r,i} \right)$$

as $T \rightarrow \infty$. Then

$$\frac{\frac{1}{N} \sum_{i=1}^N Z_{ki} - E\left[\frac{1}{N} \sum_{i=1}^N Z_{ki}\right]}{\text{Var}\left[\frac{1}{N} \sum_{i=1}^N Z_{ki}\right]} \Rightarrow N(0, 1)$$

as $N \rightarrow \infty$ since $B_{k-r,i}$ and $B_{k-r,j}$ are independent for $i \neq j$. It implies that

$$\frac{\overline{LR}(\Pi_B|\Pi_A) - E[\overline{LR}(\Pi_B|\Pi_A)]}{\text{Var}[\overline{LR}(\Pi_B|\Pi_A)]} \Rightarrow N(0, 1)$$

as $T \rightarrow \infty$ followed by $N \rightarrow \infty$. The above discussion indicates that $\overline{LR}(r|k)$ and $\overline{LR}(\Pi_B|\Pi_A)$ are also equivalent when T and N are large.

Groen and Kleibergen (1999) applied $LR(\Pi_B|\Pi_A)$ to a data set of exchange rates and appropriate monetary fundamentals. They found strong evidence for the validity of the monetary exchange rate model within a panel of vector correction models for three major European countries, whereas the results based on individual vector error correction models for each of these countries separately are less supportive.

4.5 Finite Sample Properties

McCoskey and Kao (1999b) conducted Monte Carlo experiments to compare the size and power of different residual based tests for cointegration in heterogeneous panel data: varying slopes and varying intercepts. Two of the tests are constructed under the null hypothesis of no cointegration. These tests are based on the average ADF test and Pedroni's pooled tests in (28) - (29). The third test is based on the null hypothesis of cointegration which is based on the McCoskey and Kao LM test in (26). Wu and Yin (1999) performed a similar comparison for panel tests in which they consider only tests for which the null hypothesis is that of no cointegration. Wu and Yin (1999) compared *ADF* statistics with maximum eigenvalue statistics in pooling information on means and *p*-values respectively. They found that the average *ADF* performs better with respect to power and their maximum eigenvalue based *p*-value performs better with regards to size.

The test of the null hypothesis was originally proposed in response to the low power of the tests of the null of no cointegration, especially in the time series case. Further, in cases where economic theory predicted a long run steady state relationship, it seemed that a test of the null of cointegration rather than the null of no cointegration would be appropriate. The results from the Monte Carlo study showed that the McCoskey and Kao LM test outperforms the other two tests.

Of the two reasons for the introduction of the test of the null hypothesis of cointegration, low power and attractiveness of the null, the introduction of the cross-section dimension of the panel solves one: all of the tests show decent power when used with panel data. For those applications where the null of cointegration is more logical than the null of no cointegration, McCoskey and Kao (1999b), at a minimum, conclude that using the McCoskey and Kao LM test does not compromise the ability of the researcher in determining the underlying nature of the data.

Recently, Hall et al. (1999) proposed a new approach based on principal components analysis to test for the number of common stochastic trends driving the non-stationary series in a panel data set. The test is consistent even if there is a mixture of $I(0)$ and $I(1)$ series in the sample. This makes it unnecessary to pretest the panel for unit root. It also has the advantage of solving the problem of dimensionality encountered in large panel data sets.

5 Estimation and Inference in Panel Cointegration Models

5.1 Panel OLS

In this section, we provide a brief discussion of the OLS estimation methods in a panel cointegrated model. Consider the following panel regression:

$$y_{it} = x'_{it}\beta + z'_{it}\gamma + u_{it}, \quad (33)$$

where $\{y_{it}\}$ are 1×1 , β is a $k \times 1$ vector of the slope parameters, z_{it} is the deterministic component, 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, (33) describes a system of cointegrated regressions, i.e., y_{it} is cointegrated with x_{it} . Next, we characterize the innovation vector $w_{it} = (u_{it}, \varepsilon'_{it})'$. Assume the partial sum process $\frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} w_{it}$ satisfies the following multivariate invariance principle:

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[Tr]} w_{it} \Rightarrow B_i(r) \equiv BM_i(\Omega) \text{ as } T \rightarrow \infty, \quad (34)$$

where

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

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

$$\begin{aligned} \Omega_i &= \sum_{j=-\infty}^{\infty} E(w_{ij}w'_{i0}) \\ &= \Pi_i(1)\Sigma_\varepsilon\Pi_i(1)' \\ &= \Sigma_i + \Gamma_i + \Gamma_i' \\ &= \begin{bmatrix} \Omega_{ui} & \Omega_{u\varepsilon i} \\ \Omega_{\varepsilon ui} & \Omega_{\varepsilon i} \end{bmatrix}, \end{aligned}$$

where

$$\Gamma_i = \sum_{j=1}^{\infty} E(w_{ij}w'_{i0}) \equiv \begin{bmatrix} \Gamma_{ui} & \Gamma_{u\varepsilon i} \\ \Gamma_{\varepsilon ui} & \Gamma_{\varepsilon i} \end{bmatrix} \quad (35)$$

and

$$\Sigma_i = E(w_{i0}w'_{i0}) \equiv \begin{bmatrix} \Sigma_{ui} & \Sigma_{u\varepsilon i} \\ \Sigma_{\varepsilon ui} & \Sigma_{\varepsilon i} \end{bmatrix} \quad (36)$$

are partitioned conformably with w_{it} . Define

$$\Omega_{u.\varepsilon i} = \Omega_{ui} - \Omega_{u\varepsilon i} \Omega_{\varepsilon i}^{-1} \Omega_{\varepsilon ui}. \quad (37)$$

Then, B_i can be rewritten as

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

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

$$\begin{aligned} \Delta_i &= \Sigma_i + \Gamma_i \\ &= \sum_{j=0}^{\infty} E(w_{ij} w'_{i0}) \end{aligned}$$

with

$$\Delta_i = \begin{bmatrix} \Delta_{ui} & \Delta_{u\varepsilon i} \\ \Delta_{\varepsilon ui} & \Delta_{\varepsilon i} \end{bmatrix}.$$

The OLS estimator of β is

$$\widehat{\beta}_{OLS} = \left[\sum_{i=1}^N \sum_{t=1}^T \tilde{x}_{it} \tilde{x}'_{it} \right]^{-1} \left[\sum_{i=1}^N \sum_{t=1}^T \tilde{x}_{it} \tilde{y}_{it} \right]. \quad (38)$$

It follows that

$$\begin{aligned} &\sqrt{NT} \left(\widehat{\beta}_{OLS} - \beta \right) \\ &= \left[\frac{1}{N} \sum_{i=1}^N \frac{1}{T^2} \sum_{t=1}^T \tilde{x}_{it} \tilde{x}'_{it} \right]^{-1} \left[\sqrt{N} \frac{1}{N} \sum_{i=1}^N \frac{1}{T} \sum_{t=1}^T \tilde{x}_{it} \tilde{u}_{it} \right] \\ &= \left[\frac{1}{N} \sum_{i=1}^N \zeta_{2iT} \right]^{-1} \left[\sqrt{N} \frac{1}{N} \sum_{i=1}^N \zeta_{1iT} \right] \\ &= [\zeta_{2NT}]^{-1} \sqrt{N} \zeta_{1NT}, \end{aligned}$$

where $\zeta_{1iT} = \frac{1}{T} \sum_{t=1}^T \tilde{x}_{it} \tilde{u}_{it}$, $\zeta_{2iT} = \frac{1}{T^2} \sum_{t=1}^T \tilde{x}_{it} \tilde{x}'_{it}$, $\zeta_{1NT} = \frac{1}{N} \sum_{i=1}^N \zeta_{1iT}$, and $\zeta_{2NT} = \frac{1}{N} \sum_{i=1}^N \zeta_{2iT}$. It is easy to show that

$$\begin{aligned} \zeta_{2iT} &= \frac{1}{T^2} \sum_{t=1}^T \tilde{x}_{it} \tilde{x}'_{it} \\ &\Rightarrow \Omega_{\varepsilon i}^{1/2} \left(\int W_{iZ} W_{iZ}' \right) \left(\Omega_{\varepsilon i}^{1/2} \right)' = \zeta_{2i}, \end{aligned} \quad (39)$$

as $T \rightarrow \infty$ for all i , and

$$\begin{aligned} \zeta_{1iT} &= \frac{1}{T} \sum_{t=1}^T \tilde{x}_{it} \tilde{u}_{it} \\ &\Rightarrow \Omega_{\varepsilon i}^{1/2} \left(\int W_{iZ} dV_{iZ} \right) \Omega_{u.\varepsilon i}^{1/2} + \Omega_{\varepsilon i}^{1/2} \left(\int W_{iZ} dW_{iZ}' \right) \Omega_{\varepsilon i}^{-1/2} \Omega_{\varepsilon ui} + \Delta_{\varepsilon ui} \\ &= \zeta_{1i}, \end{aligned} \quad (40)$$

as $T \rightarrow \infty$ for all i , where $\Delta_{\varepsilon ui} = \Sigma_{\varepsilon ui} + \Gamma_{\varepsilon ui}$.

Using

$$\begin{array}{lll} z_{it} & E[\int W_{iZ} dW_{iZ}] & E\left[\left(\int W_{iZ} W_{iZ}'\right)\right] \\ 0 & 0 & \frac{1}{2} \\ 1 & 0 & \frac{1}{2} \\ \mu_i & -\frac{1}{2}\mathbf{I}_k & \frac{1}{6}\mathbf{I}_k \\ (\mu_i, t) & -\frac{1}{2}\mathbf{I}_k & \frac{1}{15}\mathbf{I}_k \end{array}$$

where \mathbf{I}_K is a $k \times k$ identity matrix. It then follows that

$$\begin{array}{lll} z_{it} & E[\zeta_{1i}] & E[\zeta_{2i}] \\ 0 & 0 & \frac{1}{2} \\ 1 & 0 & 0 \\ \mu_i & -\frac{1}{2}\Omega_{\varepsilon ui} + \Delta_{\varepsilon ui} & \frac{1}{6}\Omega_{\varepsilon i} \\ (\mu_i, t) & -\frac{1}{2}\Omega_{\varepsilon ui} + \Delta_{\varepsilon ui} & \frac{1}{15}\Omega_{\varepsilon i} \end{array}$$

Assume that both ζ_{1i} and ζ_{2i} are independent across i . Finally, by a law of large numbers

$$\frac{1}{N} \sum_{i=1}^N \zeta_{1i} \xrightarrow{p} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N E[\zeta_{1i}],$$

and

$$\frac{1}{N} \sum_{i=1}^N \zeta_{2i} \xrightarrow{p} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N E[\zeta_{2i}],$$

as $N \rightarrow \infty$. It follows that

$$\xi_{1NT} = \frac{1}{N} \sum_{i=1}^N \zeta_{1iT} \xrightarrow{p} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N E[\zeta_{1i}],$$

and

$$\xi_{2NT} = \frac{1}{N} \sum_{i=1}^N \zeta_{2iT} \xrightarrow{p} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N E[\zeta_{2i}].$$

Then

$$\begin{array}{lll} z_{it} & \xi_{1NT} & \xi_{2NT} \\ 0 & 0 & \frac{1}{2} \\ 1 & 0 & 0 \\ \mu_i & -\frac{1}{2}\Omega_{\varepsilon ui} + \Delta_{\varepsilon ui} & \frac{1}{6}\Omega_{\varepsilon i} \\ (\mu_i, t) & -\frac{1}{2}\Omega_{\varepsilon ui} + \Delta_{\varepsilon ui} & \frac{1}{15}\Omega_{\varepsilon i} \end{array}$$

Define

$$\zeta_{1iT}^* = \zeta_{1iT} - \Omega_{\varepsilon i}^{1/2} \left(\int W_{iZ} dW_{iZ}' \right) \Omega_{\varepsilon i}^{-1/2} \Omega_{\varepsilon ui} - \Delta_{\varepsilon ui}$$

and

$$\xi_{1NT}^* = \frac{1}{N} \sum_{i=1}^N \zeta_{1iT}^*.$$

First, we note

$$\zeta_{1iT}^* \Rightarrow \Omega_{\varepsilon i}^{1/2} \left(\int W_{iZ} dV_{iZ} \right) \Omega_{u,\varepsilon i}^{1/2}$$

as $T \rightarrow \infty$. Clearly,

$$E \left[\int W_{iZ} dV_{iZ} \right] = 0.$$

It follows that

$$\frac{1}{N} \sum_{i=1}^N \zeta_{1iT}^* \xrightarrow{p} 0$$

by the sequential limit theory. The variance-covariance of $\Omega_{\varepsilon i}^{1/2} \left(\int W_{iZ} dV_{iZ} \right) \Omega_{u,\varepsilon i}^{1/2}$

$$\text{Var} \left[\Omega_{\varepsilon i}^{1/2} \left(\int W_{iZ} dV_{iZ} \right) \Omega_{u,\varepsilon i}^{1/2} \right] \quad (41)$$

can be found easily. For example,

$$\text{Var} \left[\Omega_{\varepsilon i}^{1/2} \left(\int W_{iZ} dV_{iZ} \right) \Omega_{u,\varepsilon i}^{1/2} \right] = \frac{1}{6} \Omega_{u,\varepsilon i} \Omega_{\varepsilon i}$$

with $z_{it} = \{\mu_i\}$. Since $\frac{1}{6} \Omega_{u,\varepsilon i} \Omega_{\varepsilon i}$ in (41) is finite, from the multivariate Lindeberg-Levy central limit theorem we have

$$\sqrt{N} \xi_{1NT}^* \Rightarrow N \left(0, \frac{1}{6} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \Omega_{u,\varepsilon i} \Omega_{\varepsilon i} \right)$$

as $N \rightarrow \infty$. Using the Slutsky theorem, we obtain

$$[\xi_{2NT}]^{-1} \sqrt{N} \xi_{1NT}^* \Rightarrow N \left(0, 6\Omega_{\varepsilon}^{-1} \left(\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \Omega_{u,\varepsilon i} \Omega_{\varepsilon i} \right) \Omega_{\varepsilon}^{-1} \right).$$

where

$$\Omega_{\varepsilon} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \Omega_{\varepsilon i}.$$

Hence,

$$\sqrt{NT} \left(\widehat{\beta}_{OLS} - \beta \right) - \sqrt{N} \delta_{NT} \Rightarrow N \left(0, \Omega_{\varepsilon}^{-1} \left(\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \Omega_{u,\varepsilon i} \Omega_{\varepsilon i} \right) \Omega_{\varepsilon}^{-1} \right),$$

by the sequential limit theory, where

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

Kao and Chiang (1999) in this volume derived limiting distributions for the fully-modified (FM), and dynamic least squares (DOLS) estimators in a cointegrated regression and showed they are asymptotically normal. Phillips and Moon (1999a) and Pedroni (1996) also obtained similar results for the FM estimator. The reader is referred to the cited papers for further details and discussions. Kao and Chiang (1999) also investigated the finite sample properties of the OLS, FM, and DOLS estimators. They found that (i) the OLS estimator has a non-negligible bias in finite samples, (ii) the FM estimator does not improve over the OLS estimator in general, and (iii) the DOLS estimator may be more promising than OLS or FM estimators in estimating the cointegrated panel regressions.

Choi (1999b) extended Kao and Chiang (1999) to study asymptotic properties of OLS, Within and GLS estimators for an error component model. The error component model involves both stationary and nonstationary regressors. Choi's simulation results indicated that the feasible GLS estimator is more efficient than the Within estimator. Choi (1999c) studied instrumental variable estimation for an error component model with stationary and nearly nonstationary regressors.

Phillips and Moon (1999a) studied various regressions between two panel vectors that may or may not have cointegrating relations, and present a fundamental framework for studying sequential and joint limit theories in nonstationary panel data. In particular, Phillips and Moon (1999a) studied regression limit theory of nonstationary panels when both N and T go to infinity. Their limit theory allows for both sequential limits, where $T \rightarrow \infty$ followed by $N \rightarrow \infty$ and joint limits, where $T, N \rightarrow \infty$ simultaneously. Phillips and Moon (1999a) require that $N/T \rightarrow 0$, so that this theory applies for moderate N and large T macro panel data and not large N and small T micro panel data. The panel models considered allow for no cointegration, heterogeneous cointegration, homogeneous, and near-homogeneous cointegration. Phillips and Moon (1999a) showed that in all cases considered the pooled estimator is consistent and has a normal limiting distribution. They also showed that for the homogeneous and near homogeneous cointegration cases, a consistent estimator of the long-run regression coefficient can be constructed which they call a pooled FM estimator. They showed that this estimator has faster convergence rate than the simple cross-section and time series estimators. See also Phillips and Moon (1999b) for a review of some recent developments in panel data asymptotics.

Moon and Phillips (1999) investigate the asymptotic properties of the Gaussian MLE of the localizing parameter in local to unity dynamic panel regression models with deterministic and stochastic trends. Moon and Phillips find that for the homogeneous trend model, the Gaussian MLE of the common localizing parameter is \sqrt{N} consistent, while for the heterogeneous trends model, it is inconsistent. The latter inconsistency is due to the presence of an infinite number of incidental parameters (as $N \rightarrow \infty$) for the individual trends. Unlike the fixed effects dynamic panel data model where this inconsistency due to the incidental parameter

problem disappears as $T \rightarrow \infty$, the inconsistency of the localizing parameter in the Moon and Phillips model persists even when both N and T go to infinity.

Pesaran, Shin, and Smith (1999) derived the asymptotics of a pooled mean group (PMG) estimator. The PMG estimation constrains the long-run coefficients to be identical, but allow the short run and adjustment coefficients as the error variances to differ across the cross-sectional dimension.

Kauppi (1999) discussed issues that arise in the estimation and inference of panel cointegrated regressions with near integrated regressors. Kauppi showed that a bias corrected pooled OLS for a common cointegrating parameter has an asymptotic normal distribution centered on the true value irrespective of whether the regressor has near or exact unit root. However, if the regression model contains individual effects and/or deterministic trends, then Kauppi's bias corrected pooled OLS still produces asymptotic bias. Kauppi also showed that the panel FM estimator is subject to asymptotic bias regardless of how individual effects and/or deterministic trends are contained if the regressor has near unit root. This indicates that much care should be taken in interpreting empirical results achieved by the recent panel cointegration methods that assume exact unit root.

Kao et al. (1999) apply the asymptotic theory of panel cointegration developed by Kao and Chiang (1999) to the Coe and Helpman (1995) international R&D spillover regression. Using a sample of 21 OECD countries and Israel, they re-examine the effects of domestic and foreign R&D capital stocks on total factor productivity of these countries. They find that OLS with bias-correction, the fully modified (FM) and the dynamic OLS (DOLS) estimators produce different predictions about the impact of foreign R&D on total factor productivity (TFP) although all the estimators support the result that domestic R&D is related to TFP. Kao et al.'s empirical results indicate that the estimated coefficients in the Coe and Helpman's regressions are subject to estimation bias. Given the superiority of the DOLS over FM as suggested by Kao and Chiang (1999), Kao et al. leaned towards rejecting the Coe and Helpman hypothesis that international R&D spillovers are trade related.

Funk (1998) examined the relationship between trade patterns and international R&D spillovers among the OECD countries using the panel cointegration methods developed by Kao (1999), Kao and Chiang (1999), and Pesaran, Shin, and Smith (1999). Using randomly simulated bilateral trade patterns, Funk found that the choice of weights used in constructing foreign R&D stocks is informative of the avenue of spillover transmission when panel cointegration methods are employed. A re-examination of the relationship between import patterns and R&D spillovers found no evidence to link the patterns of R&D spillovers to the patterns of imports. Funk found strong evidence indicating that exporters receive substantial R&D spillovers from their customers.

6 Dynamic Panel Data Models

This section surveys recent developments in dynamic panel data models. The dynamic panel data regression is characterized by two sources of persistence over time. Autocorrelation due to the presence of a lagged dependent variable among the regressors and individual effects characterizing the heterogeneity among the individuals

$$y_{it} = \delta y_{i,t-1} + X'_{it}\beta + \mu_i + \nu_{it} \quad (42)$$

for $i=1,2,\dots,N$; and $t=1,2,\dots,T$. δ is a scalar, X'_{it} is $1 \times K$, μ_i denotes the i -th individuals effect and ν_{it} is the remainder disturbance. Basic introductions to this topic are found in Hsiao (1986), Baltagi (1995) and Matyas and Sevestre (1996). Applications using this model are too many to enumerate. These include employment equations, see Arellano and Bond (1991), liquor demand, see Baltagi and Griffin (1995), growth convergence, see Islam (1995) and Nerlove (1999), life cycle labor supply models, see Ziliak (1992), and demand for gasoline, see Baltagi and Griffin (1997) to mention a few.

It is well known that for typical micro-panels where there are a large number of firms or individuals (N) observed over a short period of time (T), the fixed effects (FE) estimator is biased and inconsistent (since T is fixed and $N \rightarrow \infty$), see Nickell (1981) and more recently Kiviet (1995,1999). Monte Carlo results have shown that first order asymptotic properties do not necessarily yield correct inference in finite samples. Therefore, Kiviet (1995) examined higher order asymptotics which may approximate the actual finite sample properties more closely and lead to better inference. In fact, Kiviet (1995) considered the simple dynamic linear panel data model with serially uncorrelated disturbances and strongly exogenous regressors and derived an approximation for the bias of the FE estimator. When a consistent estimator of this bias is subtracted from the original FE estimator, a corrected FE estimator results. This corrected FE estimator performed well in simulations when compared with eight other consistent instrumental variable or GMM estimators.²

In macro-panels studying for example long run growth, the data covers a large number of countries N over a moderate size T . In this case, T is not very small relative to N . Hence, some researchers may still favor the FE estimator arguing that its bias may not be large. Judson and Owen (1999) performed some Monte Carlo experiments for $N=20$ or 100 and $T=5,10,20$ and 30 and found that the bias in the FE can be sizeable, even when $T=30$. The bias of the FE estimator increases with δ and decreases with T . But even for $T=30$, this bias could be as much as 20% of the true value of the coefficient of interest. Judson and Owen (1999) recommend the corrected FE estimator proposed by Kiviet (1995) as the best choice, GMM being

²Kiviet(1999) extends this derivation to the case of weakly exogenous variables and examines to what degree this order of approximation is determined by the initial conditions of the dynamic panel model.

second best and for long panels, the computationally simpler Anderson and Hsiao (1982) estimator. This last estimator first differences the data to get rid of the individual effects and then uses lagged predetermined variables in levels as instruments.³ Arellano and Bond (1991) proposed GMM procedures that are more efficient than the Anderson and Hsiao (1982) estimator. Ahn and Schmidt (1995) derive additional nonlinear moment restrictions not exploited by the Arellano and Bond (1991) GMM estimator.⁴ Ahn and Schmidt (1995, 1997) also give a complete count of the set of orthogonality conditions corresponding to a variety of assumptions imposed on the disturbances and the initial conditions of the dynamic panel data model. While many of the moment conditions are nonlinear in the parameters, Ahn and Schmidt (1997) propose a linearized GMM estimator that is asymptotically as efficient as the nonlinear GMM estimator. They also provide simple moment tests of the validity of these nonlinear restrictions. In addition, they investigate the circumstances under which the optimal GMM estimator is equivalent to a linear instrumental variable estimator. They find that these circumstances are quite restrictive and go beyond uncorrelatedness and homoskedasticity of the errors. Ahn and Schmidt (1995) provide some evidence on the efficiency gains from the nonlinear moment conditions which provide support for their use in practice. By employing all these conditions, the resulting GMM estimator is asymptotically efficient and has the same asymptotic variance as the MLE under normality. In fact, Hahn (1997) showed that GMM based on an increasing set of instruments as $N \rightarrow \infty$ would achieve the semiparametric efficiency bound.

Hahn (1997) considers the asymptotic efficient estimation of the dynamic panel data model with sequential moment restrictions in an environment with i.i.d. observations. Hahn (1997) shows that the GMM estimator with an increasing set of instruments as the sample size grows attains the semiparametric efficiency bound of the model. Hahn (1997) explains how Fourier series or polynomials may be used as the set of instruments for efficient estimation. In a limited Monte Carlo comparison, Hahn finds that this estimator has similar finite sample properties as the Keane and Runkle (1992) and/or Schmidt, et. al. (1992) estimators when the latter estimators are efficient. In cases where the latter estimators are not efficient, the Hahn efficient estimator outperforms both estimators in finite samples.

Recently, Wansbeek and Bekker (1996) considered a simple dynamic panel data model with no exogenous regressors and disturbances ν_{it} and random effects μ_i that are independent and normally distributed. They derived an expression for the optimal instrumental variable estimator, i.e., one with minimal asymptotic variance. A striking result is the difference in efficiency between the IV and ML estimators. They find that

³Arellano(1989) found that using lagged differences of predetermined variables as instruments is not recommended since it has a singularity point and very large variances over a significant range of the parameter values.

⁴See also Arellano and Bover(1995), chapter 8 of Baltagi (1995) and chapters 6 and 7 of Matyas and Sevestre (1996) for more details.

for regions of the autoregressive parameter δ which are likely in practice, ML is superior. The gap between IV (or GMM) and ML can be narrowed down by adding moment restrictions of the type considered by Ahn and Schmidt (1995). Hence, Wansbeek and Bekker (1996) find support for adding these nonlinear moment restrictions and warn against the loss in efficiency as compared with MLE by ignoring them.

Blundell and Bond (1998) revisit the importance of exploiting the initial condition in generating efficient estimators of the dynamic panel data model when T is small. They consider a simple autoregressive panel data model with no exogenous regressors

$$y_{it} = \delta y_{i,t-1} + \mu_i + \nu_{it} \quad (43)$$

with $E(\mu_i) = 0$; $E(\nu_{it}) = 0$; and $E(\mu_i \nu_{it}) = 0$ for $i=1,2,\dots,N$; $t=1,2,\dots,T$. Blundell and Bond (1998) focus on the case where $T=3$ and therefore there is only one orthogonality condition given by $E(y_{i1} \Delta \nu_{i3}) = 0$, so that δ is just-identified. In this case, the first stage IV regression is obtained by running Δy_{i2} on y_{i1} . Note that this regression can be obtained from (2) evaluated at $t=2$ by subtracting y_{i1} from both sides of this equation, i.e.,

$$\Delta y_{i2} = (\delta - 1)y_{i,1} + \mu_i + \nu_{i2} \quad (44)$$

Since we expect $E(y_{i1} \mu_i) > 0$, $(\delta - 1)$ will be biased upwards with

$$plim(\hat{\delta} - 1) = (\delta - 1) \frac{c}{c + (\sigma_\mu^2 / \sigma_\nu^2)} \quad (45)$$

where $c = (1 - \delta)/(1 + \delta)$. The bias term effectively scales the estimated coefficient on the instrumental variable y_{i1} towards zero. They also find that the F-statistic of the first stage IV regression converges to χ_1^2 with noncentrality parameter

$$\tau = \frac{(\sigma_\nu^2 c)^2}{\sigma_\mu^2 + \sigma_\nu^2 c} \rightarrow 0 \text{ as } \delta \rightarrow 1 \quad (46)$$

As $\tau \rightarrow 0$, the instrumental variable estimator performs poorly. Hence, Blundell and Bond attribute the bias and the poor precision of the first difference GMM estimator due to the problem of weak instruments described in Nelson and Startz (1990) and Staiger and Stock (1997) and characterize this weak IV by its concentration parameter τ .

Next, Blundell and Bond (1998) show that an additional mild stationarity restriction on the initial conditions process allows the use of an extended system GMM estimator that uses lagged differences of y_{it} as instruments for equations in levels, in addition to lagged levels of y_{it} as instruments for equations in first differences, see Arellano and Bover (1995). The system GMM estimator is shown to have dramatic

efficiency gains over the basic first difference GMM as $\delta \rightarrow 1$ and $(\sigma_\mu^2/\sigma_v^2)$ increases. In fact, for $T=4$ and $(\sigma_\mu^2/\sigma_v^2) = 1$, the asymptotic variance ratio of the first difference GMM estimator to this system GMM estimator is 1.75 for $\delta = 0$ and increases to 3.26 for $\delta = 0.5$ and 55.4 for $\delta = 0.9$. This clearly demonstrates that the levels restrictions suggested by Arellano and Bover (1995) remain informative in cases where first differenced instruments become weak. Things improve for first difference GMM as T increases. However, with short T and persistent series, the Blundell and Bond findings support the use of the extra moment conditions. Hahn (1999) examines the role of the initial condition imposed by the Blundell and Bond (1998) estimator. This is done by numerically comparing the semiparametric information bounds for the case that incorporates the stationarity of the initial condition and the case that does not. Hahn (1999) finds that the efficiency gain can be substantial.

Ziliak (1997) asks the question whether the bias/efficiency trade-off for the GMM estimator considered by Tauchen (1986) for the time series case is still binding in panel data where the sample size is normally larger than 500. For time series data, Tauchen (1986) shows that even for $T=50$ or 75 there is a bias/efficiency trade-off as the number of moment conditions increase. Therefore, Tauchen recommends the use of sub-optimal instruments in small samples. This result was also corroborated by Andersen and Sorensen (1996) who argue that GMM using too few moment conditions is just as bad as GMM using too many moment conditions. This problem becomes more pronounced with panel data since the number of moment conditions increase dramatically as the number of strictly exogenous variables and the number of time series observations increase. Even though it is desirable from an asymptotic efficiency point of view to include as many moment conditions as possible, it may be infeasible or impractical to do so in many cases. For example, for $T=10$ and five strictly exogenous regressors, this generates 500 moment conditions for GMM. Ziliak (1997) performs an extensive set of Monte Carlo experiments for a dynamic panel data model and finds that the same tradeoff between bias and efficiency exists for GMM as the number of moment conditions increase, and that one is better off with sub-optimal instruments. In fact, Ziliak finds that GMM performs well with suboptimal instruments, but is not recommended for panel data applications when all the moments are exploited for estimation.⁵ Ziliak estimates a life cycle labor supply model under uncertainty based on 532 men observed over 10 years of data (1978-1987) from the panel study of income dynamics. The sample was restricted to continuously married, continuously working prime age men aged 22-51 in 1978. These men were paid an

⁵For a Hausman and Taylor (1981) type model, Metcalf (1996) shows that using less instruments may lead to a more powerful Hausman specification test. Asymptotically, more instruments lead to more efficient estimators. However, the asymptotic bias of the less efficient estimator will also be greater as the null hypothesis of no correlation is violated. Metcalf argues that if the bias increases at the same rate as the variance (as the null is violated) for the less efficient estimator, then the power of the Hausman test will increase. This is due to the fact that the test statistic is linear in variance but quadratic in bias.

hourly wage or salaried and could not be piece-rate workers or self-employed. Ziliak finds that the downward bias of GMM is quite severe as the number of moment conditions expands, outweighing the gains in efficiency. Ziliak reports estimates of the intertemporal substitution elasticity which is the focal point of interest in the labor supply literature. This measures the intertemporal changes in hours of work due to an anticipated change in the real wage. For GMM, this estimate changes from .519 to .093 when the number of moment conditions used in GMM are increased from 9 to 212. The standard error of this estimate drops from .36 to .07. Ziliak attributes this bias to the correlation between the sample moments used in estimation and the estimated weight matrix. Interestingly, Ziliak finds that the forward filter 2SLS estimator proposed by Keane and Runkle (1992) performs best in terms of the bias/efficiency trade-off and is recommended. Forward filtering eliminates all forms of serial correlation while still maintaining orthogonality with the initial instrument set. Schmidt, Ahn and Wyhowski (1992) argued that filtering is irrelevant if one exploits all sample moments during estimation. However, in practice, the number of moment conditions increases with the number of time periods T and the number of regressors K and can become computationally intractable. In fact for $T=15$ and $K=10$, the number of moment conditions for Schmidt, et al. (1992) is $T(T-1)K/2$ which is 1040 restrictions, highlighting the computational burden of this approach. In addition, Ziliak argues that the overidentifying restrictions are less likely to be satisfied possibly due to the weak correlation between the instruments and the endogenous regressors.⁶ In this case, the forward filter 2SLS estimator is desirable yielding less bias than GMM and sizeable gains in efficiency. In fact, for the life cycle labor example, the forward filter 2SLS estimate of the intertemporal substitution elasticity was .135 for 9 moment conditions compared to .296 for 212 moment conditions. The standard error of these estimates dropped from .32 to .09.

The practical problem of not being able to use more moment conditions as well as the statistical problem of the trade-off between small sample bias and efficiency prompted Ahn and Schmidt (1999a) to pose the following questions: “Under what conditions can we use a smaller set of moment conditions without incurring any loss of asymptotic efficiency? In other words, under what conditions are some moment conditions redundant in the sense that utilizing them does not improve efficiency?” These questions were first dealt with by Im, Ahn, Schmidt and Wooldridge (1999) who considered panel data models with strictly exogenous explanatory variables. They argued that, for example, with ten strictly exogenous time-varying variables and six time periods, the moment conditions available for the random effects (RE) model is 360 and this reduces to 300 moment conditions for the fixed effects (FE) model. GMM utilizing all these moment conditions leads to an efficient estimator. However, these moment conditions exceed what the simple RE and FE

⁶See the growing literature on weak instruments by Nelson and Startz (1990), Bekker (1994), Angrist and Kreuger (1995), Bound, Jaeger and Baker (1995) and Staiger and Stock (1997) to mention a few.

estimators use. Im et al. (1999) provide the assumptions under which this efficient GMM estimator reduces to the simpler FE or RE estimator. In other words, Im et al. (1999) show the redundancy of the moment conditions that these simple estimators do not use. Ahn and Schmidt (1999a) provide a more systematic method by which redundant instruments can be found and generalize this result to models with time-varying individual effects. However, both papers deal only with strictly exogenous regressors. In a related paper, Ahn and Schmidt (1999b) consider the cases of strictly and weakly exogenous regressors. They show that the GMM estimator takes the form of an instrumental variables estimator if the assumption of no conditional heteroskedasticity (NCH) holds. Under this assumption, the efficiency of standard estimators can often be established showing that the moment conditions not utilized by these estimators are redundant. However, Ahn and Schmidt (1999b) conclude that the NCH assumption necessarily fails if the full set of moment conditions for the dynamic panel data model are used. In this case, there is clearly a need to find modified versions of GMM, with reduced set of moment conditions that lead to estimates with reasonable finite sample properties.

Crepon, Kramarz and Trognon (1997) argue that for the dynamic panel data model, when one considers a set of orthogonal conditions, the parameters can be divided into parameters of interest (like δ) and nuisance parameters (like the second order terms in the autoregressive error component model). They show that the elimination of such nuisance parameters using their empirical counterparts does not entail an efficiency loss when only the parameters of interest are estimated. In fact, Sevestre and Trognon in chapter 6 of Matyas and Sevestre (1996) argue that if one is only interested in δ , then one can reduce the number of orthogonality restrictions without loss in efficiency as far as δ is concerned. However, the estimates of the other nuisance parameters are not generally as efficient as those obtained from the full set of orthogonality conditions.

The Alonso-Borrego and Arellano (1999) paper is also motivated by the finite sample bias in panel data instrumental variable estimators when the instruments are weak. The dynamic panel model generates many overidentifying restrictions even for moderate values of T. Also, the number of instruments increases with T, but the quality of these instruments is often poor because they tend to be only weakly correlated with first differenced endogenous variables that appear in the equation. Limited information maximum likelihood (LIML) is strongly preferred to 2SLS if the number of instruments gets large as the sample size tends to infinity. Hillier (1990) showed that the alternative normalization rules adopted by LIML and 2SLS are at the root of their different sampling behavior. Hillier (1990) also showed that a symmetrically normalized 2SLS estimator has properties similar to those of LIML. Following Hillier (1990), Alonso-Borrego and Arellano (1999) derive a symmetrically normalized GMM (SNM) and compare it with ordinary GMM and LIML analogues by means of simulations. Monte Carlo and empirical results show that GMM can exhibit large

biases when the instruments are poor, while LIML and SNM remain unbiased. However, LIML and SNM always had a larger interquartile range than GMM. For $T=4$, $N=100$, $\sigma_\mu^2 = 0.2$ and $\sigma_\nu^2 = 1$, the bias for $\delta = 0.5$ was 6.9% for GMM, 1.7% for SNM and 1.7% for LIML. This bias increases to 17.8% for GMM, 3.7% for SNM and 4.1% for LIML for $\delta = 0.8$.

Alvarez and Arellano (1997) studied the asymptotic properties of FE, one-step GMM and non-robust LIML for a first-order autoregressive model when both N and T tend to infinity with $(N/T) \rightarrow c$ for $0 \leq c < 2$. For $T < N$, GMM bias is always smaller than FE and LIML bias is smaller than the other two. In fixed T framework, GMM and LIML are asymptotically equivalent, but as T increases, LIML has a smaller asymptotic bias than GMM. These results provide some theoretical support for LIML over GMM.⁷

Wansbeek and Knaap (1997) consider a simple dynamic panel data model with a time trend and heterogeneous coefficients on the lagged dependent variable and the time trend, i.e.,

$$y_{it} = \delta_i y_{i,t-1} + \xi_i t + \mu_i + \nu_{it} \quad (47)$$

This model results from Islam's (1995) version of Solow's model on growth convergence among countries. Wansbeek and Knaap (1997) show that double differencing gets rid of the individual country effects (μ_i) on the first round of differencing and the heterogeneous coefficient on the time trend (ξ_i) on the second round of differencing. Modified OLS, IV and GMM methods are adapted to this model and LIML is suggested as a viable alternative to GMM to guard against the small sample bias of GMM. Macroeconomic data are subject to measurement error and Wansbeek and Knaap (1997) show how these estimators can be modified to account for measurement error that is white noise. For example, GMM is modified so that it discards the orthogonality conditions that rely on the absence of measurement error.

Jimenez-Martin (1998) performs Monte Carlo experiments to study the performance of the Holtz-Eakin (1988) test for the presence of individual heterogeneity effects in dynamic small T unbalanced panel data models. The design of the experiment includes both endogenous and time-invariant regressors in addition to the lagged dependent variable. The test behaves correctly for a moderate autoregressive coefficient. However, when this coefficient approaches unity, the presence of an additional regressor sharply affects the power and the size of the test. The power of this test is higher when the variance of the specific effects increases (they are easier to detect), when the sample size increases, when the data set is balanced (for a given number of cross-section units) and when the regressors are strictly exogenous.

⁷ An alternative one-step method that achieves the same asymptotic efficiency as robust GMM or LIML estimators is the maximum empirical likelihood estimation method, see Imbens (1997). This maximizes a multinomial pseudo-likelihood function subject to the orthogonality restrictions. These are invariant to normalization because they are maximum likelihood estimators.

6.1 Heterogeneous Dynamic Panel Data Models

The fundamental assumption underlying pooled homogeneous parameter models has been called into question. Robertson and Symons (1992) warned about the bias from pooled estimators when the estimated model is dynamic and homogeneous when in fact the true model is static and heterogeneous. Pesaran and Smith (1995) argued in favor of dynamic heterogeneous models when N and T are large. In this case, pooled homogeneous estimators are inconsistent whereas an average estimator of heterogeneous parameters can lead to consistent estimates as N and T tend to infinity. Maddala, Srivastava and Li (1994) argued that shrinkage estimators are superior to either heterogeneous or homogeneous parameter estimates especially for prediction purposes. In fact, Maddala, Trost, Li and Joutz (1997) considered the problem of estimating short run and long run elasticities of residential demand for electricity and natural gas for each of 49 states over the period 1970-1990.⁸ They conclude that individual heterogeneous state estimates were hard to interpret and had the wrong signs. Pooled data regressions were not valid because the hypothesis of homogeneity of the coefficients was rejected. They recommend shrinkage estimators if one is interested in obtaining elasticity estimates for each state since these give more reliable results.

Baltagi and Griffin (1997) compare short run and long run estimates as well as forecasts for pooled homogeneous, individual heterogeneous and shrinkage estimators of a dynamic demand model for gasoline across 18 OECD countries over the period 1960-1990. Based on one, five and ten year forecasts and plausibility of the short run and long run elasticity estimates, the results are in favor of pooling. Similar results were obtained for a dynamic model for cigarette demand across 46 states over the period 1963-1992, see Baltagi, Griffin and Xiong (1999).

In chapter 8 of Matyas and Sevestre (1996), Pesaran, Smith and Im investigated the small sample properties of various estimators of the long-run coefficients for a dynamic heterogeneous panel data model using Monte Carlo experiments. Their findings indicate that the mean group estimator performs reasonably well for large T . However, when T is small, the mean group estimator could be seriously biased, particularly when N is large relative to T . Pesaran and Zhao (1999) examine the effectiveness of alternative bias-correction procedures in reducing the small sample bias of these estimators using Monte Carlo experiments. An interesting finding is that when the coefficient of the lagged dependent variable is greater than or equal to 0.8, none of the bias correction procedures seem to work.

Hsiao, Pesaran and Tahmiscioglu (1999) suggest a Bayesian approach for estimating the mean parameters of a dynamic heterogeneous panel data model. The coefficients are assumed to be normally distributed across

⁸Maddala, et. al. (1997) also provide a unified treatment of classical, Bayes and empirical Bayes procedures for estimating this model.

cross-sectional units and the Bayes estimator is implemented using Markov Chain Monte Carlo methods. Hsiao, et al. (1999) argue that Bayesian methods can be a viable alternative in the estimation of mean coefficients in dynamic panel data models even when the initial observations are treated as fixed constants. They establish the asymptotic equivalence of this Bayes estimator and the mean group estimator proposed by Pesaran and Smith (1995). The asymptotics are carried out for both N and $T \rightarrow \infty$ with $\sqrt{N}/T \rightarrow 0$. Monte Carlo experiments show that this Bayes estimator has better sampling properties than other estimators for both small and moderate size T . Hsiao, et al. (1999) also caution against the use of the mean group estimator unless T is sufficiently large relative to N . The bias in the mean coefficient of the lagged dependent variable appears to be serious when T is small and the true value of this coefficient is larger than 0.6. Hsiao, et al. (1999) apply their methods to estimate the q-investment model using a panel of 273 US firms over the period 1972-1993.

7 Conclusion

This survey gives a brief overview of some of the main results in the econometrics nonstationary panels as well as recent developments in dynamic panels. There has been an immense amount of research in this area recently with the demand for empirical studies exceeding the supply of econometric theory developed for these models. As this survey indicates, several issues have been resolved but a lot remains to be done.

References

- Ahn, S.C., and P. Schmidt. 1995. "Efficient Estimation of Models for Dynamic Panel Data." *Journal of Econometrics* 68: 5-27.
- Ahn, S.C., and P. Schmidt. 1997. "Efficient Estimation of Dynamic Panel Data Models: Alternative Assumptions and Simplified Estimation." *Journal of Econometrics* 76: 309-321.
- Ahn, S.C., and P. Schmidt. 1999a. "Modified Generalized Instrumental Variables Estimation of Panel Data Models with Strictly Exogenous Instrumental Variables." In *Analysis of Panel Data and Limited Dependent Variable Models*, edited by C. Hsiao, K. Lahiri, L.F. Lee and M.H. Pesaran, Cambridge: Cambridge University Press, pp. 171-198.
- Ahn, S.C., and P. Schmidt. 1999b. "Estimation of Linear Panel Data Models Using GMM." In *Generalized Method of Moments Estimation*, Cambridge: Cambridge University Press, pp. 211-247.

- Alonso-Borrego, C., and M. Arellano. 1999. "Symmetrically Normalized Instrumental Variable Estimation Using Panel Data." *Journal of Business and Economic Statistics* 17: 36-49.
- Alvarez, J., and M. Arellano. 1997. "The Time Series and Cross-section Asymptotics of Dynamic Panel Data Estimators." Working paper, CEMFI, Madrid.
- Andersen, T.G., and R.E. Sørensen. 1996. "GMM Estimation of a Stochastic Volatility Model: A Monte Carlo Study." *Journal of Business and Economic Statistics* 14: 328-352.
- Anderson, T.W., and C. Hsiao. 1982. "Formulation and Estimation of Dynamic Models Using Panel Data." *Journal of Econometrics* 18: 47-82.
- Andersson, J., and J. Lyhagen. 1999. "A Long Memory Panel Unit Root Test: PPP Revisited." Working paper, Economics and Finance, No. 303, Stockholm School of Economics, Sweden.
- Angrist, J.D., and A.B. Krueger. 1995. "Split Sample Instrumental Variable Estimates of Return to Schooling." *Journal of Business and Economic Statistics* 13: 225-235.
- Arellano, M. 1989. "A Note on the Anderson-Hsiao Estimator for Panel Data." *Economics Letters* 31: 337-341.
- Arellano, M., and S. Bond. 1991. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and An Application to Employment Equations." *Review of Economic Studies* 58: 277-297.
- Arellano, M., and O. Bover. 1995. "Another Look at the Instrumental Variables Estimation of Error-Component Models." *Journal of Econometrics* 68: 29-51.
- Baltagi, B.H. 1995. *Econometric Analysis of Panel Data*. Chichester: Wiley.
- Baltagi, B.H., and J.M. Griffin. 1995. "A Dynamic Demand Model for Liquor: The Case for Pooling." *Review of Economics and Statistics* 77: 45-553.
- Baltagi, B.H., and J.M. Griffin. 1997. "Pooled Estimators v.s. Their Heterogeneous Counterparts in the Context of Dynamic Demand for Gasoline." *Journal of Econometrics* 77:303-327.
- Baltagi, B.H., J.M. Griffin, and W. Xiong. 1999. "To Pool or Not to Pool: Homogeneous Versus Heterogeneous Estimators Applied to Cigarette Demand." *Review of Economics and Statistics*, forthcoming.
- Banerjee, A. 1999. "Panel Data Units and Cointegration: An Overview." *Oxford Bulletin of Economics and Statistics* 61: 607-629.

- Bekker, P.A. 1994. "Alternative Approximations to the Distributions of Instrumental Variables Estimators." *Econometrica* 62: 657-682.
- Bernard, A., and C. Jones. 1996. "Productivity Across Industries and Countries: Time Series Theory and Evidence." *Review of Economics and Statistics* 78: 135-146.
- Bhargava, A., L. Franzini, and W. Narendranathan. 1982. "Serial Correlation and Fixed Effects Models." *Review of Economic Studies* 49: 533-549.
- Blundell, R.W., and S. Bond. 1998. "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models." *Journal of Econometrics* 87: 115-143.
- Boumahdi, R., and A. Thomas. 1991. "Testing for Unit Roots Using Panel Data." *Economics Letters* 37: 77-79.
- Bound, J., D.A. Jaeger, and R.M. Baker. 1995. "Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variables is Weak." *Journal of the American Statistical Association* 90: 443-450.
- Breitung, J. 1999. "The Local Power of Some Unit Root Tests for Panel Data." Discussion paper, Humboldt University, Berlin.
- Breitung, J., and W. Meyer. 1994. "Testing for Unit Roots in Panel Data: Are Wages on Different Bargaining Levels Cointegrated?" *Applied Economics* 26: 353-361.
- Canzoneri, M.B., E.E. Cumby, and B. Diba. 1999. "Relative Labor Productivity and the Real Exchange Rate in the Long Run: Evidence for a Panel of OECD Countries." *Journal of International Economics* 47: 245-266.
- Choi, I. 1999a. "Unit Root Tests for Panel Data." Manuscript, Kookmin University, Korea.
- Choi, I. 1999b. "Asymptotic Analysis of a Nonstationary Error Component Model." Manuscript, Kookmin University, Korea.
- Choi, I. 1999c. "Instrumental Variables Estimation of a Nearly Nonstationary Error Component Model." Manuscript, Kookmin University, Korea.
- Coakley, J., and A.M. Fuertes. 1997. "New Panel Unit Root Tests of PPP." *Economics Letters* 57: 17-22.
- Coe, D., and E. Helpman. 1995. "International R&D Spillovers." *European Economic Review* 39: 859-887.

- Conley, T.G. 1999. "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics* 92: 1-45.
- Crepon, B., F. Kramarz, and A. Trognon. 1997. "Parameters of Interest, Nuisance Parameters and Orthogonality Conditions: An Application to Autoregressive Error Components Models." *Journal of Econometrics* 82: 135-156.
- Culver, S. E., and D.H. Papell. 1997. "Is There a Unit Root in the Inflation Rate? Evidence from Sequential Break and Panel Data Model." *Journal of Applied Econometrics* 35: 155-160.
- Driscoll, J.C., and A.C. Kraay. 1998. "Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data." *Review of Economics and Statistics* 80: 549-560.
- Evans, P., and G. Karras. 1996. "Convergence Revisited." *Journal of Monetary Economics* 37: 249-265.
- Entorf, H. 1997. "Random Walks with Drifts: Nonsense Regression and Spurious Fixed-Effect Estimation." *Journal of Econometrics* 80: 287-296.
- Frankel, J.A., and A.K. Rose. 1996. "A Panel Project on Purchasing Power Parity: Mean Reversion Within and Between Countries." *Journal of International Economics* 40: 209-224.
- Funk, M. 1998. "Trade and International R&D Spillovers Among OECD Countries." Working paper, Department of Economics, St. Louis University, St. Louis.
- Gerdtham, U.G., and M. Löthgren. 1998. "On Stationarity and Cointegration of International Health Expenditure and GDP." Working paper, Economics and Finance, No. 232, Stockholm School of Economics, Sweden.
- Groen, J.J.J. 1999. "The Monetary Exchange Rate Model as A Long-run Phenomenon." *Journal of International Economics*, forthcoming.
- Groen, J.J.J., and F. Kleibergen. 1999. "Likelihood-Based Cointegration Analysis in Panels of Vector Error Correction Models." Discussion paper 99-055/4, Tinbergen Institute, The Netherlands.
- Hadri, K. 1999. "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root in Panel Data with Serially Correlated Errors." Manuscript, Department of Economics and Accounting, University of Liverpool, United Kingdom.
- Hahn, J. 1997. "Efficient Estimation of Panel Data Models With Sequential Moment Restrictions." *Journal of Econometrics* 79: 1-21.

- Hahn, J. 1999. "How Informative is the Initial Condition in the Dynamic Panel Model with Fixed Effects?" *Journal of Econometrics* 93: 309-326.
- Hall, S. Lazarova and G. Urga. 1999. "A Principal Components Analysis of Common Stochastic Trends in Heterogeneous Panel Data: Some Monte Carlo Evidence." *Oxford Bulletin of Economics and Statistics* 61: 749-767.
- Harris, D., and B. Inder. 1994. "A Test of the Null Hypothesis of Cointegration," In *Nonstationary Time Series Analysis and Cointegration*," edited by Hargreaves, Colin P. New York: Oxford University Press.
- Harris, R. D. F., and E. Tzavalis. 1999. "Inference for Unit Roots in Dynamic Panels Where the Time Dimension is Fixed." *Journal of Econometrics* 91: 201-226.
- Hausman, J.A., and W.E. Taylor. 1981. "Panel Data and Unobservable Individual Effects." *Econometrica* 49: 1377-1398.
- Hillier, G.H. 1990. "On the Normalization of Structural Equations: Properties of Direction Estimators." *Econometrica* 58: 1181-1194.
- Holtz-Eakin, D. 1988. "Testing for Individual Effects in Autoregressive Models." *Journal of Econometrics* 39: 297-307.
- Hsiao, C. 1986. *Analysis of Panel Data*. Cambridge: Cambridge University Press.
- Hsiao, C., M.H. Pesaran, and K. Tahmiscioglu. 1999. "Bayes Estimation of Short-run Coefficients in Dynamic Panel Data Models." In *Analysis of Panel Data and Limited Dependent Variable Models*, edited by C. Hsiao, K. Lahiri, L.F. Lee and M.H. Pesaran. Cambridge: Cambridge University Press, pp. 268-296.
- Im, K.S., S.C. Ahn, P. Schmidt, and J.M. Wooldridge. 1999. "Efficient Estimation of Panel Data Models with Strictly Exogenous Explanatory Variables." *Journal of Econometrics* 93: 177-201.
- Im, K.S., M.H. Pesaran, and Y. Shin. 1997. "Testing for Unit Roots in Heterogeneous Panels." Manuscript, Department of Applied Economics, University of Cambridge, United Kingdom.
- Imbens, G. 1997. "One-Step Estimators for Over-identified Generalized Method of Moments Models." *Review of Economic Studies* 64: 359-383.

- Islam, N. 1995. "Growth Emprics: A Panel Data Approach." *Quarterly Journal of Economics* 110: 1127-1170.
- Jimenez-Martin, S. 1998. "On the Testing of Heterogeneity Effects in Dynamic Unbalanced Panel Data Models." *Economics Letters* 58: 157-163.
- Johansen, S. 1995. *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford: Oxford University Press.
- Jorion, P., and R. Sweeney. 1996. "Mean Reversion in Real Exchange Rates: Evidence and Implications for Forecasting." *Journal of International Money and Finance* 15: 535-550.
- Judson, R.A., and A.L. Owen. 1999. "Estimating Dynamic Panel Data Models: A Guide for Macroeconomists." *Economics Letters* 65: 9-15.
- Kao, C. 1999. "Spurious Regression and Residual-Based Tests for Cointegration in Panel Data." *Journal of Econometrics* 90: 1-44.
- Kao, C., and M.H. Chiang. 1999. "On the Estimation and Inference of a Cointegrated Regression in Panel Data." Working paper, Center for Policy Research, Syracuse University, New York.
- Kao, C., M.H. Chiang, and B. Chen. 1999. "International R&D Spillovers: An Application of Estimation and Inference in Panel Cointegration." *Oxford Bulletin of Economics and Statistics* 61: 691-709.
- Karlsson, S., and M. Löthgren. 1999. "On the Power and Interpretation of Panel Unit Root Tests." Working paper, Economics and Finance, No. 299, Stockholm School of Economics, Sweden.
- Kauppi, H. 1999. "On Estimation and Inference of Cointegrated Panel Regressions with Near Integrated Regressors." Department of Economics, University of Helsinki, Finland.
- Keane, M.P., and D.E. Runkle. 1992. "On the Estimation of Panel-data Models with Serial Correlation When Instruments are Not Strictly Exogenous." *Journal of Business and Economics Statistics* 10: 1-9.
- Kiviet, J.F. 1995. "On Bias, Inconsistency and Efficiency of Some Estimators in Dynamic Panel Data Models." *Journal of Econometrics* 68: 53-78.
- Kiviet, J.F. 1999. "Expectations of Expansions for Estimators in a Dynamic Panel Data Model: Some Results for Weakly Exogenous Regressors." In *Analysis of Panel Data and Limited Dependent Variable*

- Models*, edited by C. Hsiao, K. Lahiri, L.F. Lee and M.H. Pesaran. Cambridge: Cambridge University Press, pp. 199-225.
- Larsson, R., J. Lyhagen, and M. Löthgren. 1998. "Likelihood-Based Cointegration Tests In Heterogeneous Panels." Working paper, Economics and Finance, No. 250, Stockholm School of Economics, Sweden.
- Lee, K., M.H. Pesaran, and R. Smith. 1997. "Growth and Convergence in a Multi-Country Empirical Stochastic Solow Model." *Journal of Applied Econometrics* 12: 357-392.
- Levin, A., and C.F. Lin. 1992. "Unit Root Test in Panel Data: Asymptotic and Finite Sample Properties." Discussion paper #92-93, University of California at San Diego.
- Lothian, J.R. 1996. "Multi-Country Evidence on the Behavior of Purchasing Power Parity Under the Current Float." *Journal of International Money and Finance* 16: 19-35.
- MacDonald, R. 1996. "Panel Unit Root Tests and Real Exchange Rates" *Economics Letters* 50: 7-11.
- Maddala, G.S. 1999. "On the Use of Panel Data Methods with Cross Country Data." *Annales d'Economie et de Statistique*, forthcoming.
- Maddala, G.S., V.K. Srivastava, and H. Li. 1994. "Shrinkage Estimators for the Estimation of Short-run and Long-run Parameters From Panel Data Models." Working paper, Ohio State University, Ohio.
- Maddala, G.S., R.P. Trost, H. Li, and F. Joutz. 1997. "Estimation of Short-run and Long-run Elasticities of Energy Demand from Panel Data Using Shrinkage Estimators." *Journal of Business and Economic Statistics* 15: 90-100.
- Maddala, G.S., and S. Wu. 1999. "A Comparative Study of Unit Root Tests with Panel Data and A New Simple Test." *Oxford Bulletin of Economics and Statistics* 61: 631-652.
- Maddala, G.S., S. Wu and P. Liu. 1999. "Do Panel Data Rescue Purchasing Power Parity (PPP) Theory?" In *Panel Data Econometrics: Future Directions*, edited by J. Krishnakumar and E. Ronchetti. Elsevier Science, forthcoming.
- Mátyás, L., and P. Sevestre. 1996. eds., *The Econometrics of Panel Data: A Handbook of Theory and Applications*. Dordrecht: Kluwer Academic Publishers.
- McCoskey, S., and C. Kao. 1998. "A Residual-Based Test of the Null of Cointegration in Panel Data." *Econometric Reviews* 17: 57-84.

- McCoskey, S., and C. Kao. 1999a. "Testing the Stability of a Production Function with Urbanization as a Shift Factor: An Application of Non-Stationary Panel Data Techniques." *Oxford Bulletin of Economics and Statistics* 61: 671-690.
- McCoskey, S., and C. Kao. 1999b. "Comparing Panel Data Cointegration Tests with an Application of the Twin Deficits Problems." Working paper, Center for Policy Research, Syracuse University, New York.
- McCoskey, S., and T. Selden. 1998. "Health Care Expenditures and GDP: Panel Data Unit Root Test Results." *Journal of Health Economics* 17: 369-376.
- Metcalfe, G.E. 1996. "Specification Testing in Panel Data with Instrumental Variables." *Journal of Econometrics* 71: 291-307.
- Moon, H.R., and P.C.B. Phillips. 1999. "Maximum Likelihood Estimation in Panels with Incidental Trends." *Oxford Bulletin of Economics and Statistics* 61: 711-747.
- Nelson, C. and R. Startz. 1990. "The Distribution of the Instrumental Variables Estimator and Its t-ratio When the Instrument is A Poor One." *Journal of Business* 63: S125-S140.
- Nerlove, M. 1999. "Properties of Alternative Estimators of Dynamic Panel Models: An Empirical Analysis of Cross-country Data for the Study of Economic Growth." In *Analysis of Panel Data and Limited Dependent Variable Models*, edited by C. Hsiao, K. Lahiri, L.F. Lee and M.H. Pesaran. Cambridge: Cambridge University Press, pp. 136-170.
- Nickell, S. 1981. "Biases in Dynamic Models with Fixed Effects." *Econometrica* 49: 1417-1426.
- O'Connell, P.G.J. 1998. "The Overvaluation of Purchasing Power Parity." *Journal of International Economics* 44: 1-19.
- Oh, K.Y. 1996. "Purchasing Power Parity and Unit Roots Tests Using Panel Data." *Journal of International Money and Finance* 15: 405-418.
- Papell, D. 1997. "Searching for Stationarity: Purchasing Power Parity Under the Current Float." *Journal of International Economics* 43: 313-332.
- Pedroni, P. 1996. "Fully Modified OLS for Heterogeneous Cointegrated Panels and the Case of Purchasing Power Parity." Working paper, Economics Department, Indiana University.

- Pedroni, P. 1997. "Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP Hypothesis." Working paper, Economics Department, Indiana University.
- Pedroni, P. 1999. "Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors." *Oxford Bulletin of Economics and Statistics* 61: 653-678.
- Pesaran, M.H., and R. Smith. 1995. "Estimating Long-run Relationships From Dynamic Heterogeneous Panels." *Journal of Econometrics* 68: 79-113.
- Pesaran, M.H., Y. Shin, and R. Smith. 1999. "Pooled Mean Group Estimation of Dynamic Heterogeneous Panels." *Journal of the American Statistical Association* 94: 621-634.
- Pesaran, M.H., and Z. Zhao. 1999. "Bias Reduction in Estimating Long-run Relationships From Dynamic Heterogeneous Panels." In *Analysis of Panels and Limited Dependent Variable Models*, edited by C. Hsiao, K. Lahiri, L.F. Lee and M.H. Persaran. Cambridge: Cambridge University Press, pp. 297-322.
- Phillips, P.C.B., and B.E. Hansen. 1990. "Statistical Inference in Instrumental Variables Regression with I(1) Processes." *Review of Economic Studies* 57: 99-125.
- Phillips, P.C.B., and H. Moon. 1999a. "Linear Regression Limit Theory for Nonstationary Panel Data." *Econometrica* 67: 1057-1111.
- Phillips, P.C.B., and H. Moon. 1999b. "Nonstationary Panel Data Analysis: An Overview of Some Recent Developments." *Econometric Reviews*, forthcoming.
- Phillips, P.C.B., and S. Ouliaris. 1990. "Asymptotic Properties of Residual Based Tests for Cointegration." *Econometrica* 58: 165-193.
- Quah, D. 1994. "Exploiting Cross Section Variation for Unit Root Inference in Dynamic Data." *Economics Letters* 44: 9-19.
- Quah, D. 1996. "Empirics for Economic Growth and Convergence." *European Economic Review* 40: 1353-1375.
- Robertson, D., and J. Symons. 1992. "Some Strange Properties of Panel Data Estimators." *Journal of Applied Econometrics* 7: 175-189.
- Saikkonen, P. 1991. "Asymptotically Efficient Estimation of Cointegrating Regressions." *Econometric Theory* 58: 1-21.

- Sala-i-Martin, X. 1996. "The Classical Approach to Convergence Analysis." *Economic Journal* 106: 1019-1036.
- Schmidt, P., S.C. Ahn, and D. Wyhowski. 1992. "Comment." *Journal of Business and Economic Statistics* 10: 10-14.
- Shin, Y. 1994. "A Residual Based Test of the Null of Cointegration Against the Alternative of No Cointegration." *Econometric Theory* 10: 91-115.
- Staiger, D., and J.H. Stock. 1997. "Instrumental Variables Regression With Weak Instruments." *Econometrica* 65: 557-586.
- Stock, J. 1993. "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems." *Econometrica* 61: 783-820.
- Stock, J. and M. Watson. 1993. "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems." *Econometrica* 61: 783-820.
- Tauchen, G. 1986. "Statistical Properties of Generalized Method of Moments Estimators of Structural Parameters Obtained From Financial Market Data." *Journal of Business and Economic Statistics* 4: 397-416.
- Wansbeek, T., and P. Bekker. 1996. "On IV, GMM and ML In A Dynamic Panel Data Model." *Economics Letters* 51: 145-152.
- Wansbeek, T., and T. Knaap. 1997. "Econometric Aspects of Estimating Solow's Growth Model Using Panel Data." *Annales d'Economie et de Statistique*, forthcoming.
- Wooldridge, J.M. 1997. "Multiplicative Panel Data Models Without the Strict Exogeneity Assumption." *Econometric Theory* 13: 667-678.
- Wu, S., and Y. Yin. 1999. "Tests for Cointegration in Heterogeneous Panel: A Monte Carlo Study." Working paper, Department of Economics, State University of New York at Buffalo, New York.
- Wu, Y. 1996. "Are Real Exchange Rates Nonstationary? Evidence from a Panel Data Set." *Journal of Money, Credit and Banking* 28: 54-63.
- Ziliak, J.P. 1997. "Efficient Estimation with Panel Data When Instruments are Predetermined: An Empirical Comparison of Moment-condition Estimators." *Journal of Business and Economic Statistics* 15: 419-431.