6-1-1999

Nonstationary Panel Data Analysis: An Overview of Some Recent Developments

Peter C.B. Phillips

Hyungsik Roger Moon

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 1221

NONSTATIONARY PANEL DATA ANALYSIS: AN OVERVIEW OF SOME RECENT DEVELOPMENTS

Peter C. B. Phillips and Hyungsik R. Moon

June 1999

Nonstationary Panel Data Analysis: An Overview of Some Recent Developments*

Peter C.B. Phillips
Cowles Foundation for Research in Economics
Yale University
and
Hyungsik R. Moon
Department of Economics
University of California, Santa Barbara

May, 1999

Abstract

This paper overviews some recent developments in panel data asymptotics, concentrating on the nonstationary panel case and gives a new result for models with individual effects. Underlying recent theory are asymptotics for multi-indexed processes in which both indexes may pass to infinity. We review some of the new limit theory that has been developed, show how it can be applied and give a new interpretation of individual effects in nonstationary panel data. Fundamental to the interpretation of much of the asymptotics is the concept of a panel regression coefficient which measures the long run average relation across a section of the panel. This concept is analogous to the statistical interpretation of the coefficient in a classical regression relation. A variety of nonstationary panel data models are discussed and the paper reviews the asymptotic properties of estimators in these various models. Some recent developments in panel unit root tests and stationary dynamic panel regression models are also reviewed.

1 Introduction

Since the early 1960's, economists have found panel data to be useful in studying a wide range of economic issues that involve individual economic agents over time. A variety of important and useful panel data sets have been constructed and are now widely available in electronic form. Some of these panel data sets, like the Penn-World tables, cover different individuals, industries, and countries over long time periods and have been useful in assessing and comparing growth characteristics, like real per capita GDP growth. One of the distinguishing features of these data sets is that they sometimes have an appreciable time series dimension (T) as well as a large cross section dimension (n). In some cases, the orders of magnitude of the cross section and time series dimensions are similar.

*We thank Essie Maasoumi, an anonymous referee, and Seung C. Ahn for helpful comments. Parts of this paper were presented at the 8^{th} Panel Data Conference, Göteborg, Sweden, June 1998. Phillips thanks the NSF for research support under Grant Nos. SBR 94–22922 & SBR97-30295 and Moon thanks the Cowles Foundation for financial support. The paper was typed by the authors in SW2.5.

These large n, large T panels have different characteristics and implications for theoretical and empirical analysis from the large n, small T panel data sets which have been the traditional object of study in panel data analysis.[1] For example, large n, large T panel regression models call for the use of large n, T asymptotics rather than just large n asymptotics. When T is large, there is an obvious need also to consider serial correlation patterns in the panel more generally, including both short memory and persistent components. In some panel data sets like the Penn-World Table, the time series components have strongly evident nonstationarity, a feature which received virtually no attention in traditional panel regression analysis. In order to properly analyze large n and large T panel data with such characteristics, it will generally be inadequate to appeal to conventional methods of analysis which are based on large n, small T data configurations.

Since the beginning of the 1990's, there has been a burgeoning of theoretical and applied research on the use of large n and large T panels allowing for nonstationarity in the data over time. Without attempting a general purpose review, this paper overviews some of the theoretical developments that have taken place in this type of panel data analysis, concentrating on recent advances in the econometric theory of panel regression. Our main focus of attention will be on some new asymptotics that have been developed in our own recent work but the paper also attempts to put this work in the broader context of the rapidly emerging literature on nonstationary panel data.

The organization of the paper is as follows. Section 2 discusses some concepts of multi-index asymptotics that were introduced in Phillips and Moon (1999). Depending on the asymptotic behavior of the two sample size indexes, n and T, a taxonomy of potential limit theories is provided and the relationships among them are briefly discussed. Section 3 gives a new interpretation of individual effects in nonstationary panel data and suggests a natural data generating process for nonstationary panel data with individual effects. Section 4 defines the concept of a long-run average relation in an analogous way to that of a regression coefficient in a classical linear regression, as in Phillips and Moon (1999). Section 5 reviews various linear regression models with nonstationary panel data and limit theories of various estimators of the long-run average parameter. This section extends the panel spurious regression framework studied in Phillips and Moon (1999) by allowing for individual effects and gives some new asymptotic results for the expanded model. Section 6 briefly reviews conventional dynamic panel regression models, some recent developments on panel unit root tests and the estimation of the localizing parameter in near integrated panels. Section 7 cites some recent and ongoing empirical work with nonstationary panels. Section 8 concludes and mathematical derivations are given in the Appendix.

2 Multi-Index Asymptotics

In regressions with large n, large T panels most of the interesting test statistics and estimators inevitably depend on both n and T. In consequence, a limit theory for such tests and estimators generally needs to allow both indexes to pass to infinity. Conventional limit theorems, by contrast, rely on the passage to infinity of a single index and are therefore not directly applicable in a panel context where there are twin indexes of sample size. This section reviews some concepts of multi-index asymptotics that were introduced in Phillips and Moon (1999) that are useful in the development of panel limit theory.

[1] Chamberlain(1984), Hsiao(1986), Matyas and Sevestre(1992), and Baltagi(1995) review much of the past research on conventional large n but small T panel data.

A typical multi-indexed process of the type that occurs in much panel data analysis has the following linear form

$$X_{n,T} = \frac{1}{k_n} \sum_{i=1}^{n} Y_{i,T},$$

where $Y_{i,T}$ are independent random vectors across i and usually a typical $Y_{i,T}$ component is a standardized sum of time series component of panel data. An example of a typical double indexed process is the following simple panel regression model with individual effects,

$$y_{i,t} = \alpha_i + \beta x_{i,t} + u_{i,t}.$$

In this model, attention is often given to the following pooled OLS estimator (sometimes called the within estimator):

$$\hat{\beta} = \beta + \frac{\sum_{i=1}^{n} \sum_{t=1}^{T} \left(u_{i,t} - \frac{1}{T} \sum_{s=1}^{T} u_{i,s} \right) \left(x_{i,t} - \frac{1}{T} \sum_{s=1}^{T} x_{i,s} \right)}{\sum_{i=1}^{n} \sum_{t=1}^{T} \left(x_{i,t} - \frac{1}{T} \sum_{s=1}^{T} x_{i,s} \right)^2}. \tag{2.1}$$

In developing a large sample theory for $\hat{\beta}$, the primary objects of interest involve the limits of quantities of the form

$$X_{n,T} = \frac{1}{n} \sum_{i=1}^{n} Y_{i,T},$$

where, for the numerator of (2.1) , we have

$$Y_{i,T} = \begin{cases} \frac{1}{T} \sum_{t=1}^{T} \left(x_{i,t} - \frac{1}{T} \sum_{s=1}^{T} x_{i,s} \right) \left(u_{i,t} - \frac{1}{T} \sum_{s=1}^{T} u_{i,t} \right) & \text{stationary case} \\ \frac{1}{T^2} \sum_{t=1}^{T} \left(x_{i,t} - \frac{1}{T} \sum_{s=1}^{T} x_{i,s} \right) \left(u_{i,t} - \frac{1}{T} \sum_{s=1}^{T} u_{i,t} \right) & \text{nonstationary case} \end{cases}$$

and, for the denominator of (2.1) , we have

$$Y_{i,T} = \begin{cases} \frac{1}{T} \sum_{t=1}^{T} \left(x_{i,t} - \frac{1}{T} \sum_{s=1}^{T} x_{i,t} \right)^2 & \text{stationary case} \\ \frac{1}{T^2} \sum_{t=1}^{T} \left(x_{i,t} - \frac{1}{T} \sum_{t=1}^{T} x_{i,t} \right)^2 & \text{nonstationary case} \end{cases}$$

In general, limits for double indexed processes like $X_{n,T}$ depend on the treatment of the two indexes, n and T, which tend to infinity together. Several approaches are possible, depending on the passage to infinity of the two indexes. These are reviewed below.

(a) Sequential Limits A sequential approach is to fix one index, say n, and allow the other, say T, to pass to infinity, giving an intermediate limit. Then, by letting n pass to infinity subsequently, a sequential limit theory is obtained. We write sequential limits of this type as $(T, n \to \infty)_{\text{seq}}$. In practice, when a double indexed process is of the typical form, *i.e.*,

$$X_{n,T} = \frac{1}{k_n} \sum_{i=1}^{n} Y_{i,T},$$

and the limit of $Y_{i,T}$ is Y_i as $T \to \infty$, we derive the sequential limit of $X_{n,T}$ as follows. By passing $T \to \infty$ for fixed n , an intermediate limit $X_n = \frac{1}{k_n} \sum_{i=1}^{n} Y_i$ is found. Then, by letting $n \to \infty$ and by applying an appropriate limit theory to the standardized sum

$X_n = \frac{1}{k_n}\sum_{i=1}^{n} Y_i$, the final sequential limit X is obtained. Usually, when $k_n = n$, a law of large numbers can be applied, and when $k_n = \sqrt{n}$, a central limit theorem can be applied. In many applications, sequential limits are easy to derive and helpful in extracting quick asymptotics. However, sometimes sequential limits can give misleading asymptotic results, and at the end of this section a simple example is provided that illustrates the type of problem that can arise with sequential limits.

(b) Diagonal Path Limits A second approach is to allow the two indexes, n and T, to pass to infinity along a specific diagonal path in the two dimensional array. This path can be determined by a monotonically increasing functional relation of the type $T = T(n)$ which applies as the index $n \to \infty$. This approach also simplifies the asymptotic theory by replacing $X_{n,T}$ with the single indexed process $X_{n,T(n)}$. Quah (1994) and Levin and Lin (1993) used this approach in finding the limits of panel unit root test statistics. One drawback of diagonal path limit theory is that the assumed expansion path $(T(n), n \to \infty)$ may not provide an appropriate approximation for a given (T, n) situation. Moreover, the limit theory that is obtained by this approach can depend on the specific functional relation $T = T(n)$ that is used in the asymptotic development. Again, we refer to the example at the end of this section for an illustration of what can happen.

(c) Joint Limits A joint limit theory allows both indexes, n and T, to pass to infinity simultaneously without placing specific diagonal path restrictions on the divergence, although it may still be necessary to exercise some control over the relative rate of expansion of the two indexes in order to get definitive results. Diagonal path limit theory turns out to be a special case of joint limit theory. In general, a joint limit will give a more robust result than either a sequential limit or diagonal path limit, but will also be substantially more difficult to derive and will usually apply only under stronger conditions, such as existence of higher moments, that will allow for uniformity in the convergence arguments. More importantly, it is not generally true that a sequential limit is equal to a joint limit. In the case of real number sequences, there are many such examples in real analysis (e.g., Apostol, 1974, p. 200). Here we give a simple example for a double sequence of random variables to illustrate what can happen.

Example Define the array of random variables

$$Z_{i,t} = \begin{cases} N(0,i) & \text{if } i \geq t \geq 1 \\ N(0,1) & \text{if } i < t, \end{cases}$$

where $Z_{i,t}$ is independent across i and over t. Let $Y_{i,T} = \frac{1}{\sqrt{T}}\sum_{t=1}^{T} Z_{i,t}$ and $X_{n,T} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_{i,T}$. Then, it is easy to verify that $X_{n,T}$ converges in distribution sequentially as $(T, n \to \infty)_{\text{seq}}$ to $N(0,1)$. Now let $n = T^r$. Then, $X_{n,T}$ $(= X_{T^r,T})$ has different limit distributions as the values of r change. In particular, as $T \to \infty$

$$X_{T^r,T} \Rightarrow \begin{cases} N(0,1) & \text{if } r < \frac{1}{2} \\ N(0,\frac{4}{3}) & \text{if } r = \frac{1}{2} \\ \text{does not converge} & \text{if } r > \frac{1}{2}. \end{cases}$$

So, sequential limits and diagonal path limits can give quite different results in this case. There is no general joint limit theory here for $X_{n,T}$ because $X_{n,T}$ diverges if n increases too fast (i.e., $n/\sqrt{T} \to \infty$). Indeed, a different normalization from $\sqrt{n}$ is required in this part of the array to obtain a well defined limit.

A fundamental question to ask is which are the cases where sequential limits will be equivalent to joint limits. Some intuition can be gleaned from known results for a double indexed real number sequence. In that case, if first stage convergence in the sequential limit holds uniformly in the other index, then the sequential limit will be a joint limit. That is, if $X_{n,T}$ converges to X_n uniformly in n as $T \to \infty$, then the sequential limit of $X_{n,T}$ is the same as the joint limit of $X_{n,T}$.

Phillips and Moon (1999) gave a generalization of this uniform convergence condition for random variable sequences that is applicable to multi-indexed asymptotics. That paper discussed the two cases of convergence in probability and convergence in distribution. The conditions given in the paper are relatively easy to verify and hold under what are fairly conventional regularity conditions. This approach to an asymptotic theory enables us to establish the joint limit of a double indexed sequence rather easily: we first derive the sequential limit and then verify the sufficient conditions that ensure the joint limit theory applies. The multi-index asymptotic theory in Phillips and Moon (1999) is applied to joint limits in which $T, n \to \infty$, and $\frac{T}{n} \to \infty$, i.e, to situations where the time series sample is large relative to the cross section sample. However, the general approach given there is also applicable to situations in which $\frac{T}{n} \to 0$, although different limit results will generally obtain in that case.

3 Individual Effects in Nonstationary Panel Data

Modeling, interpreting and dealing with individual effects is a crucial element in much panel data analysis and the same is true of the nonstationary case. In a simple dynamic panel regression model such as

$$z_{i,t} = \alpha_i + \beta z_{i,t-1} + u_{i,t}, \tag{3.2}$$

where the α_i are time invariant individual effects, if $|\beta| < 1$, the time series components of $z_{i,t}$ are stationary. In this case, depending on the underlying econometric application, conventional panel analysis offers different interpretations to the α_i, which sometimes appear as incidental parameters and sometimes as random components. When $\beta = 1$, the time series components of $z_{i,t}$ are nonstationary and this section offers an interpretation of the individual effects in terms of individual specific deterministic trends.

To motivate, we start with the simple dynamic model (3.2) with $\beta = 1$. Recursive substitution leads to

$$\begin{aligned} z_{i,t} &= \alpha_i + z_{i,t-1} + u_{i,t} \\ &= \alpha_i t + \sum_{s=1}^{t} u_{i,s} + z_{i,0} \\ &= \alpha_i t + z_{i,t}^0, \text{ say,} \end{aligned} \tag{3.3}$$

where $z_{i,t}^0 = z_{i,t-1}^0 + u_{i,t}$ so that $z_{i,t}^0$ is a pure unit root process. The reformulation in (3.3) reveals that nonstationary panel data with individual effects are composed of two components: (i) stochastic trends represented by $z_{i,t}^0$ whose time series components are pure unit root processes; and (ii) individually different (sometimes randomly different, depending on the assumptions concerning α_i) deterministic trends $\alpha_i t$, which are the realizations of the individual effects. This suggests that a natural interpretation of individual effects in nonstationary time series is as individually specific deterministic trends. Such formulations seem particularly useful in modelling aggregate macroeconomic time series like GDP per capita across countries which may have some individually specific growth characteristics while at the same time all having stochastic trends or autoregressive roots near unity.

Extending model (3.3) to the vector case is straightforward. Simply let $Z_{i,t}$ be an $m-$vector panel series and $A_{i,0}$ and $A_{i,1}$ be $m-$vector coefficients. Then, a natural data generating process for such panels with individual effects and allowance for nonstationarity would be the components model[2]

$$\begin{aligned} Z_{i,t} &= A_{i,0} + A_{i,1}t + Z_{i,t}^0, \\ Z_{i,t}^0 &= Z_{i,t-1}^0 + U_{i,t}. \end{aligned} \tag{3.4}$$

Models of this type may be useful in modelling several aggregate or financial series simultaneously over time and across countries. In such cases, we may also wish to allow for the elements of $A_{i,1}$ to be restricted in some way, perhaps by a functional dependence on another parameter, while still allowing for variation across i.

4 Long-Run Average Relationships

If (Y, X) is bivariate normally distributed as $N(0, \Sigma)$ with

$$\Sigma = \begin{bmatrix} \Sigma_{yy} & \Sigma_{yx} \\ \Sigma_{yx} & \Sigma_{xx} \end{bmatrix}$$

then the regression coefficient of Y on X is defined as the ratio $\beta = \Sigma_{yx}\Sigma_{xx}^{-1}$. Similarly, in the classical linear regression model

$$Y_t = \beta X_t + U_t, \tag{4.5}$$

where $EX_t = EU_t = 0$, and X_t and U_t are uncorrelated, the regression coefficient β satisfies the moment condition between Y_t and X_t given by

$$\beta = \left(EY_t X_t'\right)\left(EX_t X_t'\right)^{-1} = \Sigma_{yx}\Sigma_{xx}^{-1}. \tag{4.6}$$

It turns out that this type of classical regression coefficient can be extended to regression models with nonstationary time series variables. Suppose the dependent variable Y_t and independent variable X_t are unit root nonstationary and satisfy

$$\begin{pmatrix} Y_t \\ X_t \end{pmatrix} = \begin{pmatrix} Y_{t-1} \\ X_{t-1} \end{pmatrix} + \begin{pmatrix} U_{y,t} \\ U_{x,t} \end{pmatrix}$$

[2] Usually, the individual intercept terms $A_{i,0}$ can be absorbed into the intial condition of $Z_{i,0}^0$. Even if the $A_{i,0}$ can be identified by specifying a DGP for $Z_{i,0}^0$, usually the $A_{i,0}$ are not consistently estimable using time series components. See Phillips and Lee(1996) for some further discussion of this issue.

with stationary errors $U_t = \left(U'_{y,t}, U'_{x,t}\right)'$. Let the long-run variance matrix of U_t be given by

$$\Omega = \lim_T E\left(\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} U_t\right)\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} U'_t\right)\right) = \sum_{k=-\infty}^{\infty} E\left(U_0 U'_k\right),$$

and partitioned as

$$\Omega = \begin{pmatrix} \Omega_{yy} & \Omega_{xy} \\ \Omega_{yx} & \Omega_{xx} \end{pmatrix}.$$

Within this framework, it is possible to define a classical long run regression coefficient between Y and X for the long run covariance matrix Ω that is analogous to the coefficient β in (4.6). In particular, we can define

$$\beta = \lim_T E\left(\frac{Y_T}{\sqrt{T}}\frac{X'_T}{\sqrt{T}}\right)\left[\lim_T E\left(\frac{X_T}{\sqrt{T}}\frac{X'_T}{\sqrt{T}}\right)\right]^{-1} = \text{lrcov}\left(Y_T, X_T,\right)\left[\text{lrvar}\left(X_T\right)\right]^{-1} = \Omega_{yx}\Omega_{xx}^{-1} \tag{4.7}$$

and in this case β is interpreted as a coefficient that defines a long-run relation between two nonstationary variables Y_t and X_t.

When Ω has deficient rank, it now a standard result in the nonstationary time series literature (*e.g.*, Engle and Granger, 1987 and Phillips, 1986) that β is a cointegrating coefficient in the sense that the particular linear combination $Y_t - \beta X_t$ is stationary. A more interesting and remarkable feature of β is that it measures a statistical long-run correlation between two nonstationary variables X_t and Y_t even in the absence of time series cointegration. To see this point more clearly, suppose that two nonstationary time series variables Y_i and X_t have the following relation,

$$\begin{aligned} Y_t &= F_t + W_t, \\ X_t &= F_t \end{aligned}$$

with

$$\begin{pmatrix} W_t \\ F_t \end{pmatrix} = \begin{pmatrix} W_{t-1} \\ F_{t-1} \end{pmatrix} + \begin{pmatrix} U_{w,t} \\ U_{f,t} \end{pmatrix},$$

and where $U_{w,s}$ is independent of $U_{f,t}$ for all t and s and has non-zero long-run variance. In this example, F_t may correspond to a nonstationary common factor variable of Y_t and X_t, and W_t to a nonstationary idiosyncratic factor variable. Since W_t is nonstationary over time, it is apparent that there is no cointegrating relation between Y_t and X_t. However, since the two nonstationary variables Y_t and X_t share a common contributory nonstationary source in F_t, we may still expect to find evidence of long run correlation between X_t and Y_t, and this is what is measured by the regression coefficient β in (4.7).

Phillips and Moon (1999) extended this concept further to that of a *long-run average relation.* As we will discuss in the next section, this concept turns out to be useful in interpreting panel regressions with nonstationary data of the form

$$Y_{i,t} = \hat{\beta}_{n,T} X_{i,t} + \breve{U}_{i,t}, \tag{4.8}$$

or

$$Y_{i,t} = \hat{\alpha}_i + \hat{\beta}_{n,T} X_{i,t} + \breve{U}_{i,t},$$

where time series components of $\left(Y_{i,t}', X_{i,t}'\right)'$ are nonstationary[3]. To explain this notion, suppose panel observations of $Y_{i,t}$ and $X_{i,t}$ are available. In many applications it will be realistic to allow for some heterogeneity across individuals i in the population. This cross section heterogeneity can be characterized by heterogeneous long-run variance matrices Ω_i. The Ω_i can be taken to be randomly drawn from a population whose mean is $\Omega = E\Omega_i$. In this context, it is a natural extension of the usual classical concept of a regression coefficient to define the long-run average regression coefficient β as

$$\beta = E\left(\Omega_{y_i x_i}\right) E\left(\Omega_{x_i x_i}\right)^{-1} = \Omega_{yx}\Omega_{xx}^{-1}, \tag{4.9}$$

or

$$\beta = \left(\lim_{n,T} \frac{1}{n}\sum_{i=1}^{n} E\left(\frac{Y_{i,t}}{\sqrt{T}}\frac{X_{i,t}'}{\sqrt{T}}\right)\right)\left(\lim_{n,T} \frac{1}{n}\sum_{i=1}^{n} E\left(\frac{X_{i,t}}{\sqrt{T}}\frac{X_{i,t}'}{\sqrt{T}}\right)\right)^{-1} = \Omega_{yx}\Omega_{xx}^{-1},$$

which is simply the regression coefficient corresponding to the average long-run covariance matrix Ω.[4]

5 Linear Regression with Nonstationary Panel Data

Depending on the time series structure of the panel data, panel regressions of $Y_{i,t}$ on $X_{i,t}$ can be categorized into four cases: (i) panel spurious regression, where there is no time series cointegration; (ii) heterogeneous panel cointegration, where each individual has its own specific cointegrating relation; (iii) homogeneous panel cointegration, where individuals have the same cointegrating relation; and (iv) near-homogeneous panel cointegration, where individuals have slightly different cointegrating relations determined by the value of a localizing parameter. Assuming that the time series components of the panel $Z_{i,t} = \left(Y_{i,t}', X_{i,t}'\right)'$ is integrated, Phillips and Moon (1999) investigated these four models and developed panel asymptotics for regression coefficients and tests using both sequential and joint limit arguments. This section briefly reviews the main findings that relate to these nonstationary panel regression models. The final part of the section extends the asymptotics to the case where the DGP for panel vector integrated contains individual effects, that is panel data $Z_{i,t} = \left(Y_{i,t}', X_{i,t}'\right)'$ that are generated by (3.4).

(a) Panel Spurious Regression In the nonstationary time series literature, when the long-run covariance matrix Ω of the differences of a nonstationary vector $Z_{i,t} = (Y_t', X_t')'$ has full rank, an OLS regression of Y_t and X_t is said to be spurious (Granger and Newbold, 1974 and Phillips, 1986) and there is no cointegrating relation between them. Now consider the panel regression of two such component random vectors, $Y_{i,t}$ and $X_{i,t}$, for which there

[3] For expositional convenience, this paper considers the case of (4.8), where there is no individual effect in fitted regressions. For the more general case with the individual effect in the regression, refer to Phillips and Moon (1999).

[4] When the nonstationary panel data are generated by vector integrated process with individual effects

$$\begin{aligned} Z_{i,t} &= A_{i,0} + A_{i,1}t + Z_{i,t}^0, \\ Z_{i,t}^0 &= Z_{i,t-1}^0 + U_{i,t}, \end{aligned}$$

the average long-run relation between the elements of $Z_{i,t}$, say $Y_{i,t}$ and $X_{i,t}$, can be formulated by the regression coefficient of the long-run average covariance matrix of $Z_{i,t}^0$.

is no cointegrating relation among the elements of $Z_{i,t}$ for any i (*i.e.*, the conditional long-run covariance matrix of $\Delta Z_{i,t}$, Ω_i, is positive definite almost surely for all i). The regression has the form

$$Y_{i,t} = \hat{\beta} X_{i,t} + \hat{U}_{i,t} \tag{5.10}$$

or

$$Y_{i,t} = \hat{\alpha}_i + \hat{\beta} X_{i,t} + \hat{U}_{i,t}. \tag{5.11}$$

In this case, Phillips and Moon (1999) showed that under quite weak regularity conditions the pooled least squares estimator $\hat{\beta}$ is $\sqrt{n}-$consistent for the long-run average relation parameter β and has a limiting normal distribution. In other work, Moon and Phillips (1998a) showed that in (5.11) a limiting cross section regression with time averaged data is also $\sqrt{n}$-consistent for β and again has a limiting normal distribution. In some related work on nonstationary panel cointegration tests, Pedroni(1995) and Kao and McCoskey(1998) looked at model (5.11) under more restrictive conditions than Phillips and Moon (1999). In this section, for convenience, we consider the simple model (5.10) .

The idea behind the consistent estimation of the long run average coefficient in a spurious panel is simple and intuitive and can be explained as follows. In a time series spurious regression of Y_t on X_t the limit of the OLS estimator $\hat{\beta}_{sp}$ is a nondegenerate random variate that is a functional of Brownian motions as in (Phillips (1986))

$$\hat{\beta}_{sp} = \frac{1}{T^2}\sum_{t=1}^{T} Y_t X_t' \left(\frac{1}{T^2}\sum_{t=1}^{T} X_t X_t'\right)^{-1} \Rightarrow \int B_y B_x' \left(\int B_x B_x'\right)^{-1} \neq \beta = \Omega_{yx}\Omega_{xx}^{-1},$$

where $\left(B_y', B_x'\right)'$ is a vector Brownian motion with covariance matrix Ω and integrals here and elsewhere in the paper are taken over the interval $[0,1]$. In this case $\hat{\beta}_{sp}$ is not consistent for β. However, by a simple calculation

$$E\left(\int B_y B_x'\right) = \frac{1}{2}\Omega_{yx}, \quad E\left(\int B_x B_x'\right) = \frac{1}{2}\Omega_{xx},$$

so that

$$\beta = E\left(\int B_y B_x'\right)\left[E\left(\int B_x B_x'\right)\right]^{-1} = \Omega_{yx}\Omega_{xx}^{-1}.$$

The idea in Phillips and Moon (1999) is that independent cross section data in the panel adds information and this leads to a stronger overall signal than that of the pure time series case. More specifically, in the panel case, we have the estimator

$$\hat{\beta}_{n,T} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{T^2}\sum_{t=1}^{T} Y_{i,t}X_{i,t}'\right)\left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{T^2}\sum_{t=1}^{T} X_{i,t}X_{i,t}'\right)\right]^{-1},$$

which pools information across individuals i. If we fix n and let $T \to \infty$ first and then let $n \to \infty$, we have

$$\begin{aligned} \hat{\beta}_{n,T} &\Rightarrow \frac{1}{n}\sum_{i=1}^{n}\int B_{y_i}B_{x_i}' \left[\frac{1}{n}\sum_{i=1}^{n}\int B_{x_i}B_{x_i}'\right]^{-1} \text{ as } T \to \infty \text{ for fixed } n \\ &\to_p E\left(\int B_y B_x'\right)\left[E\left(\int B_x B_x'\right)\right]^{-1} = \Omega_{yx}\Omega_{xx}^{-1} = \beta \text{ as } n \to \infty. \end{aligned} \tag{5.12}$$

Similarly, the OLS estimator with time averaged data, which is called a limit cross section estimator,

$$\begin{aligned}\tilde{\beta}_{n,T} &= \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{T\sqrt{T}}\sum_{t=1}^{T}Y_{i,t}\right)\left(\frac{1}{T\sqrt{T}}\sum_{t=1}^{T}X_{i,t}'\right) \\ &\quad\times\left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{T\sqrt{T}}\sum_{t=1}^{T}X_{i,t}\right)\left(\frac{1}{T\sqrt{T}}\sum_{t=1}^{T}X_{i,t}'\right)\right]^{-1}\end{aligned}$$

has the following sequential limit as $(T, n \to \infty)_{\text{seq}}$

$$\begin{aligned}\tilde{\beta}_{n,T} &\Rightarrow \left[\frac{1}{n}\sum_{i=1}^{n}\int B_{y_i}\int B_{x_i}'\right]\left[\frac{1}{n}\sum_{i=1}^{n}\int B_{x_i}\int B_{x_i}'\right]^{-1} \quad \text{as } T\to\infty \text{ for fixed } n \\ &\to_p \left[E\int B_{y_i}\int B_{x_i}'\right]\left[E\int B_{x_i}\int B_{x_i}'\right]^{-1} = \Omega_{yx}\Omega_{xx}^{-1} = \beta \text{ as } n\to\infty. \quad (5.13)\end{aligned}$$

So both $\hat{\beta}_{n,T}$ and $\tilde{\beta}_{n,T}$ are consistent for β.

(b) Heterogeneous Panel Cointegration Regression When the long run covariance matrices Ω_i have deficient rank, there will exist time series cointegrating relations among the elements of $Z_{i,t}$. These relations will, in general, be heterogeneous and the model will then be a heterogenous panel cointegrating regression of the form

$$Y_{i,t} = \beta_i X_{i,t} + E_{i,t}, \quad (5.14)$$

where $\beta_i = \Omega_{y_i x_i}\Omega_{x_i x_i}^{-1}$ are randomly differing individual cointegration coefficients and the $E_{i,t}$ are stationary. Phillips and Moon (1999) showed that in model (5.14) the pooled least squares estimator $\hat{\beta}$ is $\sqrt{n}-$consistent for β and has a normal limit distribution.

Pesaran and Smith(1995) studied limiting cross section regressions with time averaged data and established consistency with restrictive assumptions on model (5.14). An important difference between these studies is that Pesaran and Smith(1995) use an average of the cointegration coefficients, given by $\beta_{PS} = E(\beta_i)$, which is generally different from the long run average regression coefficient β because[5]

$$E\left(\Omega_{y_i x_i}\Omega_{x_i x_i}^{-1}\right) \neq E\left(\Omega_{y_i x_i}\right)\left[E\left(\Omega_{x_i x_i}\right)\right]^{-1}.$$

In order to define β_{PS}, there needs to exist cointegrating time series relations, whereas β is defined irrespective of the existence of individual cointegrating relations and relies only on the long run average variance matrix of the panel. In this aspect, the use of the long run average regression coefficient β seems to be a more robust concept than the average coefficient β_{PS} in empirical works.

(c) Homogeneous Panel Cointegration Regression The third model assumes that the same cointegrating relation between $Y_{i,t}$ and $X_{i,t}$ applies for all i, viz.

$$Y_{i,t} = \beta X_{i,t} + E_{i,t}, \quad (5.15)$$

[5] Of course, there are some cases where two long-run average relations are equivalent. For example $\Omega_{x_i x_i} = \Omega_{xx}$ for all i.

where $\beta = \Omega_{yx}\Omega_{xx}^{-1}$ is a homogeneous cointegrating coefficient. This common cointegrating relation may be suggested by an underlying economic theory.

Among recent contributions to the nonstationary panel data literature, Phillips and Moon (1999), Pedroni(1996), and Kao and Chiang(1998) have all investigated limit theories for various estimators of β in model (5.15). Each of these studies has found that the pooled OLS estimator of β is $\sqrt{n}$- consistent or $\sqrt{n}T$ - consistent depending on whether or not there exists serial correlation in the time series component of $(E_{i,t}, \Delta X_{i,t})$. This happens because serial correlation in the time series component of $(E_{i,t}, \Delta X_{i,t})$ generates a second order bias arising from a one-sided long-run covariance between $E_{i,t}$ and $\Delta X_{i,t}$ (see Phillips and Durlauf, 1986, Park and Phillips, 1988, and Phillips and Moon, 1998), but in the panel case this bias is serious enough to alter the rate of convergence of the estimator. To address the bias problem in the time series case, some specialized estimation procedures have been suggested. Among these the fully modified (FM) method has attracted the most interest and is the most used in empirical research.

The FM method was originally suggested by Phillips and Hansen(1990) to eliminate endogeneity in the regressors (arising from correlation between $X_{i,t}$ and $E_{i,t}$ in model (5.15)) and serial correlation in the errors (i.e., serial correlation in $(E_{i,t}, \Delta X_{i,t})$), both of which generate the second order bias (see Phillips and Hansen, 1990, and Phillips, 1995, for a full exposition). In the nonstationary time series literature, it is well known that the FM method yields an optimal estimator for the cointegrating coefficient in Gaussian cointegration regression models (Phillips, 1991). To show how to apply the FM approach in a panel regression model, we consider the simple model (5.15), where the error process $(E_{i,t}, \Delta X_{i,t})$ is iid across i. Let Ω and Λ denote the long-run covariance matrix and one-sided long-run covariance matrix of $(E_{i,t}, \Delta X_{i,t})$, respectively, and partition these matrices as follows:

$$\begin{aligned}\Omega &= \begin{pmatrix} \Omega_{ee} & \Omega_{ex} \\ \Omega_{xe} & \Omega_{xx} \end{pmatrix} = \sum_{j=-\infty}^{\infty} \Gamma_j,\ \Lambda = \begin{pmatrix} \Lambda_{ee} & \Lambda_{ex} \\ \Lambda_{xe} & \Lambda_{xx} \end{pmatrix} = \sum_{j=0}^{\infty} \Gamma_j, \\ \Gamma_j &= E \begin{pmatrix} E_{i,t}E_{i,t+j} & E_{i,t}X_{i,t+j} \\ X_{i,t}E_{i,t+j} & X_{i,t}X_{i,t+j} \end{pmatrix}.\end{aligned}$$

To construct a panel FM estimator, we need to obtain consistent time series estimators $\hat{\Omega}$ and $\hat{\Lambda}$ of Ω and Λ. In our case, consistent estimates may be constructed using averages (over $i = 1, .., n$) of the usual consistent (as $T \to \infty$) nonparametric kernel estimates of the corresponding long-run quantities for each i. More specifically, let $\hat{\Gamma}_i(j) = \frac{1}{T}\sum_t F_{i,t}F'_{i,t+j}$, where the summation is over $1 \leq t,\ t+j \leq T$, and define the averaged kernel estimators

$$\begin{aligned}\hat{\Omega} &= \tfrac{1}{n}\sum_{i=1}^{n} \hat{\Omega}_i,\ \hat{\Omega}_i = \sum_{j=-T+1}^{T-1} w\left(\tfrac{j}{K}\right)\hat{\Gamma}_i(j), \\ \hat{\Lambda} &= \tfrac{1}{n}\sum_{i=1}^{n} \hat{\Lambda}_i,\ \hat{\Lambda}_i = \sum_{j=0}^{T-1} w\left(\tfrac{j}{K}\right)\hat{\Gamma}_i(j),\end{aligned}$$

where $w(x)$ is a lag kernel (for detailed conditions on suitable kernels, see Phillips and Moon, 1998). Then, define the modified dependent variables

$$Y_{i,t}^{+} = Y_{i,t} - \hat{\Omega}_{ex}\hat{\Omega}_{xx}^{-1}\Delta X_{i,t} \tag{5.16}$$

and serial dependence eliminator

$$\hat{\Lambda}_{ex}^{+} = \hat{\Lambda}_{ex} - \hat{\Omega}_{ex}\hat{\Omega}_{xx}^{-1}\hat{\Lambda}_{xx}. \tag{5.17}$$

Equation (5.16) gives the panel endogeneity correction and equation (5.17) gives the panel serial correlation correction. Using these corrections, a pooled FM (PFM) estimator can be defined as follows:

$$\hat{\beta}_{PFM} = \left(\sum_{i=1}^{n}\sum_{t=1}^{T} Y_{i,t}^{+}X_{i,t}' - nT\hat{\Lambda}_{ex}^{+}\right)\left(\sum_{i=1}^{n}\sum_{t=1}^{T} X_{i,t}X_{i,t}'\right)^{-1}.$$

Then, as shown in Phillips and Moon (1999), $\hat{\beta}_{PFM}$ is $\sqrt{n}T$ - consistent for β and has a normal limit distribution.

(d) Near-Homogenous Panel Cointegration: This model has the form

$$Y_{i,t} = \beta_i X_{i,t} + E_{i,t}, \tag{5.18}$$

where

$$\beta_i = \beta + \frac{\theta_i}{\sqrt{n}T}. \tag{5.19}$$

The model allows each individual to possess a slightly different cointegrating relation in a localizing sense. As shown in Phillips and Moon (1999), the PFM estimator for β is $\sqrt{n}T$-consistent and a normal limit distribution with an asymptotic bias that depends on the average noncentrality, $E(\theta_i)$, in (5.19).

5.1 An Extension to Models with Individual Effects

This section provides some new results that show how to extend the above ideas to models with individual effects in the DGP. Suppose that the $m-$vector of nonstationary panel data $Z_{i,t}$ is generated by an integrated process with individual specific deterministic trends as in

$$\begin{aligned} Z_{i,t} &= A_{i,0} + A_{i,1}t + Z_{i,t}^{0}, \\ Z_{i,t}^{0} &= Z_{i,t-1}^{0} + U_{i,t}, \end{aligned} \tag{5.20}$$

where the conditional long-run covariance matrix of $Z_{i,t}^{0}$ is positive definite almost surely. As we discussed above, the trends $A_{i,0} + A_{i,1}t$ reflect individual specific effects in the panel data $Z_{i,t}$. This section shows how to estimate the long-run average relation between two elements of $Z_{i,t} = \left(Y_{i,t}', X_{i,t}'\right)'$ in the presence of individual specific trends. Only the case of panel spurious regression is covered here, the panel cointegrating case being very similar[6]

Consistent estimation of the long-run average relation β is quite straightforward. First, we estimate the trend coefficients[7] $A_{i,0}$ and $A_{i,1}$, then detrend $Z_{i,t}$ by taking the residuals

[6] In the nonstationary time series literature, models like (5.20) have been widely studied. For example, Phillips(1989) investigated spurious regressions and Johansen(1991) studied cointegration regressions in this context.

[7] As mentioned earlier, it is impossible to consistently estimate the intercept coefficient $A_{i,0}$. But this fact does not affect consistent estimation of β.

from this regression. Next, the detrended data is pooled and used in least squares regression to estimate β. Following recent work in the time series literature (Phillips and Lee, 1996; Canjels and Watson, 1997), we consider two different detrending procedures based on OLS and GLS regression. The OLS detrended data denoted as $\tilde{Z}_{i,t}$ is constructed as

$$\begin{aligned}
\tilde{Z}_{i,t} &:= Z_{i,t} - \sum_{s=1}^{T} Z_{i,s} g_s' \left(\sum_{s=1}^{T} g_s g_s' \right)^{-1} g_t \\
&= Z_{i,t}^0 - \sum_{s=1}^{T} Z_{i,s}^0 g_s' \left(\sum_{t=1}^{T} g_s g_s' \right)^{-1} g_t \quad (t = 1, ..., T \ ; i = 1, ..., n),
\end{aligned} \tag{5.21}$$

where $g_t = (1, t)'$, so that $\tilde{Z}_{i,t}$ is the time series OLS residual of $Z_{i,t}$ on the linear trend $g_t = (1, t)'$. When the component time series $Z_{i,t}^0$ is a pure unit root process, time series differencing forms the basis of the GLS transformation, as it is not necessary to use any of the short memory serial correlation properties in the formation of an efficient detrending procedure (Phillips and Lee, 1996). Thus, we have

$$\Delta Z_{i,t} = A_{i,1} + \Delta Z_{i,t}^0 \quad \text{for } 2 \leq t \leq T.$$

and then the GLS estimator of $A_{i,1}$ is

$$\bar{A}_{i,1} = \frac{1}{T-1} \sum_{t=2}^{T} \Delta Z_{i,t} = \frac{1}{T-1} \left(Z_{i,T} - Z_{i,1} \right).$$

as in Schmidt and Phillips (1992). The intercept coefficient $A_{i,0}$ is estimated by taking an average of $Z_{i,t} - \bar{A}_{i,1}t$, *i.e.*,

$$\begin{aligned}
\bar{A}_{i,0} &= \frac{1}{T} \sum_{t=1}^{T} \left(Z_{i,t} - \bar{A}_{i,1} t \right) \\
&= \frac{1}{T} \sum_{t=1}^{T} \left(Z_{i,t} - \frac{t}{T-1} \left(Z_{i,T} - Z_{i,1} \right) \right).
\end{aligned}$$

The GLS detrended process, $\bar{Z}_{i,t}$, is then defined as[8]

$$\begin{aligned}
\bar{Z}_{i,t} &:= Z_{i,t} - \tfrac{t}{T-1} \left(Z_{i,T} - Z_{i,1} \right) - \tfrac{1}{T} \sum_{s=1}^{T} \left(Z_{i,s} - \tfrac{s}{T-1} \left(Z_{i,T} - Z_{i,1} \right) \right), \qquad 1 \leq t \leq T \\
&= Z_{i,t}^0 - \tfrac{t}{T-1} \left(Z_{i,T}^0 - Z_{i,1}^0 \right) - \tfrac{1}{T} \sum_{s=1}^{T} \left(Z_{i,s}^0 - \tfrac{s}{T-1} \left(Z_{i,T}^0 - Z_{i,1}^0 \right) \right).
\end{aligned}$$

Now, let $\tilde{Z}_{i,t} = \left(\tilde{Y}_{i,t}', \tilde{X}_{i,t}' \right)'$ and $\bar{Z}_{i,t} = \left(\bar{Y}_{i,t}', \bar{X}_{i,t}' \right)'$, where the partitions are conformable. Using this detrended data, the following two estimators of β can be constructed:

$$\tilde{\beta}_s = \left(\sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{Y}_{i,t} \tilde{X}_{i,t}' \right) \left(\sum_{i=1}^{n} \sum_{t=1}^{T} \tilde{X}_{i,t} \tilde{X}_{i,t}' \right)^{-1}$$

and

$$\bar{\beta}_s = \left(\sum_{i=1}^{n} \sum_{t=1}^{T} \bar{Y}_{i,t} \bar{X}_{i,t}' \right) \left(\sum_{i=1}^{n} \sum_{t=1}^{T} \bar{X}_{i,t} \bar{X}_{i,t}' \right)^{-1}.$$

The following result, whose proof is in the Appendix gives a sequential limit theory for these estimators.

[8] This particular detrending procedure has been used in much recent work on nonstationary time series – see Bhargava(1986), Stock(1991), and Schmidt and Phillips(1992).

Theorem 1 *In sequential limits as* $(T, n \to \infty)_{seq}$,

(a) $\tilde{\beta}_s$, $\bar{\beta}_s \to_p \beta$.

(b) $\sqrt{n}\left(\tilde{\beta}_s - \beta\right) \Rightarrow N\left(0, 225\left(\Omega_{xx}^{-1} \otimes I_{m_y}\right)\tilde{\Theta}_s\left(\Omega_{xx}^{-1} \otimes I_{m_y}\right)\right)$
$\sqrt{n}(\bar{\beta}_s - \beta) \Rightarrow N\left(0, 144\left(\Omega_{xx}^{-1} \otimes I_{m_y}\right)\bar{\Theta}_s\left(\Omega_{xx}^{-1} \otimes I_{m_y}\right)\right)$.

Theorem 1 shows that the pooled least squares estimators using the detrended data are $\sqrt{n}$- consistent for the long-run average parameter β and they both have normal asymptotic distributions. From the expressions for the variance matrices given in the appendix and writing

$$\Omega_{x_i x_i} = C_{x_i} C_{x_i}', \quad \Omega_{y_i y_i} = C_{y_i} C_{y_i}'$$

we deduce that

$$\begin{aligned}
& 144\bar{\Theta}_s - 225\tilde{\Theta}_s \\
= & \left(\frac{144}{720} - \frac{225 \times 11}{12600}\right) E\left(\Omega_{x_i x_i} \otimes \left(\Omega_{y_i y_i} - \beta\Omega_{x_i y_i} - \Omega_{y_i x_i}\beta' + \beta\Omega_{x_i x_i}\beta'\right)\right) \\
& + \left(\frac{144}{720} - \frac{225 \times 11}{12600}\right) E\left(\left(\Omega_{x_i y_i} - \Omega_{x_i x_i}\beta'\right) \otimes \left(\Omega_{y_i x_i} - \beta\Omega_{x_i x_i}\right) K_{m_y m_x}\right) \\
= & 0.0036E\left[\left(C_{x_i} \otimes \left(C_{y_i} - \beta C_{x_i}\right)\right)\left(I_{m^2} + K_m\right)\left(C_{x_i} \otimes \left(C_{y_i} - \beta C_{x_i}\right)\right)'\right] > 0.
\end{aligned}$$

It follows that OLS detrending leads to an asymptotically more efficient estimator of the long run average coefficient β in pooled regression than GLS detrending. This result, which may seem unexpected, is explained as follows. GLS detrending produces a more efficient estimator of the trend coefficient than OLS detrending in time series regression. However, the residuals after time series GLS detrending have more cross section variation than they do after OLS detrending and this produces great variation in the limit distribution of the pooled regression estimator of the long run average coefficient.

Theorem 1 can also be shown to hold when we take joint limits of the indexes provided $\frac{n}{T} \to 0$. We do not give a proof for the joint limit here, but refer readers who are interested in joint limit arguments to Phillips and Moon (1999).

5.2 Issues of Pooling and Cross Section Regression with Nonstationary Panel Data

In panel regression models, when the parameter of interest is the average effect between two variables and both n and T are large, different ways of aggregating panel data can lead to different results. For instance, in a random coefficient dynamic stationary panel regression model, the pooled estimator is inconsistent for the average of the random coefficient, while a limiting cross section regression gives a consistent estimate (*e.g.*, see Pesaran and Smith, 1995). However, in nonstationary panel regression models, both the pooled least squares estimator and the limiting cross section estimator are *sometimes* consistent for the long-run average relation parameter. An example is the spurious panel regression model in (5.10), and the results given in (5.12) and (5.13) confirm this.

6 Dynamic Panel Regression

6.1 Conventional Methods

One of important regression models in panel data analysis is a dynamic panel regression model with individual effects, and the simplest form is

$$z_{i,t} = \alpha_i + \beta z_{i,t-1} + u_{i,t}, \tag{6.22}$$

where α_i denote individual effects. In conventional dynamic panel analysis it is usually assumed that the size of time series component T is fixed while the cross-sectional dimension n goes to infinity. In this case, it is well known that widely used estimators such as the within estimator and the first difference estimator are inconsistent and generate asymptotic biases (*e.g.*, Nickell (1981) calculated the asymptotic bias of the within estimator for β and showed that it vanishes to zero as $T \to \infty$.) This problem is quite similar to the classical incidental parameter problem found by Neyman and Scott (1948) (Recently Lancaster (1998) surveys various incidental parameter problems that arise in econometric models.).

A simple way to overcome the incidental parameter problem is to treat the individual effect parameters α_i as random variable whose distribution belongs to a finite dimensional parameter family. Depending on different specification of the joint distribution of α_i and $z_{i,0}$ (initial observations), we may have different likelihood functions (*e.g.*, Anderson and Hsiao, 1981, Anderson and Hsiao, 1982, Bhargava and Sargan, 1983). In this case, according to the studies mentioned above, the maximum likelihood estimators are usually consistent, although there are some exceptions (*e.g.*, see Hsiao, 1986 Chapter 4 for exceptional cases.).

A drawback of the MLE approach is that it often requires quite strong assumptions on distributions of the individual effect α_i and the initial condition $z_{i,0}$. Also the computation of the MLE is usually not simple. An alternative method to overcome these problems of the MLE is to use an simple instrumental variable (IV). For example, Anderson and Hsiao (1981) used $z_{i,t-2}$ as an IV in the first difference regression. The IV estimation method is usually easy to implement, even though IV estimators in the early literature are not efficient in general. During the 80's, as a generalization of the IV estimation method, the Generalized Method of Moments (GMM) estimation method was developed (Hansen, 1982). The GMM estimation utilizes the information on the population moment conditions implied by economic theories or underlying data generating processes. Since than, recent dynamic panel studies have applied the GMM approach in estimating the dynamic panel regression model. From the assumptions imposed on the data generating process of the dynamic panel data, relevant population moment conditions are found. Based on these conditions, one constructs an efficient GMM estimator that is consistent and asymptotic normal under quite weak conditions. In this case, various GMM estimators are possible, depending on the different moment conditions. (*e.g.*, Ahn and Schmidt, 1995, Ahn and Schmidt, 1997, Arellano and Bond, 1991, Arellano and Bover, 1995, Chamberlain, 1992, and Hahn, 1997, Blundell and Bond, 1998).

On the other hand, the availability of large n and large T panel data sets initiates new studies of the dynamic panel data analysis for large n and large T panel data. In this case, the modeling of time series components in the panel data are important in analyzing asymptotic theories. In particular, depending on the value β, the time series in the panel are stationary or nonstationary and totally different limit theories are applied.

Assuming $|\beta| < 1$ and normality in the disturbances, Alvarez and Arellano (1998) and Hahn (1998) studied the asymptotic properties of various estimators for β in the dynamic panel regression model when both n and T are comparable. According to Alvarez and Arellano (1998), when T/n tends to a nonzero constant, the within, GMM, and LIML (Limited Information Maximum Likelihood) estimators have negative asymptotic biases of order T, n, and $(2n - T)$, respectively. Parts of these results are also obtained in Hahn (1998) independently. Hahn (1998) first showed that when T/n tends to a nonzero constant, both the MLE and IV estimators have asymptotic bias. In the case of the MLE he developed a device to fix the asymptotic bias and, using a convolution theorem, he also showed that the bias corrected MLE is asymptotically efficient.

6.2 Nonstationary Dynamic Panel and Panel Unit Root Tests

When β in (6.22) equals to one, the time series components in the panel are nonstationary, and in this case it requires to apply different approximation theories from the stationary case. An important model for nonstationary panel data where there has been active recent interest is the following dynamic system

$$\Delta z_{i,t} = \alpha_i + (\beta - 1) z_{i,t-1} + u_{i,t}, \tag{6.23}$$

or

$$\begin{aligned} z_{i,t} &= \alpha_{i,0} + \alpha_{i,1} t + z_{i,t}^0, \\ z_{i,t}^0 &= \beta z_{i,t-1}^0 + u_{i,t}, \end{aligned} \tag{6.24}$$

where $\beta = 1$, and so the time series components of the panel $z_{i,t}$ have unit roots for all individuals i. This type of model has been widely investigated especially in studies concerned with panel unit root testing.

Quah(1994) first suggested a simple panel unit root test statistic and indicated its usefulness in applications such as tests of growth convergence theories in macroeconomics. Using a simple panel unit root regression model

$$z_{i,t} = \beta z_{i,t-1} + u_{i,t},$$

where $\beta = 1$, he suggested a simple panel unit root test using the pooled OLS estimator. Assuming that the $u_{i,t}$ are iid across i and over t and the order of magnitude of the cross section and time series dimension is the same, *i.e.*, $n = T$, Quah showed that the panel unit root test statistic has a normal limit distribution. Levin and Lin(1993) extended Quah(1994)'s panel unit root test using an augmented version of the panel unit root model (6.23). In deriving large n and large T asymptotics, they allowed for a more general relationship between n and T, $T = T(n)$, and considered heterogenous error processes. Using a dynamic panel regression model of type (6.23), Im *et al.*(1997) developed panel unit root tests based on an average of individual LM tests assuming Gaussian errors. Considering a more general alternative hypothesis, Choi(1997) and Maddala and Wu(1997) independently suggested panel unit root test statistics based on various combinations of the p-values of unit root tests applied to each individual. Following an approach in the time series literature suggested by Kwiatowski *et al.*(1992), Hadri(1998) proposed a panel version of a residual based LM test for a null of trend stationary against the alternative of a unit root for panel data.

On the other hand, using model (6.24), where $\beta = \exp\left(\frac{c}{T}\right)$, Moon and Phillips (1998b) focused on estimating the localizing parameter c in β. The local parameter c characterizes the local behaviour of the unit root process. Information about this parameter is useful in the context of several different econometric procedures. A few examples are the analysis of the power properties of unit root tests (Phillips, 1987) and cointegration tests (Phillips, 1989, Johansen, 1991), the construction of confidence intervals for the long-run autoregressive coefficient (Stock, 1991), the development of efficient detrending methods (Phillips and Lee, 1996, Canjels and Watson, 1997), and the construction of point optimal invariant tests for a unit root (Elliot *et al.*, 1996) and cointegrating rank (Xiao and Phillips, 1999). However, the local to unity parameter c in $\beta = \exp\left(\frac{c}{T}\right)$ cannot be consistently estimated using time series data (although the reader is referred to Phillips *et al.*, 1998, for an alternate block local to unity model in which the consistent estimation of c is possible). The paper by Moon and Phillips (1998b) developed procedures for the estimation of the local parameter using panel data. They showed that when $c \leq 0$, it is possible to estimate c consistently with panel data and derived asymptotic properties of the estimators. As an application, they showed how to extract the deterministic trend efficiently using consistent estimates of c.

The panel unit root tests we have reviewed above are also closely related to residual based panel cointegration tests for a null of no cointegration. To test for time series cointegration, Phillips and Ouliaris (1990) proposed various test statistics for cointegration between two nonstationary time series, y_t and x_t, say, by applying unit root tests to the residual of the regression y_t on x_t. Recently, Pedroni (1995), McKoskey and Kao (1998), Kao (1999) applied this idea to test for cointegration in nonstationary panel data and investigated some properties of cointegration statistics in pooled time series panels for the null of no cointegration.

7 Empirical Applications

The econometric methods reviewed above, especially panel unit root tests, panel cointegration tests and the estimation of long-run average relations, have formed the basis of some recent empirical econometric studies with large n, large T panels. For example, using panel unit root tests, Bernard and Jones (1996) tested growth convergence theories, and MacDonald (1996), Oh (1996), Pedroni (1997), Wu (1996), and Wu (1997) tested various forms of purchasing power parity relations using both panel unit root tests and residual based panel cointegration tests. Coakley *et al.*(1996) developed an economic model where panels of savings and investments are cointegrated and tested the theory using residual based panel cointegration tests. On the same topic and allowing for savings and investment rates to be nonstationary but not requiring that there exist time series cointegration at the individual level, Moon and Phillips (1998a) argued that what Feldstein and Horioka (1980) and many subsequent authors have estimated in cross section regression of time averaged savings and investment rates is a long-run average relation between savings and investment rates. The existence of a long run average relation in such a context is justified by the arguments outlined in Section 4. In rerunning regressions of the Feldstein–Horioka type and using asymptotically valid econometric tests, Moon and Phillips (1998a) found evidence that continues to support the original conclusions of the Feldstein–Horioka study.

8 Concluding Remarks

This paper provides an outline of recent developments in the econometric theory of nonstationary panel data. The field offers some interesting new asymptotic theory for multi-indexed processes and introduces the new concept of a cross section average long run relation that relies both on the time series notion of the long run and the cross section notion of a statistical regression coefficient. These techniques and conceptual apparatus provide a basis for performing and interpreting econometric analyses of panel regressions with nonstationary data and with large n and T sample dimensions. Panel data facilitates the study of individual economic behaviour over time. When the individuals are nation states and the data are macroeconomic aggregates or financial asset prices whose time series behaviour is typically nonstationary, the scope for the empirical use of these methods seems to be substantial. And, in the future, the scope for these methods may be even more important as interest in inter-country comparisons of economic performance heightens and more extensive panel data sets become available.

9 Appendix

Proof of Theorem 1 Before we start the analysis, we assume that the initial condition $Z_{i,0}$ satisfies $Z_{i,0} = O_p(1)$ as $T \to \infty$. Let $D_T = diag\left(1, T^{-1}\right)$ and $g(r) = (1, r)'$. If $t = [Tr]$, then, as $T \to \infty$, $D_T g_t \to g(r)$ uniformly in $r \in [0,1]$. We also assume that as $T \to \infty$

$$\frac{Z_{i,t}^0}{\sqrt{T}} \Rightarrow \Omega_i^{1/2} W_i(r) := M_i(r),$$

where the $W_i(r)$ are standard vector Brownian motions. Phillips and Moon (1999) give sufficient conditions for this functional central limit theorem. Then, by the continuous mapping theorem, it follows that as $T \to \infty$

$$\begin{aligned} \frac{\tilde{Z}_{i,t}}{\sqrt{T}} &= \frac{Z_{i,t}^0}{\sqrt{T}} - \frac{1}{T}\sum_{t=1}^{T} \frac{Z_{i,t}^0}{\sqrt{T}} D_T g_t \left(\frac{1}{T}\sum_{t=1}^{T} D_T g_t g_t' D_T'\right)^{-1} D_T g_t \\ &\Rightarrow M_i(r) - \int M_i(r) g(r)' dr \left(\int g(r) g(r)'\right)^{-1} g(r) := \tilde{M}_i(r), \end{aligned} \tag{9.25}$$

and

$$\begin{aligned} \frac{\bar{Z}_{i,t}}{\sqrt{T}} &= \frac{Z_{i,t}^0}{\sqrt{T}} - \frac{t}{T-1}\frac{Z_{i,T}^0}{\sqrt{T}} - \frac{1}{T}\sum_{s=1}^{T}\left(\frac{Z_{i,s}^0}{\sqrt{T}} - \frac{s}{T-1}\frac{Z_{i,T}^0}{\sqrt{T}}\right) + o_p(1) \\ &\Rightarrow M_i(r) - rM_i(1) - \int \left(M_i(s) - sM_i(1)\right) ds := \bar{M}_i(r), \end{aligned} \tag{9.26}$$

where the $o_p(1)$ holds by the initial condition assumption. Denoting $\tilde{W}_i(r) = W_i(r) - \int W_i(r) g(r)' dr \left(\int g(r) g(r)'\right)^{-1} g(r)$ and $\bar{W}_i(r) = W_i(r) - rW_i(1) - \int \left(W_i(s) - sW_i(1)\right) ds$, we can write

$$\tilde{M}_i(r) = \Omega_i^{1/2} \tilde{W}_i(r) \text{ and } \bar{M}_i(r) = \Omega_i^{1/2} \bar{W}_i(r).$$

The limit distribution $\tilde{M}_i(r)$ of the standardized OLS detrended process is a randomly scaled detrended Brownian motion and the limit distribution $\bar{M}_i(r)$ of the standardized GLS detrended process is a randomly scaled demeaned Brownian bridge. We partition conformably $\tilde{M}_i(r)$ and $\bar{M}_i(r)$ into $\tilde{M}_i(r) = \left(\tilde{M}_{y_i}(r)', \tilde{M}_{x_i}(r)'\right)'$ and $\bar{M}_i(r) = \left(\bar{M}_{y_i}(r)', \bar{M}_{x_i}(r)'\right)'$.

By (9.25), (9.26), and the continuous mapping theorem, the pooled estimators $\tilde{\beta}$ and $\bar{\beta}$ constructed from detrended data have the following limit distributions, as $T \to \infty$ for any fixed n,

$$\tilde{\beta}_s \Rightarrow \left(\frac{1}{n}\sum_{i=1}^{n}\int \tilde{M}_{y_i}\tilde{M}'_{x_i}\right)\left(\frac{1}{n}\sum_{i=1}^{n}\int \tilde{M}_{x_i}\tilde{M}'_{x_i}\right)^{-1} \tag{9.27}$$

and

$$\bar{\beta}_s \Rightarrow \left(\frac{1}{n}\sum_{i=1}^{n}\int \bar{M}_{y_i}\bar{M}'_{x_i}\right)\left(\frac{1}{n}\sum_{i=1}^{n}\int \bar{M}_{x_i}\bar{M}'_{x_i}\right)^{-1}. \tag{9.28}$$

Write $h(s,r) = g(s)'\left(\int g(r)g(r)'dr\right)^{-1} g(r)$. Let $W(r)$ be a standard m-vector Brownian motion and $W^o(r)$ an m-vector Brownian Bridge. Let $\mathcal{F}$ be the sigma field generated by Ω_i, which is iid across i, and let $E_{\mathcal{F}}$ denote a conditional expection. Then, using the fact that $E\left(W(r)W'(s)\right) = (r \wedge s)\, I_m$ and $E\left(W^o(r)W^o(s)\right) = \left((r \wedge s) - rs\right) I_m$, we have

$$\begin{aligned} E\left(\int \tilde{M}_i\tilde{M}'_i\right) &= E\left[E_{\mathcal{F}}\left(\int M_i(r)M'_i(r)dr - \int\int M_i(s)h(s,r)M'_i(r)dsdr\right)\right] \\ &= \left(\int_0^1 rdr - \int_0^1\int_0^1 (r \wedge s)\, h(s,r)dsdr\right) E\left(\Omega_i\right) = \frac{1}{15}\Omega, \end{aligned}$$

and, by writing $M_i^o(r) = M_i(r) - rM_i(1)$, we have

$$\begin{aligned} E\left(\int \bar{M}_i\bar{M}'_i\right) &= E\left[E_{\mathcal{F}}\left(\int M_i^o(r)M_i^{o\prime}(r)dr - \int M_i^o(s)ds\int M_i^{o\prime}(r)ds\right)\right] \\ &= \left(\int_0^1 \left(r - r^2\right) dr - \int_0^1\int_0^1 \left((r \wedge s) - rs\right) drds\right) E\left(\Omega_i\right) = \frac{1}{12}\Omega. \end{aligned}$$

Thus, if we apply the Kolmogorov strong law to the numerators and the denominators of the matrix quotients in (9.27) and (9.28), we have as $n \to \infty$

$$\frac{1}{n}\sum_{i=1}^{n}\int \tilde{M}_{y_i}\tilde{M}'_{x_i} \overset{\text{a.s.}}{\to} \frac{1}{15}\Omega_{yx}, \qquad \frac{1}{n}\sum_{i=1}^{n}\int \tilde{M}_{x_i}\tilde{M}'_{x_i} \overset{\text{a.s.}}{\to} \frac{1}{15}\Omega_{xx} \tag{9.29}$$

and

$$\frac{1}{n}\sum_{i=1}^{n}\int \bar{M}_{y_i}\bar{M}'_{x_i} \overset{\text{a.s.}}{\to} \frac{1}{12}\Omega_{yx}, \qquad \frac{1}{n}\sum_{i=1}^{n}\int \bar{M}_{x_i}\bar{M}'_{x_i} \overset{\text{a.s.}}{\to} \frac{1}{12}\Omega_{xx}. \tag{9.30}$$

Thus, it is straightforward to see that in sequential limits as $(n, T \to \infty)_{seq}$

$$\tilde{\beta}_s, \bar{\beta}_s \overset{p}{\to} \beta.$$

To find the limit distributions of $\tilde{\beta}_s$ and $\bar{\beta}_s$, we center the estimators at β and rescale them by $\sqrt{n}$. By letting $T \to \infty$ for fixed n, we have

$$\sqrt{n}\left(\tilde{\beta}_s - \beta\right) \Rightarrow \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\int \tilde{M}_{y_i}\tilde{M}'_{x_i} - \beta\int \tilde{M}_{x_i}\tilde{M}'_{x_i}\right)\left(\frac{1}{n}\sum_{i=1}^{n}\int \tilde{M}_{x_i}\tilde{M}'_{x_i}\right)^{-1} \tag{9.31}$$

and

$$\sqrt{n}\left(\tilde{\beta}_s - \beta\right) \Rightarrow \tfrac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\int \bar{M}_{y_i} \bar{M}_{x_i}' - \beta \int \bar{M}_{x_i} \bar{M}_{x_i}'\right) \left(\tfrac{1}{n} \sum_{i=1}^{n} \int \bar{M}_{x_i} \bar{M}_{x_i}'\right)^{-1}. \tag{9.32}$$

It is not difficult to see that

$$E\left(\int \tilde{M}_{y_i} \tilde{M}_{x_i}' - \beta \int \tilde{M}_{x_i} \tilde{M}_{x_i}'\right) = \frac{1}{15} E\left(\Omega_{y_i x_i} - \Omega_{yx} \Omega_{xx}^{-1} \Omega_{x_i x_i}\right) = 0,$$

and

$$E\left(\int \bar{M}_{y_i} \bar{M}_{x_i}' - \beta \int \bar{M}_{x_i} \bar{M}_{x_i}'\right) = \frac{1}{12} E\left(\Omega_{y_i x_i} - \Omega_{yx} \Omega_{xx}^{-1} \Omega_{x_i x_i}\right) = 0.$$

After some lengthy calculations, we derive the variance matrices

$$\begin{aligned}
& E\left(\operatorname{vec}\left(\int \tilde{M}_{y_i} \tilde{M}_{x_i}' - \beta \int \tilde{M}_{x_i} \tilde{M}_{x_i}'\right) \operatorname{vec}\left(\int \tilde{M}_{y_i} \tilde{M}_{x_i}' - \beta \int \tilde{M}_{x_i} \tilde{M}_{x_i}'\right)'\right) \\
= & \frac{11}{12600} E\left(\Omega_{x_i x_i} \otimes \left(\Omega_{y_i y_i} - \beta \Omega_{x_i y_i} - \Omega_{y_i x_i} \beta' + \beta \Omega_{x_i x_i} \beta'\right)\right) \\
& + \frac{11}{12600} E\left(\left(\Omega_{x_i y_i} - \Omega_{x_i x_i} \beta'\right) \otimes \left(\Omega_{y_i x_i} - \beta \Omega_{x_i x_i}\right) K_{m_y m_x}\right) \\
& + \frac{1}{225} E\left(vec\left(\Omega_{y_i x_i} - \beta \Omega_{x_i x_i}\right)\left(vec\left(\Omega_{y_i x_i} - \beta \Omega_{x_i x_i}\right)\right)'\right) \\
= & \tilde{\Theta}_s, \text{ say,}
\end{aligned} \tag{9.33}$$

and

$$\begin{aligned}
& E\left(\operatorname{vec}\left(\int \bar{M}_{y_i} \bar{M}_{x_i}' - \beta \int \bar{M}_{x_i} \bar{M}_{x_i}'\right) \operatorname{vec}\left(\int \bar{M}_{y_i} \bar{M}_{x_i}' - \beta \int \bar{M}_{x_i} \bar{M}_{x_i}'\right)'\right) \\
= & \frac{1}{720} E\left(\Omega_{x_i x_i} \otimes \left(\Omega_{y_i y_i} - \beta \Omega_{x_i y_i} - \Omega_{y_i x_i} \beta' + \beta \Omega_{x_i x_i} \beta'\right)\right) \\
& + \frac{1}{720} E\left(\left(\Omega_{x_i y_i} - \Omega_{x_i x_i} \beta'\right) \otimes \left(\Omega_{y_i x_i} - \beta \Omega_{x_i x_i}\right) K_{m_y m_x}\right) \\
& + \frac{1}{144} E\left(vec\left(\Omega_{y_i x_i} - \beta \Omega_{x_i x_i}\right)\left(vec\left(\Omega_{y_i x_i} - \beta \Omega_{x_i x_i}\right)\right)'\right) \\
= & \bar{\Theta}_s, \text{ say.}
\end{aligned} \tag{9.34}$$

>From the multivariate Lindeberg–Levy theorem, we then obtain as $n \to \infty$

$$\tfrac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\int \tilde{M}_{y_i} \tilde{M}_{x_i}' - \beta \int \tilde{M}_{x_i} \tilde{M}_{x_i}'\right) \Rightarrow N(0, \tilde{\Theta}_s) \tag{9.35}$$

and

$$\tfrac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\int \bar{M}_{y_i} \bar{M}_{x_i}' - \beta \int \bar{M}_{x_i} \bar{M}_{x_i}'\right) \Rightarrow N(0, \bar{\Theta}_s). \tag{9.36}$$

Combining (9.35) and (9.36) with the limits (9.29) and (9.30) as $n \to \infty$, we have the desired limit distributions for $\sqrt{n}(\tilde{\beta}_s - \beta)$ and $\sqrt{n}(\bar{\beta}_s - \beta)$. ■

References

[1] Ahn, S. C. and P. Schmidt. (1995): Efficient Estimation of Models for Dynamic Panel Data, *Journal of Econometrics*, 68, 5–28.

[2] 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.

[3] Anderson, T. W. and C. Hsiao. (1981): Estimation of Dynamic Models with Error Compoments, *Journal of the American Statistical Association*, 76, 598–606.

[4] Anderson, T. W. and C. Hsiao. (1982): Formulation and Estimation of Dynamic Models Using Panel Data, *Journal of Econometrics*, 18, 47–82.

[5] Alvarez, J. and M. Arellano. (1998): The Time Series and Cross-Section Asymptotics of Dynamic Panel Data Estimators, Mimeo.

[6] Arellano, M. and S. Bond. (1991): Some Tests of Specification for Panel Data; Monte Carlo Evidence and an Application to Employment Equation, *Review of Economic Studies*, 58, 277–297.

[7] Arellano, M. and O. Bover. (1995): Another Look at the Instrumental Variables Estimation of Error-Component Models, *Journal of Econometrics*, 68, 29–51.

[8] Apostol, T. (1974): *Mathematical Analysis,* Reading: Addison-Wesley.

[9] Baltagi, B. (1995): *Econometric Analysis of Panel Data*, New York, Wiley.

[10] Bernard, A. and C. Jones (1996): Productivity Across Industries and Countries: Time Series Theory and Evidence, *Review of Economics and Statistics,* 78, 135–146.

[11] Bhargava, A. (1986): On the Theory of Testing for Unit Roots in Observed Time Series, *Review of Economic Studies,* 53, 369–384.

[12] Bhargava, A. and D. Sargan. (1983): Estimation Dynamic Random Effects Models from Panel Data Covering Short Time Periods, *Econometrica*, 51, 1635–1659.

[13] Blundell, R. and S. Bond. (1998): Initial Conditions and Moment Restrictions in Dynamic Panel Data Models, *Journal of Econometrics*, 87, 115-143.

[14] Canjels, E. and M. Watson (1997): Estimating Deterministic Trends in the Presence of Serially Correlated Errors, *Review of Economics and Statistics,* 79, 184–200.

[15] Cavanagh, C., G. Elliott, and J.H. Stock (1995): Inference in Models with Nearly Integrated Regressors, *Econometric Theory*, 11, 1131–1147.

[16] Chamberlain, G. (1984): Panel Data, in Z. Griliches and M. Intriligator (eds.), *Handbook of Econometrics*, Vol. 2, Amsterdam, North-Holland.

[17] Chamberlain, G. (1992): Comment: Sequential Moment Restrictions in Panel Data, *Journal of Business and Economics Statistics*, 10, 20–26.

[18] Choi, I. (1997): Unit Root Tests for Panel Data, Mimeo.

[19] Coakely, J., F. Kulasi, and R. Smith (1996): Current Account Solvency and the Feldstein-Horioka Puzzle, *Economic Journal*, 106, 620–627.

[20] Elliott, G., T.J. Rothenberg, and J.H. Stock (1996): Efficient Tests of an Autoregressive Unit Root, *Econometrica*, 64, 813–836.

[21] Feldstein, M. and C. Horioka (1980): Domestic Saving and International Capital Flows, *Economic Journal*, 90, 314–329.

[22] Granger, C. W. J. and P. Newbold (1974): Spurious Regressions in Econometrics, *Journal of Econometrics*, 2, 111–120.

[23] Hadri, K. (1998): Testing for Stationarity in Heterogenous Panel Data, Mimeo.

[24] Hahn, J. (1997): Efficient Estimation of Panel Data Models with Sequential Moment Restrictions, *Journal of Econometrics,* 79, 1–21.

[25] Hansen, L. (1982): Large Sample Properties of Generalized Method of Moments Estimators, *Econometrica*, 50, 1029–1054.

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

[27] Im, K., H. Pesaran, and Y. Shin (1997): Testing for Unit Roots in Heterogeneous Panels, Mimeo.

[28] Johansen, S. (1991): Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models, *Econometrica*, 59, 1551–1580.

[29] Kao, C. (1999): Spurious Regression and residual Based Tests for Cointegration in Panel Data, *Journal of Econometrics*, 90, 1–45.

[30] Kao, C. and M. Chiang (1998): On the Estimation and Inference of a Cointegrated Regression in Panel Data, Mimeo.

[31] Kwiatkowski, D, P. C. B. Phillips, P. Schmidt, and Y. Shin (1992): Testing the null hypothesis of stationarity against the alternative of a unit root, *Journal of Econometrics*, 54, 91–115.

[32] Langcaster, T. (1998): The Incidental Parameter Problem Since 1948, Mimeo, Brown University.

[33] Levin, A. and C. Lin (1993): Unit Root Tests in Panel Data: New Results, UC San Diego Working Paper.

[34] MacDonald, R. (1996): Panel Unit Root Tests and Real Exchange Rates, *Economics Letters*, 50, 7–11.

[35] Maddala, G.S. and S. Wu (1997): A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test, Mimeo.

[36] Matyas, L. and P. Sevestre (eds.) (1992): *The Econometrics of Panel Data*, Boston, MA, Kluwer Academic Publishers.

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

[38] Moon, H.R. and P.C.B. Phillips (1998a): A Reinterpretation of the Feldstein–Horioka Regressions from a Nonstationary Panel Viewpoint, Mimeo, Yale University.

[39] Moon, H.R. and P.C.B. Phillips (1998b): Estimation of Autoregressive Roots near Unity using Panel Data, Mimeo, Yale University.

[40] Neyman, J and E. L. Scott (1948): Consistent Estimates Based on Partially Consistent Observations, *Econometrica*, 16, 1–32.

[41] Nickell, S. (1981): Biases in Dynamic Models with Fixed Effects, *Econometrica*, 49, 1417–1426.

[42] Oh, K.Y. (1996): Purchasing Power Parity and the Unit Root Tests Using Panel Data, *Journal of International Money and Finance*, 15, 405–418.

[43] Park, J and P.C.B. Phillips (1998): Statistical Inference in Regressions with Integrated Processes: Part I, *Econometric Theory*, 4, 468–497.

[44] Pedroni, P. (1995): Panel Cointegration; Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP Hypothesis, Indiana University Working Papers in Economics, No. 95–013.

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

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

[47] Phillips, P. C. B. (1986): Understanding Spurious Regressions in Econometrics, *Journal of Econometrics*, 33, 311–340.

[48] Phillips, P.C.B. (1987): Time Series Regression with Unit Root, *Econometrica*, 55, 277–301.

[49] Phillips, P.C.B. (1989): Partially Identified Econometric Models, *Econometric Theory*, 5, 181–240.

[50] Phillips, P.C.B. (1991): Optimal Inference in Cointegrated System, *Econometrica*, 59, 283–306.

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

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

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

[54] Phillips, P.C.B. and C. Lee (1996): Efficiency Gains from Quasi-Differencing under Nonstationarity, in P. M. Robinson and M. Rosenblatt (eds.), *Athens Conference on Applied Probability and Time Series: volume II Time Series Analysis in memory of E. J. Hannan.* New York: Springer–Verlag.

[55] Phillips, P.C.B. and H.R. Moon (1999): Linear Regression Limit Theory for Nonstationary Panel Data, *Econometrica*, forthcoming.

[56] Phillips, P.C.B., H.R. Moon, and Z. Xiao (1998): How to Estimate Autoregressive Roots Near Unity, Cowles Foundation Discussion Paper, Yale University

[57] Phillips, P.C.B. and S. Ouliaris (1990): Asymptotic Properties of Residual Based Tests for Cointegration, *Econometrica*, 58, 165–193.

[58] Quah, D. (1994): Exploiting Cross-Section Variations for Unit Root Inference in Dynamic Data, *Economic Letters*, 44, 9–19.

[59] Robertson, D. and J. Symons (1992): Some Strange Properties of Panel Data Estimators, *Journal of Applied Econometrics*, 7, 175–189.

[60] Schmidt, P. and P.C.B. Phillips (1992): LM Test for a Unit Root in the Presence of Deterministic Trends, *Oxford Bulletin of Economics and Statistics*, 54, 257–287.

[61] Stock, J.H.(1991): Confidence Intervals for the Largest Autoregressive Root in U.S. Macroeconomic Time Series, *Journal of Monetary Economics*, 28, 435–459.

[62] Wu, S. (1997): Purchasing Power Parity Under the Currency Float : New Evidence from Panel Data Unit Root Tests, Mimeo.

[63] Wu, Y. (1996): Are Real Exchange Rates Nonstationary? Evidence from A Panel Data Test, *Journal of Money, Credit, and Banking*, 28, 54–63.

[64] Xiao, Z. and P.C.B. Phillips (1999): Efficient Detrending in Cointegrating Regression, *Econometric Theory*, forthcoming.