# Peter C.B. Phillips' Contributions to Panel Data Methods

Hyungsik Roger Moon & Benoit Perron

Univ. Southern California, University of Montreal

March 7, 2008

# Outline

- SUR
- Introduction
- Asymptotic Theories for Large $n$, $T$ Panel
- General Analysis of Linear Nonstionary Panel Regression Models
- Dynamic Panels
  - Estimation
  - Testing
- Empirics

# Seemingly Unrelated Regression I

- P.C.B Phillips (1977): An Approximation to the Finite Sample Distribution of Zellner's Seemingly Unrelated Regression Estimator, Journal of Econometrics
- P.C.B Phillips (1985): The Exact Distribution of the SUR Estimator, Econometrica

# Seemingly Unrelated Regression II

- These are early contributions in the context of the multivariate linear regression model, of which seemingly unrelated regressions model is a special case
- This differs from later contributions in that $n$ is small and $T \longrightarrow \infty$.
- Phillips (1977) used an Edgeworth expansion to obtain an approximation of the feasible (two-step) GLS estimator for the multivariate regression model with exogenous regressors. This allowed a characterization of the gains, in finite samples, of using feasible GLS over system OLS.
- Phillips (1985) characterizes the *exact* distribution of the two-step estimator in a multivariate regression model, possibly with constraints. Once again, SUR is a special case. This is made possible by the development of matrix fractional calculus.

# Introduction of Large n,T Panel Analysis

- Most existing panel research deals with large $n$, small $T$ panels
  - Large $n$, fixed $T$ asymptotics.
  - Usually do not allow for serial dependence of short time series.
  - Sometimes do not distinguish stationarity and nonstationarity.
- Availability of large $n$, $T$ panels
  - The Penn-World Table, Panel of Financial Data, even PSID (more than 20 years).
- Large $n$, $T$ asymptics may provide a better approximation.
  - Motivate to develop new panel data limit theories that are more appropriate for approximating large $n$, $T$ panels
- Need to consider serial dependence in time series more explicitly
  - Dynamic panel regressions with unit roots or near unit roots
  - Weak IV problems
- Need to distinguish stationarity and nonstationarity
  - Panel unit root tests
  - Panel cointegration analysis

# Asymptotic Theories for Large n,T Panel: Literature

- P.C.B. Phillips and H.R. Moon (1999): Linear Regression Limit Theory for Nonstationary Panel Data, Econometrica.
- P.C.B. Phillips and H.R. Moon (2000): Nonstationary Panel Data Analysis: An Overview of Some Recent Developments, Econometric Reviews.
- Around the middle of the 1990's:
  - Quah (1994), Levin and Lin (1993): $T = T(n) \to \infty$.
  - Pedroni (1997), Maddala and Wu (1997), Choi (1997), Kao (1999), Kao and Chiang (1999), and some others: $T \to \infty$, then $n \to \infty$ (sequential limit)

# Asymptotic Theories for Large n,T Panel: Phillips and Moon (1999)

Phillips and Moon (1999, Econometrica)

- Large $n$, $T$ panel model
  - Estimators and test statistics depend on $n$, $T$, where both $n$, $T$ are large.
  - A typical double index process of form $X_{n,T} = \frac{1}{k_n}\sum_{i=1}^{n} Y_{i,T}$.
- Distinguish different limit concepts:
  - Sequential limit : Let $T \to \infty$ and then $n \to \infty$.
  - Diagonal path limit : As $n \to \infty$, $T \to \infty$ along path $T(n)$.
  - Joint limit : Let $n$, $T \to \infty$.
- Find sufficient conditions under which the sequential limit is equivalent to the joint limit. (Need a certain uniformity condition)
- List regularity conditions for the CLT of double index process of $X_{n,T} = \frac{1}{k_n}\sum_{i=1}^{n} Y_{i,T}$, where $Y_{i,T}$ are cross sectionally independent.

# Linear Nonstationary Panel Regression: Literature

- P.C.B. Phillips and H.R. Moon (1999): Linear Regression Limit Theory for Nonstationary Panel Data, Econometrica.
- P.C.B. Phillips and H.R. Moon (2000): Nonstationary Panel Data Analysis: An Overview of Some Recent Developments, Econometric Reviews.

# Linear Nonstationary Panel Regression: Basic Setup

- Two nonstationary components $Z_{it} = (Y_{it}', X_{it}')'$ such that

$$\Delta Z_{it} = U_{it} = \left(U_{y,it}', U_{x,it}'\right)',$$

with

$$\Omega_i = \begin{pmatrix} \Omega_{i,yy} & \Omega_{i,xy} \\ \Omega_{i,yx} & \Omega_{i,xx} \end{pmatrix}: \text{ long run covariance matrix of } U_{it}$$

- Object of interest: Long-run average relation

$$\beta = \left(E\left[\Omega_{i,xy}\right]\right)\left[E\left(\Omega_{i,xx}\right)\right]^{-1}.$$

# Linear Nonstationary Panel Regression: Models (I)

1. Panel Spurious Regression (No Cointegration):

$$Y_{it} = \beta X_{it} + E_{it}, \qquad E_{it} = I(1).$$

2. Heterogeneous Panel Cointegration:

$$\begin{aligned} Y_{it} &= \beta_i X_{it} + E_{it}, \\ E_{it} &= I(0) \text{ and } \beta_i = \Omega_{i,yx}\Omega_{i,xx}^{-1}. \end{aligned}$$

   Note:

$$E(\beta_i) = E\left(\Omega_{i,yx_i}\Omega_{i,xx}^{-1}\right) \neq E(\Omega_{i,yx})\left[E(\Omega_{i,xx})\right]^{-1} = \beta.$$

   *c.f.* The Long-run Average Relation in Pesaran and Smith (1995): $E(\beta_i)$.

3. Homogeneous Panel Cointegration:

$$\begin{aligned} Y_{it} &= \beta X_{it} + E_{it}, \\ E_{it} &= I(0) \text{ and } \beta = \Omega_{xy}\Omega_{xx}^{-1} \end{aligned}$$

# Linear Nonstationary Panel Regression: Models (II)

4. Near Homogenous Panel Cointegration:

$$\begin{aligned} Y_{it} &= \left(\beta + \frac{\theta_i}{\sqrt{n}T}\right) X_{it} + E_{it} \\ E_{it} &= I(0) \text{ and } \beta = \Omega_{xy}\Omega_{xx}^{-1}. \end{aligned}$$

5. Models with Individual Effects:

$$Y_{it} = \gamma_i + \beta_i X_{it} + E_{it}.$$

# Linear Nonstationary Panel Regression: Findings

- The pooled ordinary least squares estimators $(\hat{\beta}' s)$ (with fixed effects in case 5) are consistent for $\beta$ in all five cases as $n, T \to \infty$.
- In 1,2, and 5 $\sqrt{n}\left(\hat{\beta} - \beta\right)$ has a normal limit distribution as $n, T \to \infty$ with $\frac{n}{T} \to 0$ (even spurious case)
- In 3 and 4 $\sqrt{n}T\left(\hat{\beta}_{PFM} - \beta\right)$ has a normal limit distribution as $n, T \to \infty$ with $\frac{n}{T} \to 0$, where $\hat{\beta}_{PFM}$ is the panel fully modified estimator (much faster rate than combines the cross-sectional $\sqrt{n}$ rate and cointegration $T$ rate)
- Hypothesis tests about the long-run average parameters both within and between individuals.

# Dynamic Panel: Basic Model

- The basic model in the literature is of the following form (or its variations):

$$Z_{it} = D_{it} + Y_{it},$$

where

$$\begin{aligned} D_{it} &= \beta_i\text{: time invariant fixed effects} \\ D_{it} &= \beta_{i0} + \beta_{i1} t : \text{ incidental trends} \\ D_{it} &= \beta_i' f_t : \text{ factor model} \end{aligned}$$

and

$$\begin{aligned} Y_{it} &= \rho Y_{it-1} + U_{it} \text{ (homogeneous panel)} \\ \text{or } Y_{it} &= \rho_i Y_{it-1} + U_{it} \text{ (heterogeneous panel)} \end{aligned}$$

- Goal:
  - How to estimate $\rho$ in the presence of incidental parameters $D_{it}$
  - How to test for $\rho_i = 1$ in the presence of incidental parameters $D_{it}$.

# Dynamic Panel – Estimation: Literature


- H.R. Moon and P.C.B. Phillips (1999): Maximum Likelihood Estimation in Panels with Incidental Trends, Oxford Bulletin of Economics and Statistics.
- H.R. Moon and P.C.B. Phillips (2000): Estimation of Autoregressive Roots near Unity using Panel Data, Econometric Theory.
- P.C.B. Phillips and D. Sul (2003): Dynamic Panel Estimation and Homogeneity Testing Under Cross Section Dependence, Econometrics Journal.
- H.R. Moon and P.C.B. Phillips (2004): GMM Estimation of Autoregressive Roots Near Unity with Panel Data, Econometrica.
- P.C.B. Phillips and D. Sul (2007): Bias in Dynamic Panel Estimation with Fixed Effects, Incidental Trends and Cross Section Dependence, Journal of Econometrics.
- C. Gourieroux, P.C.B. Phillips, and J. Yu (2006): Indirect Inference for Dynamic Panel Models, Cowles Foundation Discussion Paper.
- C. Han and P.C.B. Phillips (2007): GMM Estimation for Dynamic Panels with Fixed Effects and Strong Instruments at Unity, Cowles

# Dynamic Panel – Estimation: Background

- With time invariant fixed effects and asymptotics with $T$ is fixed
  - QMLE is inconsistent (*e.g.*, Nickel (1981, Econometrica)) – "incidental parameter problem" in dynamic panels
  - $\rho$ can be consistently estimable with IV's or GMM (*e.g.*, Anderson and Hsiao (1982, JASA), Arellano and Bond (1991, RES), Arellano and Bover, Ahn and Schmidt (1995, Journal of Econometrics))
- With time invariant fixed effects and asymptotics with $n, T \longrightarrow \infty$ with $\frac{n}{T} \longrightarrow c$,
  - Hahn and Kuersteiner (2002): Derive the bias in the limit distribution of the QMLE and suggest a bias corrected estimator of $\rho$.
- Assume that the cross sections of the panel are independent.

# Dynamic Panel – Estimation: Moon and Phillips (1999, 2000, 2004)

- Moon and Phillips (1999, 2000, and 2004) consider a homogeneous panel with near unit root as

$$\rho = 1 - \frac{c}{T}$$

  and aim to estimate the local parameter $c$ consistently in the presence of

$$D_{it} = \beta_{i0} + \beta_{i1} t$$

  as

$$n, T \to \infty.$$

- Moon and Phillips (1999, OBES)
  - With known $D_{it}$, the QMLE of $c$ is consistent when $n, T \to \infty$.
  - With unknown $D_{it}$, the QMLE of $c$ is inconsistent even if $n, T \to \infty$.
  - Called it "incidental trend problem".

# Dynamic Panel – Estimation: Moon and Phillips (1999, 2000, 2004)

- Moon and Phillips (2000, ET)
  - Restrict the paremeter set for $c$ to be $[c_{\min}, c_{\max}]$, where $0 < c_{\min} < c_{\max} < \infty$.
  - Consider various methods to correct for the incidental trend problems,
  - These include an iterative ordinary least squares (OLS) procedure and a double bias corrected estimator.
  - Show the these estimators are $\sqrt{n}-$ consistent and asymptotically normal.
  - However, exclude the case $c = 0$, unit root.

# Dynamic Panel – Estimation: Moon and Phillips (1999, 2000, 2004)

- Moon and Phillips (2004, Econometrica)
  - Allow parameter $c$ to be zero (unit root) with parameter set $[0, c_{\max}]$
  - When $\rho = 1 - \frac{c}{T}$, the conventional IV's which are futher lags of panel such as $Z_{it-2}$ become weak IV's.
  - Use modifed scores as moment conditions that hold asymptotically as $T \to \infty$
  - The first moment condition is a modified score of the OLS detrended data constructred by substracting the bias of the OLS score function
  - The second moment condition is a modified score of the GLS detrended data constructed by substracting the bias of the GLS detrended score function
  - Show that the GMM estimator of these two moment conditions is consistent and converges at rate $n^{1/6}$, much slower than the usual $n^{1/2}$
  - The paper finds an implication that the local power would be low in the presence of the incidental trends, which is confirmed later by Ploberger and Phillips and Moon, Perron, and Phillips.

# Dynamic Panel – Estimation: Bias

- A well known fact in time series is that the OLS estimator of the AR (1) coefficient has a large downward bias, especially close to one.
- An important issue in the dynamic panel literature is to understand the bias of the various estimators of $\rho$ in the presence of incidental parameters.

# Dynamic Panel – Estimation: Bias (Philips and Sul (2003))

- Phillips and Sul (2003)
  - Consider a dynamic panel model with both incidental trends and factors.
  - The factors play a role of modelling cross sectional dependence (*e.g.*, Bai and Ng (2004) and Moon and Perron (2004))
  - Illustrate how much the pooled panel OLS estimators are biased in small samples.
  - Propose new estimators under cross sectional dependence through factors: a pooled feasible generalized median unbiased estimator and a seemingly unrelated median unbiased estimator.

# Dynamic Panel – Estimation: Bias (Philips and Sul (2007))

- Phillips and Sul (2007)
  - Extend Nickell (1981)'s results.
  - Compute the bias of the QMLE of $\rho$ with incidental trends, unit root, predetermined and exogenous regressors, and errors that may be cross sectionally dependent through factors.
  - Find that the bias is large when incidental trends are fitted and $T$ is small.
  - With factors, the conventional QMLE has a random probability limit.

# Dynamic Panel – Estimation: Bias (Gourieroux, Phillips, and Yu (2006))

- Gourieroux, Phillips, and Yu (2006)
  - Use the built-in bias-reduction feature of indirect inference to reduce the bias of MLE in a dynamic panel model
  - As opposed to other bias-reduction techniques, this does not require compliacted analytical expression expressions for the bias

# Dynamic Panel – Estimation: IV (Han and Phillips (2007)

- Han and Phillips (2007)
  - Propose a simple GMM estimation method of $\rho$ with moment conditions based on first differences (with fixed effet) or second differences (with incidental trends)
  - The moment conditions used do not suffer from the weak IV or identification problem when $\rho \simeq 1$ and $\rho = 1$.
  - The estimator has a normal limit distribution for any $\rho \in (-1, 1]$ as $nT \to \infty$ (any combination of $n$ and $T$).

# Dynamic Panel – Testing: Literature


- W. Ploberger and P.C.B. Phillips (2002): Optimal testing for unit roots in panel data. Mimeo.
- P.C.B. Phillips and D. Sul (2003): Dynamic Panel Estimation and Homogeneity Testing under Cross Section Dependence, Econometrics Journal.
- H.R. Moon, B. Perron, and P.C.B. Phillips (2006): On the Breitung Test for Panel Unit Roots and Local Asymptotic Power, Econometric Theory.
- H.R. Moon, B. Perron, and P.C.B. Phillips (2007): Incidental Trends and the Power of Panel Unit Root Tests, Journal of Econometrics.

# Recall the Basic Model

- The basic model in the literature is of the following form (or its variations):

$$Z_{it} = D_{it} + Y_{it},$$

where

$$\begin{aligned} D_{it} &= \beta_i\text{: time invariant fixed effects} \\ D_{it} &= \beta_{i0} + \beta_{i1} t \text{ : incidental trends} \\ D_{it} &= \beta_i' f_t \text{ : factor model} \end{aligned}$$

and

$$\begin{aligned} Y_{it} &= \rho Y_{it-1} + U_{it} \\ \text{or } Y_{it} &= \rho_i Y_{it-1} + U_{it}. \end{aligned}$$

# Dynamic Panel – Testing: Background

- So-called first generation panel unit root tests were around (with independence).
  - Quah (1994), Levin, Lin, and Chu (2002): a modified $t-$ ratio statistic of the pooled OLS estimator with OLS detrend panel.
  - Assume cross sectional independence.
  - Im, Pesaran, Shin (2003): Based on cross section average of the individual DF test statsitics.
  - Maddala and Wu (1999), I. Choi (2001): Based on the Fisher statistics (cross section average of p-values).
  - Typically, only size is analyzed; the asymptotic distributions are obtained under the null hypothesis.
  - Usually, consistency of the test is established under a fixed alternative.
  - However, power under alternative hypothesis left to simulation: suggestive but design-dependent.
- Important issues were:
  - Panel unit root test under cross sectional dependence
  - Ananytic local power and optimality

# Dynamic Panel – Testing: Cross sectional dependence

- Phillips and Sul (2003)
  - Consider $H_0 : \rho_i = \rho$, homogeneity of the coefficient. A special case of this is when $\rho = 1$ (unit root).
  - In the case of stationarity, they propose a modified Hausman test under cross sectional dependence. The test compares the pooled feasible generalized median unbiased estimator of $\rho$ and a vector of the individual median unbiased estimator of $\rho_i$.
  - In the case of unit root, they propose a further modified Hausman type test. The test is based the same type of the estimators as in the stationary case. However, the estimators are based on orthogonalization samples where the cross sectional dependence through factors is removed.

# Dynamic Panel – Test: Power (Ploberger and Phillips(2002))

- Consider heterogeneous local alternative:

$$\rho_i = 1 - \frac{c_i}{n^\kappa T},$$

  where the value of $\kappa$ determines the neighborhood size of significant local alternatives.
- Consider a model with cross sectional independence.
- It was known that without the incidental parameters $D_{it}$, $\kappa = 1/2$ (e.g., Breitung, Moon and Perron (2004))
- Ploberger and Phillips (2002)
  - Propose an optimal invariant test that maximizes power against a weighted average of alternatives.
  - Show that with incidental trends, $D_{it} = \beta_{i0} + \beta_{i1} t$, this test has significant power with $\kappa = 1/4$ only.

# Dynamic Panel – Test: Power (Moon, Phillips, and Perron (2006,7))

- Moon, Phillips, and Perron (2006, 7)
  - Derive asymptotic power envelope assuming Gaussian innovations.
    - Without fixed effects ($D_{it}$ is known): $\kappa = 1/2$, $(n^{-1/2}T^{-1})$
    - With heterogeneous intercepts only ($D_{it} = \beta_i$): $\kappa = 1/2$, $(n^{-1/2}T^{-1})$
    - With heterogeneous intercepts and trends ($D_{it} = \beta_{i0} + \beta_{i1}t$): $\kappa = 1/4$ $\left(n^{-1/4}T^{-1}\right)$
    - With heterogenous intercepts but with homogenous trends $((D_{it} = \beta_{i0} + \beta_1 t)$: $\kappa = 1/2$, $(n^{-1/2}T^{-1})$
  - Propose feasible common point-optimal tests.
  - Compare analytical local power with other existing tests
    - Levin,-Lin-Chu test, Ploberger-Phillips test, Moon-Phillips test, Breitung test
  - Discuss the issues of initial condition and cross sectional dependence.

# Empirical contributions: Economic growth and convergence

- P.C.B. Phillips and D. Sul (2007): Some Empirics on Economic Growth under Heterogeneous Technology, Journal of Macroeconomics.
- P.C.B. Phillips and D. Sul (2007): Transition Modeling and Econometric Convergence Tests, Econometrica.

# Related literature

- Neo-classical growth mode implies that poor countries should grow faster (catch up) with larger countries
- Large cross-sections of countries (e.g. from Penn-World table) suggest overall divergence rather than convergence
- Barro and Sala-i-Martin (JPE, 1992) or Mankiw, Romer and Weil (QJE, 1992) emphasized that countries could be converging to different steady states (conditional convergence): heterogeneity is important, yet technology is still required to be homogeneous.
- The use of new panel techniques allows for more general heterogeneity

# Empirical contributions: Transition dynamics

- Decompose log per capita income (or GDP) as a nonlinear factor model:
$$\log y_{it} \left(= X_{it}\right) = \delta_{it}\mu_t$$
where $\mu_t$ is a common growth path
- The coefficients $\delta_{it}$ measure the response of country $i$ at time $t$ to this common factor. These will reflect the transition path of the economy to the common trend.
- Consider the transition relative to the average as a function of $t$ (called the transition path, or relative transition parameter):
$$h_{it} = \frac{\delta_{it}}{\frac{1}{n}\sum_{j=1}^{n}\delta_{jt}}.$$
This elminiates the effect of the common factor $\mu_t$. Convergence is then defined as:
$$\lim_{t\to\infty} h_{it} = 1 \text{ for all } i.$$
- Convergence could also occur in subgroups (convergence clubs)

# Convergence testing: log t regressions

- Time varying factor presentation

$$X_{it} = \delta_{it}\mu_t.$$

- Estimation of the factor loadings require some parametric assumption:

$$\delta_{it} = \delta_i + \frac{\sigma_i}{L(t)\, t^{\alpha}}\xi_{it}$$

where $L(t)$ is a slowly-varying function at infinity and $\xi_{it}$ is $i.i.d.(0,1)$.

- This parametrization allows the loading coefficients for a given individual to vary over time, but their variance shrinks if $\alpha > 0$. Convergence is characterized by $\delta_i = \delta$ for all $i$ and $\alpha \geq 0$.
- Clustering (or club convergence) is also allowed for if $\delta_i = \delta$ for some subset of units and $\alpha \geq 0$.

# Convergence testing: log t regressions

- With this parametrization, convergence can be tested:

$$H_0 : \delta_i = \delta \text{ and } \alpha \geq 0$$

- The statistic is just the usual (HAC) t statistic on $\hat{b}$ in the log $t$ regression:

$$\log\left(\frac{H_1}{H_t}\right) - 2\log L(t) = a + b\log t + u_t$$

where $H_t = \frac{1}{n}\sum_{i=1}^{n}(h_{it} - 1)^2$ and $b = 2\alpha$.Data starts at some fraction $r$ of the sample (.3 is recommended).

- Rejection of the convergence null does not rule out that some sub-groups converge. Phillips and Sul (2007) propose an algorithm for grouping the data among clusters (convergence clubs).
- Two applications: real GDP (JofMacro) or price levels in US cities (Econometrica).