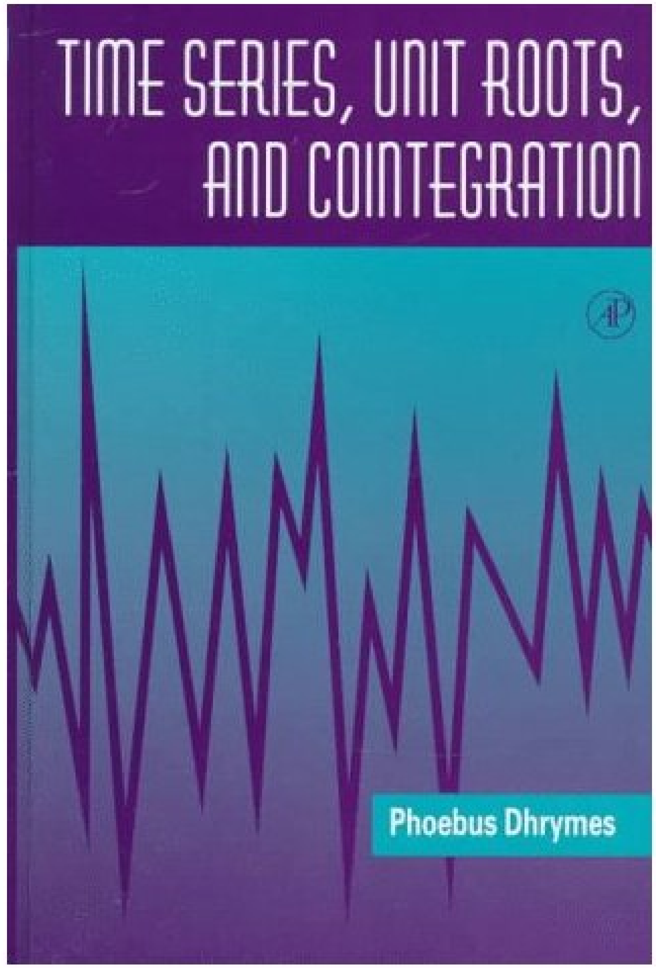
UNIT ROOT

Zhijie Xiao
Boston College

March 7, 2008

A Selective Review on Peter's Contribution to Unit Root Econometrics


TIME SERIES, UNIT ROOTS,
AND COINTEGRATION
Phoebus Dhrymes

To P.C.B. Phillips,

whose work on integrated processes

infused clarity and depth into the subject

Part I: Unit Root Time Series Asymptotics

Part II: Unit Root Tests

Part III: Other Topics in Unit Roots

Part IV: Conclusion

Part I: Unit Root Time Series Asymptotics

- Statistical foundation of unit root analysis

- The asymptotic theory applies to very general time series regression models.

- Limit theory on autoregression:

1. Stationary - Phillips and Solo (1992), Ibragimov and Phillips (2004).
2. Moderate deviations from unit root - Giraitis and Phillips (2004), Phillips and Magdalinos (2004, 2005), Ibragimov and Phillips (2004).
3. Local to unit root - Phillips (1987b, 1988), Ibragimov and Phillips (2004).
4. Unit root - Phillips (1987), Phillips and Solo (1992), Ibragimov and Phillips (2004).
5. Explosive autoregressive processes - Phillips and Magdalinos (2004, 2005), Ibragimov and Phillips (2004).

- The number of Peter's paper on this topic is very large, I will focus on three very important papers

  1. Phillips (1987) "Time Series Regression with a Unit Root", Econometrica.

  2. Phillips and Solo (1992), Asymptotics for Linear Processes, Annals of Statistics.

  3. Ibragimov and Phillips (2004), Regression Asymptotics Using Martingale Convergence Methods

Phillips (1987) “Time Series Regression with a Unit Root”, Econometrica.

1. The Semiparametric unit root tests.

2. Nonparametric correction for serial correlation and endogeneity bias

3. Unit Root Asymptotics

Asymptotics for Unit Root Autoregression

$$y_t = \alpha y_{t-1} + u_t,\ t = 1, \ldots n, \text{ with } \alpha = 1,$$

- $\mathrm{E}(u_t) = 0$

- $\sup_t \mathrm{E}|u_t|^{\beta} < \infty$, $\beta > 2$

- $\sigma^2 = \lim n^{-1} E\left(\sum_{t=1}^{n} u_t\right)^2$ exists and $\sigma^2 > 0$.

- u_t strong mixing with $\sum \alpha_m^{1-2/\beta} < \infty$

$$n\left(\widehat{\alpha} - 1\right) = \frac{n^{-1}\sum_t y_{t-1}u_t}{n^{-2}\sum_t y_{t-1}^2}$$

1. Functional limit theory for multilinear forms of weakly dependent random variables are derived by using their representations as polynomials in sample moments (via summation by parts arguments) and then using standard weak convergence results for sums of weakly dependent sequences. In particular,

   - Invariance principles for partial sums, sample variances and sample moments,

   - Convergence to stochastic integrals

2. - Classical work, extensively used in econometrics since Phillips (1987). Much of subsequent work on unit root econometric models in the last 20 years use these fundamental results.

Phillips and Solo (1992) "Asymptotics for Linear Processes", Annals of Statistics

1. An extremely **convenient and powerful** approach of time series asymptotics.

2. **Advantage**: Reducing complicate asymptotics for dependent linear processes to those of i.i.d., i.ni.d., or martingale difference sequences

   (a) An ingenious approach to develop LLN, CLT, FCLT, LIL for time series.

   (b) Applies to not only the linear process, but also products of time series, the DFT's, multivariate processes, etc.

(c) Great simplicity. e.g. No need of tightness in developing FCLT for time series.

3 Basic Idea:

simple limting theory

$$+$$

algebra of the linear operator

$$\downarrow$$

limiting theory for linear processes

- Phillips-Solo decomposition of linear processes
- Finding representation of the dependent process u_t in terms of ε_t whose asymptotics can be handled.

4 Model: Consider a very general linear process

$$u_t = C(L)\varepsilon_t,$$

and

$$C(L) = \sum_{j=0}^{\infty} c_j L^j, \ \sum_{j=0}^{\infty} j^{1/2}|c_j| < \infty, \ C(1) \neq 0, \qquad (1)$$

where ε_t can be

- $\varepsilon_t \equiv \text{iid}(0, \sigma^2)$,
- $\varepsilon_t \equiv \text{i, ni, d}(0, \sigma_t^2)$, or
- ε_t is a m.d.s. with respect to the natural filtration $(\mathcal{F}_t)$with $E(\varepsilon_t^2|\mathcal{F}_{t-1}) = \sigma_t^2$, a.s. (and $\sup_t \sigma_t^2 < \infty$).

5 **Construction**: Under the summability condition, expansion of the operator $C(L)$

$$C(L) \;=\; C(1) + \widetilde{C}(L)(L-1),$$

where $\widetilde{C}(L) = \sum_{j=0}^{\infty} \widetilde{c}_j L^j$ and $\widetilde{c}_j = \sum_{j+1}^{\infty} c_s$. This expansion gives rise to an explicit martingale difference decomposition of u_t

$$u_t \;=\; C(1)\varepsilon_t + \widetilde{\varepsilon}_{t-1} - \widetilde{\varepsilon}_t, \quad \text{with } \widetilde{\varepsilon}_t = \widetilde{C}(L)\varepsilon_t,$$

- If we consider

$$y_t = \alpha y_{t-1} + u_t,\ t = 1, ...n, \text{ with } \alpha = 1,\ u_t = C(L)\varepsilon_t,$$

$$y_t = C(1)\sum_{s=1}^{t} \varepsilon_s + \widetilde{\varepsilon}_0 - \widetilde{\varepsilon}_t + y_0 = Y_t + \eta_t,$$

where $Y_t = C(1)\sum_{s=1}^{t} \varepsilon_s$ and $\eta_t = \widetilde{\varepsilon}_0 - \widetilde{\varepsilon}_t + y_0$ are the long run and short run components of y_t, respectively.

- *Martingale decomposition*: the partial sums $\sum_{s=1}^{t} u_s$ have the leading martingale term $C(1)\sum_{s=1}^{t} \varepsilon_s$.

The decomposition was justified by **Phillips and Solo (1992)**.

Phillips and Solo (1992) showed how to use it to prove strong laws, central limit theorems, functional laws, and laws of iterated logarithms for general linear processes.

6. 2nd order Phillips-Solo decomposition for asymptotics of sample variances, sample covariances, and sample autocorrelations.

- Example

$$X_t^2 = f_0(L)\varepsilon_t^2 + 2\sum_{r=1}^{\infty} f_r(L)\varepsilon_t\varepsilon_{t-r}$$

by applying the Phillips-Solo device to $f_r(L)$, we have

$$X_t^2 = f_0(1)\varepsilon_t^2 + 2\varepsilon_t\varepsilon_{t-1}^f - (1-L)\left[\widetilde{f}_0(L)\varepsilon_t^2 + 2\sum_{r=1}^{\infty} \widetilde{f}_r(L)\varepsilon_t\varepsilon_{t-r}\right]$$

where $\varepsilon_{t-1}^f = \sum_{r=1}^{\infty} f_r(1)\varepsilon_{t-r}$.

7 Phillips-Solo decomposition in the frequency domain and applications to asymptotics in frequency domain analysis.

- Phillips-Solo decomposition in frequency domain

$$X_t = C(e^{-i\lambda})\varepsilon_t + \widetilde{\varepsilon}_{\lambda,t-1}e^{i\lambda} - \widetilde{\varepsilon}_{\lambda,t}$$

where

$$\widetilde{\varepsilon}_{\lambda,t} = \widetilde{C}_\lambda(L)\varepsilon_t,\ \widetilde{C}_\lambda(L) = \sum_{j=0}^{\infty} \widetilde{c}(\lambda)_j L^j,\ \widetilde{c}(\lambda)_j = e^{-ij\lambda}\left(\sum_{s=j+1}^{\infty} c_s e^{is\lambda}\right).$$

and decomposition to the DFT of X_t,

$$w_x(\lambda_j) = C(e^{-i\lambda_j})w_\varepsilon(\lambda_j) + r_n(\lambda_j)$$

where the remainder term $r_n(\lambda_j) = O_p(n^{-1/2})$.

Example CLT for nonparametric estimator of the LRVAR (Sun, Phillips and Jin (2006), Phillips and Xiao).

$$\widehat{f}_{xx}(0) = \frac{1}{m} \sum_{\lambda_j \in B(0)} K(\lambda_j) I_{xx}(\lambda_j),$$

Focusing on the variance component,

$$\frac{1}{m} \sum_{\lambda_j \in B(\omega)} K(\lambda_j) \left[I_{xx}(\lambda_j) - f_{xx}(\lambda_j) \right]$$

Using the Phillips-Solo device in the frequency domain, the periodogram has the following decomposition:

$$I_{xx}(\lambda_j) = I_{\varepsilon\varepsilon}(\lambda_j) \left| C(e^{-i\lambda_j}) \right|^2 + R_n(\lambda_j),$$

Note that

$$f_{xx}(\lambda_j) = \left| C(e^{-i\lambda_j}) \right|^2 \frac{\sigma^2}{2\pi}$$

Under summability and moment conditions,

$$
\begin{aligned}
& \frac{1}{\sqrt{m}} \sum_{\lambda_j \in B(0)} K(\lambda_j) \left[I_{xx}(\lambda_j) - f_{xx}(\lambda_j) \right] \\
= & \frac{1}{\sqrt{m}} \sum_{\lambda_j \in B(0)} K(\lambda_j) \left| C(e^{-i\lambda_j}) \right|^2 \left[I_{\varepsilon\varepsilon}(\lambda_j) - \frac{\sigma^2}{2\pi} \right] + O_p(\sqrt{\frac{m}{n}}) \\
= & \frac{1}{\sqrt{m}} \sum_{\lambda_j \in B(0)} K(\lambda_j) \left| C(1) \right|^2 \left[I_{\varepsilon\varepsilon}(\lambda_j) - \frac{\sigma^2}{2\pi} \right] \\
& + \frac{1}{\sqrt{m}} \sum_{\lambda_j \in B(0)} K(\lambda_j - \omega) \left[\left| C(e^{-i\lambda_j}) \right|^2 - \left| C(1) \right|^2 \right] \left[I_{\varepsilon\varepsilon}(\lambda_j) - \frac{\sigma^2}{2\pi} \right] \\
& + O_p(\sqrt{\frac{m}{n}})
\end{aligned}
$$

and the leading term

$$\begin{aligned}
& \frac{1}{\sqrt{m}} |C(1)|^2 \sum_{\lambda_j \in B(0)} K(\lambda_j) \left[I_{\varepsilon\varepsilon}(\lambda_j) - \frac{\sigma^2}{2\pi} \right] \\
\approx & \sum_{t=2}^{n} \varepsilon_t \sum_{s=1}^{t-1} \varepsilon_s \left[\frac{2}{\pi n \sqrt{m}} |C(1)|^2 \sum_{\lambda_j \in B(0)} K(\lambda_j) \cos((t-s)\lambda_j \right] \\
= & \sum_{t=2}^{n} \varepsilon_t \sum_{s=1}^{t-1} c_{t-s} \varepsilon_s + O_p(\sqrt{\frac{m}{n}}) \\
= & \sum_{t=2}^{n} z_t + O_p(\sqrt{\frac{m}{n}})
\end{aligned}$$

where

$$z_t = \varepsilon_t \sum_{s=1}^{t-1} c_{t-s} \varepsilon_s, \text{ and } c_{t-s} = \frac{\sqrt{m}}{\pi n} |C(1)|^2 \, k\left(\frac{t-s}{M}\right)$$

since the function sin(·) is symmetric around 0 and

$$k\left(\frac{h}{M}\right)=\frac{1}{m}\sum_{\lambda_j\in B(0)} K(\lambda_j)e^{-ih\lambda_j}.$$

Notice z_t is a martingale difference array, a central limiting theorem for m.d.s. can be applied if the regularity conditions hold. If $\sqrt{m}/M^q \longrightarrow 0$, as $n \longrightarrow \infty$,

$$\sqrt{m}\left[\widehat{f}_{xx}(0)-f_{xx}(0)\right] \Rightarrow N\left(0, f_{xx}(0)^2 \int k^2(x)\,dx\right)$$

Example 2. Unit Root log-periodogram regression (Phillips 1999).

- The Phillips-Solo device can also be applied to various other setting, including the 2-sided linear processes, Multivariate Time Series, etc.

- The Phillips-Solo device is now a very popular and important approach in time series asymptotic analysis.

Ibragimov and Phillips (2004): Regression Asymptotics Using Martingale Convergence Methods

- A new and conceptually simple method for obtaining weak convergence of partial sums and multilinear forms in independent random variables and linear processes to stochastic integrals.

- **Advantage**: **Great generality and wide range of applicability.**

1. Developing weak convergence of different types multilinear forms

- IP for Partial sums, sample variances and sample covariances like

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{[nr]} u_t, \frac{1}{\sqrt{n}}\sum_{t=1}^{[nr]} u_t u_{t+h},$$

- Convergence to stochastic integrals

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{[nr]}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{t-1} u_i\right) u_t \Rightarrow \int_0^r B(s)dB(s), \text{ or } \int_0^r B(s)dB(s) + r\lambda$$

- Asymptotics for general functionals of partial sums

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{[nr]} f\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{t-1} u_i\right) u_t \Rightarrow \lambda\int_0^r f^{'}(B(s))ds + \int_0^r f(B(s))dB(s)$$

2. It can be applied to a number of other fields of statistics and econometrics where convergence to Gaussian processes and stochastic integrals arises, including the study of general multilinear forms and U-statistics to multiple and stochastic integrals arises, stochastic integrals as well as asymptotics for empirical copula processes.

3 **Basic Idea**: All asymptotics reduces to the weak convergence of semimartingales, i.e. convergence of a sequence of (semi)martingales to a continuous (semi)martingale.

4 **Mechanism:** Under appropriate assumptions (say, identification of limit), reducing semimartingale convergence to **convergence of its predictable characteristics**.

5 A unified treatment of the asymptotics for both stationary, moderate deviation from unit root, local to unit root, unit root, and explosive autoregressions.

- Consider autoregression

$$y_t = \alpha y_{t-1} + u_t,\ u_t = \text{ i.i.d}(0, \sigma^2)$$

where α may be:

  - $|\alpha| < 1$: Stationary
  - $\alpha = 1 + c/n^b$, where $0 < b < 1$ and $c < 0$: Moderate deviations from unit root)
  - $\alpha = 1 + c/n$: Local to unit root
  - $\alpha = 1$: Unit root
  - $\alpha > 1$: Explosive autoregression (together with Martingale convergence theorem)

- Ibragimov and Phillips (2004) show that an identical construction delivers the limiting theory in all these cases, i.e. the same construction gives
  - CLT for stationary and moderate deviations from unit root
  - weak convergence to stochastic integrals for local to unit root, unit root
- All these cases can be transformed into martingale convergence problem and different rates of convergence are accommodated in a natural way.

6 This new approach uses a very general convergence result for semimartingales obtained by Jacod and Shiryaev (2003), and overcome some technical problems in the existing literature. General sufficient conditions for the original assumptions of semimartingale convergence theorems for multivariate diffusion processes are given.

- Semimartingales:

$$
\begin{aligned}
M(r) &= M(0) + M^0(r) + B(r) \leftarrow\text{-- by Doob-Meyer decomposition} \\
M(0) &: \text{initalization, finite valued F-measurable random variable} \\
M^0(r) &: \text{local MG} \\
B(r) &: \text{finite variation process of } M
\end{aligned}
$$

- Predictable characteristics of continuous semimartingales:

$$
\begin{aligned}
B(r) &: \text{first predictable characteristic of } M, \\
C(s) &= [M, M](s) = \text{quadratic variation of } M \\
&: \text{second predictable characteristic of } M,
\end{aligned}
$$

- Convergence of a sequence of semimartingales ($M_n \xrightarrow{d} M$) holds if

  1. Their predictable characteristics (B_n and C_n) and the initial distributions ($M_n(0)$) tend to those of the limit semimartingale (i.e. $B, C, M(0)$).

  2. The predictable characteristics of the limit process grow in a regular way

  3. The process is the only continuous semimartingale with characteristics B and C and the initial distribution.

Then

$$M_n \xrightarrow{d} M$$

7 Example: Ibragimov and Phillips (2004) show how to analyze the limiting behavior of

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{[nr]}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{t-1}u_i\right)u_t,$$

a Embedding $n^{-1}\sum_{t=1}^{[nr]}\left(\sum_{i=1}^{t-1}u_i\right)u_t$ in continuous martingale $M_n(r)$, and, under appropriate assumptions, show convergence of MG

$$M_n(r)\to M(r)=\int_0^r B\mathrm{d}B.$$

if (the predictable characteristics of $M_n(r)$) $C_n(r)=[M_n]_r\to[M]_r$.

b or, using a discrete jump version of $M_n(r)$.

8 Example: Unification of the limit theory of autoregression. Consider the OLS estimator

$$\widehat{\alpha} = \frac{\sum_t y_{t-1} y_t}{\sum_t y_{t-1}^2}$$

we consider the recursive OLS estimation

$$\widehat{\alpha}_r = \frac{\sum_{t=1}^{[nr]} y_{t-1} y_t}{\sum_{t=1}^{[nr]} y_{t-1}^2},$$

$$\widehat{\alpha}_r - \alpha = \frac{\sum_{t=1}^{[nr]} y_{t-1} u_t}{\sum_{t=1}^{[nr]} y_{t-1}^2}$$

.

If we define

$$M_n(r) \;=\; \begin{cases} \frac{1}{\sqrt{n}} \sum_{t=1}^{[nr]} y_{t-1} u_t, & |\alpha| < 1 \\ \frac{1}{n^{(1+b)/2}} \sum_{t=1}^{[nr]} y_{t-1} u_t, & \alpha = 1 + c/n^b \\ \frac{1}{n} \sum_{t=1}^{[nr]} y_{t-1} u_t, & \alpha = 1 \\ \frac{1}{n} \sum_{t=1}^{[nr]} y_{t-1} u_t, & \alpha = 1 + c/n \end{cases}$$

(Embedding $\sum_{t=1}^{[nr]} (y_{t-1}) \, u_t$ in continuous martingale $M_n(r)$)

Normalized versions of the estimation error are represented in MG form as a ratio

$$\frac{M_n(r)}{([M_n]_r)^{1/2}}$$

where $M_n(r)$ is a MG with quadratic variation $[M_n]_r$, and the limit theory is delivered by MG convergence in the form

$$\frac{M_n(r)}{([M_n]_r)^{1/2}} \xrightarrow{d} \frac{M(r)}{([M]_r)^{1/2}}$$

where $M(r)$ is the limiting MG process.

If we define MG

$$M_n(r) = \begin{cases} \frac{1}{\sqrt{n}} \sum_{t=1}^{[nr]} y_{t-1}u_t, & |\alpha| < 1 \\ \frac{1}{n^{(1+b)/2}} \sum_{t=1}^{[nr]} y_{t-1}u_t, & \alpha = 1 + c/n^b \\ \frac{1}{n} \sum_{t=1}^{[nr]} y_{t-1}u_t, & \alpha = 1 \\ \frac{1}{n} \sum_{t=1}^{[nr]} y_{t-1}u_t, & \alpha = 1 + c/n \end{cases}$$

the corresponding square bracket q.v. process $[M_n]_r$, i.e. process for which $M_n(r)^2 - [M_n]_r = MG$ is

$$[M_n]_r = \begin{cases} \frac{1}{n} \sum_{t=1}^{[nr]} y_{t-1}^2 u_t^2, & |\alpha| < 1 \\ \frac{1}{n^{1+b}} \sum_{t=1}^{[nr]} y_{t-1}^2 u_t^2, & \alpha = 1 + c/n^b \\ \frac{1}{n^2} \sum_{t=1}^{[nr]} y_{t-1}^2 u_t^2, & \alpha = 1 \\ \frac{1}{n^2} \sum_{t=1}^{[nr]} y_{t-1}^2 u_t^2, & \alpha = 1 + c/n \end{cases}$$

and the compensator $\langle M_n \rangle_r$ of $[M_n]_r$ i.e. $[M_n]_r - \langle M_n \rangle_r = MG$, (the second characteristic)

$$\langle M_n \rangle_r = \begin{cases} \frac{1}{n} \sum_{t=1}^{[nr]} y_{t-1}^2 \sigma^2, & |\alpha| < 1 \\ \frac{1}{n^{1+b}} \sum_{t=1}^{[nr]} y_{t-1}^2 \sigma^2, & \alpha = 1 + c/n^b \\ \frac{1}{n^2} \sum_{t=1}^{[nr]} y_{t-1}^2 \sigma^2, & \alpha = 1 \\ \frac{1}{n^2} \sum_{t=1}^{[nr]} y_{t-1}^2 \sigma^2, & \alpha = 1 + c/n \end{cases}$$

Then,

$$M_n(r) \to M(r) = \begin{cases} \sigma_y \sigma N(0,r) = \sigma_y \sigma W(r), & |\alpha| < 1 \\ \sqrt{\frac{\sigma^2}{-2c}} N(0,r), & \alpha = 1 + c/n^b \\ \sigma^2 \int_0^r W(s) dW(s), & \alpha = 1 \\ \sigma^2 \int_0^r W_c(s) dW(s), & \alpha = 1 + c/n \end{cases}$$

and

$$\langle M_n \rangle_r \to \langle M \rangle_r = \begin{cases} r\sigma_y^2 \sigma^2, & |\alpha| < 1 \\ \frac{1}{-2c} r\sigma^2, & \alpha = 1 + c/n^b \\ \sigma^4 \int_0^r W^2, & \alpha = 1 \\ \sigma^4 \int_0^r W_c^2, & \alpha = 1 + c/n \end{cases}$$

where $\langle M \rangle_r$ is the second predictable characteristic of the continuous martingale M.

$$\left(\frac{\sum_{t=1}^{[nr]} y_{t-1}^2}{\sigma^2}\right)^{1/2} (\widehat{\alpha}_r - \alpha)$$

$$= \frac{M_n(r)}{(\langle M_n \rangle_r)^{1/2}}$$

$$\xrightarrow{d} \frac{M(r)}{(\langle M \rangle_r)^{1/2}} = \begin{cases} N(0,1), & |\alpha| < 1 \\ N(0,1), & \alpha = 1 + c/n^b \\ \left[\int_0^r W^2\right]^{-1/2} \int_0^r W(s)dW(s), & \alpha = 1 \\ \left[\int_0^r W_c^2\right]^{-1/2} \int_0^r W_c(s)dW(s), & \alpha = 1 + c/n \end{cases}$$

which unifies the limit theory for stationary and unit root autoregression.

Part I Unit Root Time Series Asymptotics

Part II Unit Root Tests

(a) Semiparametric tests Z_α and Z_t and the approach of nonparametric treatment of serial correlation and endogeneity.

(b) Variants of the semiparametric unit root tests.

(c) Other tests

(d) Local Power of Unit Root Tests and Local to Unity Asymptotics

(e) Residual based tests for the null of no cointegration

Part III Other Topics

Part II: Unit Root Tests

1. Semiparametric unit root tests - treating the unknown correlation nonparametrically: (Phillips 1987).

$$\begin{aligned} Z_\alpha &= n(\hat{\alpha}-1) - \hat{\lambda}\left(n^{-2}\sum_{t=2}^{n} y_{t-1}^2\right)^{-1} \\ &\Rightarrow \left[\int_0^1 W\,dW\right]\left[\int_0^1 W^2\right]^{-1}, \\ Z_t &= \hat{\sigma}_u \hat{\omega}^{-1} t_\alpha - \hat{\lambda}\left\{\hat{\omega}\left(n^{-2}\sum_{t=2}^{n} y_{t-1}^2\right)^{1/2}\right\}^{-1} \\ &\Rightarrow \left[\int_0^1 W\,dW\right]\left[\int_0^1 W^2\right]^{-1/2}. \end{aligned}$$

2 Peter's device of the semiparametric tests not only provide important unit root tests, it also opens a new direction to nonparametrically dealing with serial correlation and endogeneity in nonstationary time series.

- This idea has important impact on later studies. The technique of nonparametric correction for serial correlation and endogeneity now becomes an important technique in time series applications.

- The idea was further extended to models with convergence to multivariate Brownian motions, especially the technique Fully Modification in Cointegrating Regression and VAR system. Phillips and Hansen (1991), Phillips (1995).

3 The basic idea of the semiparametric tests Z_α and Z_t can be applied to many different settings and models - leading to various versions of the semiparametric unit root tests. Phillips and Perron (1988), Ouliaris, Park and Phillips (1989), Park and Sung (1994) and many other researchers give various extensions of these semiparametric tests.

(a) With a time trend: OLS detrending.

(b) With a time trend: QD detrending.

(c) LAD estimation:

(d) M-estimation:

a. With a time trend: OLS detrending. Phillips and Perron (1988), Ouliaris, Park and Phillips (1989), etc.

$$y_t = \gamma' x_t + y_t^s, \;\; y_t^s = \alpha y_{t-1}^s + u_t, \text{ with } \alpha = 1,$$

let $\widehat{y}_t = y_t - \widehat{\gamma}' x_t$, and

$$\widehat{\alpha}_{OLS} = \arg\min \sum_t (\widehat{y}_t - \alpha \widehat{y}_{t-1})^2$$

$$\begin{aligned}
Z_{\alpha,OLS} &= n(\hat{\alpha}_{OLS} - 1) - \hat{\lambda} \left(n^{-2} \textstyle\sum_{t=2}^n y_{X,t-1}^2 \right)^{-1} \\
&\Rightarrow \left[\textstyle\int_0^1 W_X dW \right] \left[\textstyle\int_0^1 W_X^2 \right]^{-1}, \\
Z_{t,OLS} &= \hat{\sigma}_u \hat{\omega}^{-1} t_\alpha - \hat{\lambda} \left\{ \hat{\omega} \left(n^{-2} \textstyle\sum_{t=2}^n y_{X,t-1}^2 \right)^{1/2} \right\}^{-1} \\
&\Rightarrow \left[\textstyle\int_0^1 W_X dW \right] \left[\textstyle\int_0^1 W_X^2 \right]^{-1/2}.
\end{aligned}$$

where $y_{X,t}$ is the residual from a regression of y_t on x_t.

b. With a time trend: QD detrending. If we can consider alternatives that are closer to unity, for some fixed $c = \bar{c}$, define the quasi-difference operator as $\Delta_{\bar{c}}$, $\Delta_{\bar{c}} y_t = (1 - L - n^{-1}\bar{c}L)y_t = \Delta y_t - n^{-1}\bar{c}y_{t-1}$, take quasi-differences and run the detrending regression

$$\Delta_{\bar{c}} y_t = \tilde{\gamma}' \Delta_{\bar{c}} x_t + \Delta_{\bar{c}} \tilde{y}_t^s. \tag{2}$$

and

$$\tilde{y}_t = y_t - \tilde{\gamma}' x_t. \tag{3}$$

We can construct the modified semi-parametric Z_α test based on

$$\widehat{\alpha}_{QD} = \text{arg min} \sum_t (\tilde{y}_t - \alpha \tilde{y}_{t-1})^2$$

The modified Z_α test statistic has the following form

$$\tilde{Z}_{\alpha,QD} = n(\widehat{\alpha}_{QD} - 1) - \tilde{\lambda}\Big(n^{-2}\sum\nolimits_{t=2}^{n}\tilde{y}_{t-1}^2\Big)^{-1}$$
$$\Rightarrow\Big[\int_0^1 \tilde{W}_{\overline{c}}^2\Big]^{-1}\int_0^1 \tilde{W}_{\overline{c}} d\tilde{W}_{\overline{c}},$$

where

$$\tilde{W}_{\overline{c}}(r) = W(r) - \int_0^1 dW_{\overline{c}} X_{\overline{c}}' \left(\int_0^1 X_{\overline{c}} X_{\overline{c}}'\right)^{-1} X(r)$$

is the weak limit of $n^{-\frac{1}{2}}\widetilde{y}_{[nr]}$, $W_{\overline{c}}(r) = W(r) - \overline{c}\int_0^r W(s)$, $\tilde{\lambda}$ is a consistent estimator of λ, $X_{\overline{c}}(r) = X'(r) - \overline{c}X(r)$ is the limit function of the quasi-differenced deterministic trend.

Using the same idea, we can construct the modified Z_t tests and the corresponding limit theory for these tests is

$$\widetilde{Z}_t \Rightarrow \Big[\int_0^1 \tilde{W}_{\overline{c}}^2\Big]^{-1/2}\int_0^1 \tilde{W}_{\overline{c}} d\tilde{W}_{\overline{c}}. \tag{4}$$

c. LAD estimation: Herce (1996). In the case of LAD or M-estimation, we have convergence to bivariate Brownian motions and the situation is similar to the case of fully modification of Phillips and Hansen (1991), Phillips (1995). If

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{[nr]} \begin{pmatrix} u_t \\ \text{sign}(u_{t-1}) \end{pmatrix} \Rightarrow \begin{bmatrix} B_1(r) \\ B_2(r) \end{bmatrix}$$

and we consider

$$\widehat{\alpha}_{LAD} = \text{arg min} \sum_t |y_t - \alpha y_{t-1}| \,,$$

then

$$n(\hat{\alpha}_{LAD} - 1)$$

$$\Rightarrow \frac{1}{2f(F^{-1}(0))}\left[\int_0^1 B_1^2\right]^{-1}\left[\int_0^1 B_1 dB_2 + \lambda_{u.\mathrm{sign}(u)}\right]$$

$$= \frac{1}{2f(F^{-1}(0))}\left[\int_0^1 B_1^2\right]^{-1}\left[\int_0^1 B_1 dB_{2.1} + \Delta_{1.2}\int_0^1 B_1 dB_1 + \lambda_{u.\mathrm{sign}(u)}\right]$$

A modification similar to those of fully modification can be applied to both the coefficient based statistic and the t-ratio, and thus semiparametric tests for a unit root can be constructed.

d. M-estimation: Lucas 1995. Although the technique in deriving the asymptotics are different, the idea of constructing the semiparametric tests are the same. Let $\rho'(u) = \psi(u)$, If

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{[nr]}\begin{pmatrix} u_t \\ \psi(u_t) \end{pmatrix} \Rightarrow \begin{bmatrix} B_1(r) \\ B_2(r) \end{bmatrix}$$

and we consider

$$\widehat{\alpha}_M = \text{arg min} \sum_t \rho\left(y_t - \alpha y_{t-1}\right),$$

then

$$n(\hat{\alpha}_M - 1) \Rightarrow \frac{1}{E\left[\psi'(u_t)\right]}\left[\int_0^1 B_1^2\right]^{-1}\left[\int_0^1 B_1 dB_2 + \lambda_{u.\psi(u)}\right].$$

And the rest of nonparametric modifications are similar.

Rank Based test

Extensions, modifications of the original semiparametric test

Local-to-unity Asymptotics and Local Power of Unit Root Tests

- Peter's work on local-to-unit root processes plays a fundamental role to the local power analysis of unit root tests.

- Phillips, 1987b, Towards a unified asymptotic theory of autoregression, Biometrika, 74, 535-547.

- Under the local alternative hypothesis $\alpha = \exp(n^{-1}c) \sim 1 + n^{-1}c$,

$$n^{-1/2} y^s_{[nr]} \Rightarrow J_c(r) = \int_0^r e^{(r-s)c} dW(s)$$

$$Z_\alpha \Rightarrow c + \left[\int_0^1 J_{cX} dW\right]\left[\int_0^1 J_{cX}^2\right]^{-1}, \tag{5}$$

where

$$J_{cX}(r) = J_c(r) - \left(\int_0^1 J_c X'\right)\left(\int_0^1 X X'\right)^{-1} X(r).$$

- The local asymptotic theory can be used to construct asymptotic power envelopes for unit root tests.

- In addition to his work on local to unit root, Peter has several important recent papers on the study of autoregressive time series with root moderately deviating from unity:
  - Giraitis and Phillips (2004), Phillips and Magdalinos (2004, 2005).

Other Tests

- . . . .

Part II Unit Root Tests

Part III Other Topics

(a) KPSS test for Stationarity and its applications and Extensions

(b) Peter's Shortcut and Robust inference in nonstationary time series

(c) Deterministic Trend and Trend Breaks

(d) Long Memory Processes

(e) PIC and model selection based unit root testing

(f) Nonlinear and nonparametric estimation on unit root processes

(g) . . . .

KPSS test for Stationarity and its applications and Extensions

1 The KPSS test is derived based on a *components* model

$$y_t = h_t + y_t^s + v_t, \quad y_t^s = y_{t-1}^s + u_t, \tag{6}$$

which decomposes the time series y_t into a deterministic trend h_t, a stochastic trend y_t^s, and a stationary residual v_t.

2 The process is trend stationary when $\sigma_u^2 = \mathsf{var}(u_t) = 0$.

3 Under Gaussian assumptions and iid error conditions, the null hypothesis of stationarity can be tested using the LM principle. Let $\widehat{e}_t$ be the residuals from the regression of y_t on the deterministic trend x_t and $\widehat{\sigma}_v^2 = n^{-1} \sum \widehat{e}_t^2$, then the LM statistic can be constructed as follows:

$$LM = \frac{n^{-2} \sum S_t^2}{\widehat{\sigma}_v^2},$$

where $S_t = \sum_{j=1}^{t} \widehat{e}_j$.

4 Under the null hypothesis of stationarity, this LM statistic converges to $\int_0^1 V_X^2$, where

$$V_X(r) = W(r) - \left[\int_0^r X'\right] \left[\int_0^1 X X'\right]^{-1} \left[\int_0^1 X dW\right]$$

is a generalized Brownian bridge process.

5 In the case that v_t is a general stationary residual, we may replace the variance estimator by a lrvar estimator.

6 Intuitively, if y_t is a stationary time series, it has a fixed mean, finite variance and cannot grow indefinitely. However, an unstable (unit root) or explosive process has unbounded variance and grows over long period of time. As a result, the fluctuation of a unit root or explosive process is much larger than that of a stationary process. Thus we can test whether or not y_t is stationary by looking at the fluctuation in the time series. If we use a Cramer-von-Mises measure for the fluctuation, we obtain the KPSS test.

1. KPSS test has stimulated lots of extensions and modifications along various directions.

2. Several directions of extensions

   (a) Extensions to different models.

   (b) Extensions based on different measurement of fluctuation

   (c) Extensions based on different estimation methods

a. Extensions to different models. e.g.

- Leybourne and McCabe (1994) suggested a similar test for stationarity which differs from the test of Kwiatkowski et al. (1992) in its treatment of autocorrelation and applies when the null hypothesis is an AR(k) process.

- Residual based test for the null of cointegration, Shin (1995), Xiao and Phillips (2000).

b. Extensions based on different measurement of fluctuation, say Kolmogorov-Smirnoff.

c. Extensions based on different estimation methods.

- e.g. if we consider a quantile regression of y_t on a deterministic trend,

$$\min \sum_t \rho_\tau(y_t - \beta' x_t)$$

let

$$\widehat{u}_{t\tau} = y_t - \beta(\tau)' x_t$$

then we may replace $\widehat{e}_t$ by

$$\psi_\tau(\widehat{u}_{t\tau}), \text{ where } \psi_\tau(u) = \tau - I(u < 0).$$

and construct a QR-based KPSS test for stationarity. Such a test only require weaker moment conditions.

Peter's Shortcut and Robust Inference on Unit Root Models

- Phillips, P.C.B., A Shortcut to LAD Estimator Asymptotics, Econometric Theory, 7, 1991, 450-463.

- Phillips, P.C.B., Robust Nonstationary Regression, Econometric Theory, 11, 1995, 912-951.

- The approach treats nonsmooth criterion functions as generalized functions and uses generalized Taylor series expansions to represent their local behavior.

The impact of these papers go much beyond unit root problems. I focus here on unit root regressions and tests.

Example: consider autoregression

$$y_t = \alpha y_{t-1} + u_t$$

where $u_t = i.i.d.(0, \sigma^2)$, with CDF $F(\cdot)$. If we consider a QR on the above model,

$$\min \sum_t \rho_\tau(y_t - \mu - \alpha y_{t-1})$$

Let $\theta = (\mu, \alpha)^\top$, $x_t = (1, y_{t-1})$, $D_n = diag(\sqrt{n}, n)$,

$$\widehat{\theta}(\tau) = \arg\min_\theta \sum_t \rho_\tau(y_t - \theta' x_t)$$

Denote

$$\theta(\tau) = (F^{-1}(\tau), \alpha)^\top,\ \eta = D_n\left(\theta - \theta(\tau)\right),\ \widehat{\eta}(\tau) = D_n\left(\widehat{\theta}(\tau) - \theta(\tau)\right),$$

then

$$\widehat{\eta}(\tau) = \text{arg min}\, Z_n(\eta).$$

where

$$Z_n(\eta) = \sum_t \left\{\rho_\tau(u_{t\tau} - D_n^{-1} x_t^\top \eta) - \rho_\tau(u_{t\tau})\right\}$$

Following Phillips (1995), under regularity conditions, notice that

1. $\rho_\tau(u)$ can be treated as a generalized function with a (smooth) regular sequence

$$\rho_{\tau m}(u) = \int_{-\infty}^{\infty} \rho_\tau(v) S\left[m(v-u)\right] m e^{-v^2/m^2} dv$$

where $S(\cdot)$ is a smudge function whose role in $\rho_{\tau m}(u)$ is to smudge out $\rho_\tau(v)$ when v is outside the interval $(u - m^{-1}, u + m^{-1})$ (see Phillips (1995) for more discussions),

2. Then $Z_n(\eta)$ is a generalized process defined by the following regular sequence of processes

$$Z_{nm}(\eta) = \sum_t \left\{ \rho_{\tau m}(u_{t\tau} - D_n^{-1} x_t^\top \eta) - \rho_{\tau m}(u_{t\tau}) \right\}.$$

3. By Taylor expansion of $Z_{nm}(\eta)$ around $\eta = 0$, we have

$$
\begin{aligned}
Z_{nm}(\eta) &= -\sum_t \dot{\rho}_{\tau m}(u_{t\tau}) D_n^{-1} x_t^\top \eta \\
&\quad + \frac{1}{2}\eta^\top \left[\sum_t \ddot{\rho}_{\tau m}(u_{t\tau} - \lambda D_n^{-1} x_t^\top \eta) D_n^{-1} x_t x_t^\top D_n^{-1} \right] \eta
\end{aligned}
$$

where $\dot{\rho}_{\tau m}(\cdot)$ and $\ddot{\rho}_{\tau m}(\cdot)$ are first and second order derivatives of $\rho_{\tau m}(\cdot)$ and $\lambda \in (0, 1)$.

4. It can be shown that

$$\sum_t \dot{\rho}_{\tau m}(u_{t\tau}) D_n^{-1} x_t^\top \eta \Rightarrow \xi_m^\top \eta$$

where

$$\xi_m \Rightarrow \xi = \left[\int \overline{B}_y(r) dB_\psi^\tau(r) \right], \text{ as } m \to \infty.$$

For the second term, notice that the regular sequence $\delta_m(\cdot)$ is differentiable and has bounded derivative, and

$$\sum_t \ddot{\rho}_{\tau m}(u_{t\tau}) D_n^{-1} x_t x_t^\top D_n^{-1} \Rightarrow \left(\mathsf{E}\left[\ddot{\rho}_{\tau m}(u_{t\tau})\right]\right) \int \overline{B}_y(r)\overline{B}_y(r)^\top dr$$

(notice that $\mathsf{E}[\ddot{\rho}_{\tau m}(u_{t\tau})]$ is an ordinary function as long as density at $F^{-1}(\tau)$ exists and continues). and as $m \to \infty$,

$$\mathsf{E}\left[\ddot{\rho}_{\tau m}(u_{t\tau})\right] \to f(F^{-1}(\tau)) \ .$$

Then

$$Z_{nm}(\eta) \Rightarrow Z_{\cdot m}(\eta) = -\xi_m^\top \eta + \left(\mathsf{E}\left[\ddot{\rho}_{\tau m}(u_{t\tau})\right]\right) \frac{1}{2} \eta^\top \int \overline{B}_y(r)\overline{B}_y(r)^\top dr \eta,$$

and as $m \to \infty$, over η on a compact set,

$$Z_{\cdot m}(\eta) \Rightarrow Z(\eta) = -\xi^\top \eta + f(F^{-1}(\tau)) \frac{1}{2} \eta^\top \int \overline{B}_y(r)\overline{B}_y(r)^\top dr \eta,$$

Notice that

1. $Z_{\cdot m}(\eta)$ is convex

2. $Z_{\cdot m}(\eta) \Rightarrow Z(\eta)$

3. $Z(\eta)$ has an unique minimum at $\frac{1}{f(F^{-1}(\tau))}\left[\int \overline{B}_y \overline{B}_y{}^\top\right]^{-1}\xi$

$$= \arg\min\left[-\xi^\top \eta + f(F^{-1}(\tau))\frac{1}{2}\eta^\top \int \overline{B}_y(r)\overline{B}_y(r)^\top dr\eta\right]$$

we have

$$D_n\left(\widehat{\theta}(\tau) - \theta(\tau)\right) \Rightarrow \frac{1}{f(F^{-1}(\tau))}\left[\int \overline{B}_y \overline{B}_y{}^\top\right]^{-1} \int \overline{B}_y dB_\psi^\tau$$

In particular

$$n\left(\widehat{\alpha}(\tau)-1\right) \Rightarrow \frac{1}{f(F^{-1}(\tau))}\left[\int \underline{B}_y \underline{B}_y{}^{\top}\right]^{-1} \int \underline{B}_y dB_{\psi}^{\tau}$$

where $\underline{B}_y(r)$ is a demeaned BM.

The above results can then be applied to construct robust inference procedures.

Deterministic Trend and Trend Breaks

The introduction of trend break functions leads to further reductions in the power of unit root tests and to substantial finite sample size distortion in the tests.

Sample trajectories of a random walk are often similar to those of a process that is stationary about a broken trend for some particular breakpoint.

In view of the fact that Brownian motion can be represented as an infinite linear random combination of deterministic functions of time, carefully chosen trend stationary models can always be expected to provide reasonable representations of given random walk data, but such models are certain to fail in post sample projections as the post sample data drifts away from the final trend line.

- Phillips (1998a): New tools for understanding spurious regression

  Stochastic trend can be validly represented in empirical regressions in terms of deterministic function of time.

- Phillips (1998b): New unit root asymptotics in the presence of deterministic trend

  – Critical values of unit root tests diverge when the number of deterministic regressors $K \to \infty$ as sample size $n \to \infty$.

  – Serious attempts to model trends by deterministic functions will always be successful and that these functions can validly represent stochastically trending data, thereby underming conventional unit root tests.

Long Memory Processes and Unit Root Tests Against Fractional Alternatives

Summary

- Peter's contribution in unit root econometrics is enormous and fundamental.

- Peter made the most important contribution on unit roots.

- Important work in almost every sub-field in unit root studies.

- Still very actively leading the research in this field.

Happy Birthday, Peter!