UNIT ROOT TOPIC

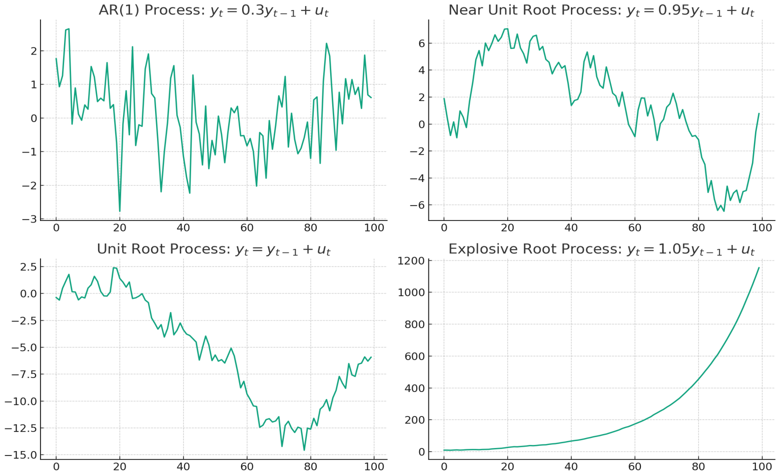
Wang Shaoping, Li Yanglin, Wang Xinyu

Huazhong University of Science and Technology, School of Economics

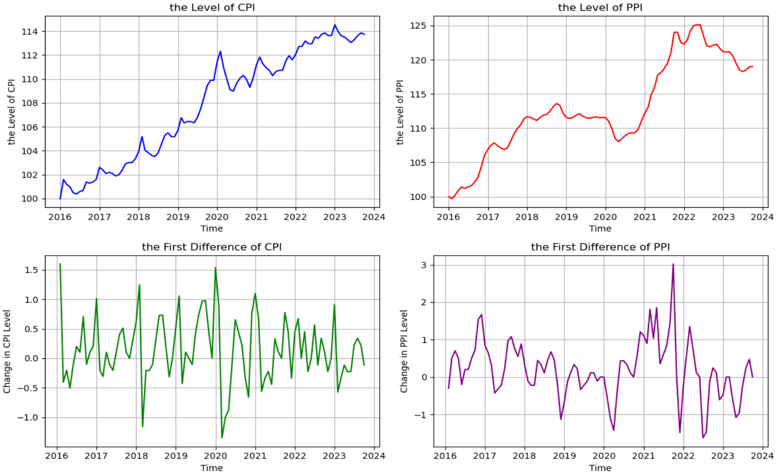
December 14, 2023

Outline

Simulation of Different Time Series Process



A Glimpse of Macroeconomics Time Series Process



A Glimpse of Macroeconomics Time Series Process

We chose Jan, 2016 - Oct, 2023 as the sample period, and transformed the time series by setting Jan, 2016 as the base period. We regarded these two series as the level of CPI and PPI, respectively. Moreover, we took the first difference on these two time series.

Stationary Time Series

Definition

Any stationary AR(1) process $y_t = \rho y_{t-1} + u_t (\rho < 1)$ has constant mean, variance and covariance.

Proof.

$y_t = \rho y_{t-1} + u_t = \rho^2 y_{t-2} + u_t + \rho u_{t-1} = \ldots = \sum_{i=0}^{\infty} \rho^i u_{t-i}$
Due to $\rho < 1$, so the sequence is absolutely summable, we have:
$E(y_t) = E(\sum_{i=0}^{\infty} \rho^i u_{t-i}) = \sum_{i=0}^{\infty} \rho^i E(u_{t-i}) = 0$
$var(y_t) = var(\sum_{i=0}^{\infty} \rho^i u_{t-i}) = \sum_{i=0}^{\infty} \rho^i var(u_{t-i}) = \sigma^2 \cdot \frac{1}{1-\rho^2}$

$$\gamma_j = E(y_t - \mu)(y_{t-j} - \mu)$$
$$= E[u_t + \rho u_{t-1} + \rho^2 u_{t-2} + \cdots + \rho^j u_{t-j} + \rho^{j+1} u_{t-j-1}$$
$$+ \rho^{j+2} u_{t-j-2} + \cdots] \times [u_{t-j} + \rho u_{t-j-1} + \rho^2 u_{t-j-2} + \cdots]$$
$$= [\rho^j + \rho^{j+2} + \rho^{j+4} + \cdots] \cdot \sigma^2 = [\rho^i/(1-\rho^2)] \cdot \sigma^2.$$

□

Test for Stationary

When $\rho < 1$ is satisfied, the limit distributions of coefficient-based statistics and testing statistics are

$$\sqrt{T}(\hat{\rho} - \rho) \sim N(0, 1 - \rho^2)$$

$$t = \frac{\hat{\rho} - \rho}{se(\hat{\rho})} \sim t(T - 1)$$

Proof I

As to AR(1) $y_t = \rho y_{t-1} + u_t$, the OLS estimator of ρ satisfies

$$\sqrt{T}(\hat{\rho} - \rho) \sim N(0, \sigma^2 (\frac{Y'_{-1} Y_{-1}}{T})^{-1}) = N(0, \sigma^2 [var(y_{t-1})]^{-1})$$

y_t is stationary under H_1: $\rho < 1$, and we have a result that

$$var(y_{t-1}) = \gamma_0 = \frac{\sigma^2}{1 - \rho^2}$$

So

$$\sqrt{T}(\hat{\rho} - \rho) \sim N(0, \sigma^2 \frac{1 - \rho^2}{\sigma^2}) = N(0, 1 - \rho^2)$$

Proof II

and $\sum_{t=1}^{T} \frac{\hat{u}_t^2}{\sigma^2} = (T-1)\frac{\hat{\sigma}^2}{\sigma^2} \sim \chi^2(T-1)$, the limit distributions of coefficient-based statistics is:

$$t = \frac{\hat{\rho}-\rho}{se(\hat{\rho})} = \frac{\hat{\rho}-\rho}{\sqrt{\hat{\sigma}^2/\sum_{t=1}^{T} y_{t-1}^2}} = \frac{1}{\sqrt{\hat{\sigma}^2/\sigma^2}} \cdot \frac{\sqrt{T}(\hat{\rho}-\rho)}{\sqrt{\sigma^2(\frac{1}{T}\sum_{t=1}^{T} y_{t-1}^2)^{-1}}}$$

$$\sim \frac{N(0,1)}{\sqrt{\chi^2(T-1)/(T-1)}} = t(T-1)$$

Unit Root Process

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

is called Random Walk or Unit Root Process, by iteration:

$$y_t = y_{t-1} + u_t = y_{t-2} + u_t + u_{t-1} = \ldots = \sum_{i=0}^{t} u_i$$

which has no drift or time trend term. It's obviously nonstationary, but its difference $\triangle y_t = y_t - y_{t-1} = u_t$ is stationary.
So a random walk without drift or trend term is difference-stationary.

Unit Root Process

There are two other unit root processes:

- With drift term but no time trend term

$$y_t = c + y_{t-1} + u_t = 2c + y_{t-2} + u_t + u_{t-1} = \ldots = ct + \sum_{i=0}^{t} u_i$$

  whose difference $\triangle y_t = \beta_0 + u_t$ is a stationary process, so a random walk with only drift term y_t is difference-stationary.

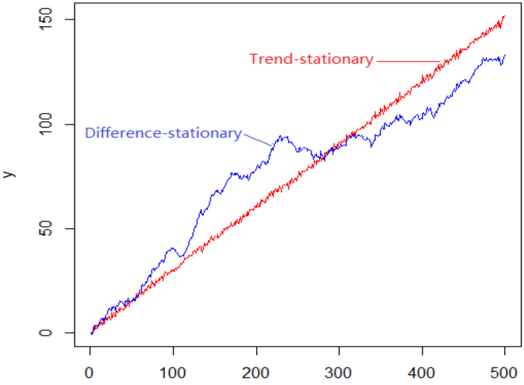
- With both drift and time trend terms

$$y_t = c + \beta t + y_{t-1} + u_t = \ldots = ct + \beta \frac{(t+1)t}{2} + \sum_{i=0}^{t} u_i$$

  whose difference $\triangle y_t = \beta_0 + \beta_1 t + u_t$ has time trend. By subtracting its mean, $\triangle y_t$ will be stationary. So $\triangle y_t$ is trend-stationary.

Graph of DGP

- Trend-stationary: $y_t = 0.5 + 0.3t + u_t$
- Difference-stationary: $y_t = 0.3 + y_{t-1} + u_t$



Test Method for Trend-stationary Process

In empirical research, the first step to distinguish trend-stationary with difference-stationary is **detrending** data's deterministic trend. After that, we can test whether the remainder is stationary. If stationary, there exists no doubt that data series is trend-stationary; otherwise, difference-stationary.

For example, running regression:

$$y_t = \beta_0 + \beta_1 t + u_t$$

regression residual $\hat{u}_t = \hat{y}_t - \hat{\beta}_0 - \hat{\beta}_1 t$ can be got. Through testing whether $\hat{u}_t$ is stationary, it's easy to judge y_t's type.

DF-test

- AR(1) process $y_t = \rho y_{t-1} + u_t$, unit root test:

$$H_0 : \rho = 1; \;\; H_1 : \rho < 1$$

- Difference $\triangle y_t = \delta y_{t-1} + u_t$, unit root test:

$$H_0 : \delta = 0; \;\; H_1 : \delta < 0$$

- When H_0 is satisfied, the asymptotic variance of coefficient-based statistics approaches to 0 and the limit distribution degenerates into non-standard distribution.

DF-test and Its Three Forms

There are also three testing forms corresponding to three unit root forms.

- (1) Without drift or trend term: $y_t = \rho y_{t-1} + u_t$
- (2) With drift term only: $y_t = \beta_0 + \rho y_{t-1} + u_t$
- (3) With drift and trend terms: $y_t = \beta_0 + \beta_1 t + \rho y_{t-1} + u_t$

Limit Distribution of Three DF-test Forms I

- (1) Without drift or trend term in the regression

$$T(\hat{\rho}-1) \xrightarrow{d} \frac{\int_0^1 W(r)dW}{\int_0^1 [W(r)]^2 dr} = \frac{(W(1)^2-1)/2}{\int_0^1 [W(r)]^2 dr}$$

Test statistics

$$\tau = \frac{\hat{\rho}-1}{se(\hat{\rho})} \xrightarrow{d} \frac{\int_0^1 W(r)dW}{[\int_0^1 [W(r)]^2 dr]^{1/2}}$$

Proof I

Definition

Standard Brownian Motion $W(\cdot)$ is a continuous-time stochastic process, associating each date $t \in [0, 1]$ with the scalar $W(t)$ such that:

(1) $W(0) = 0$;

(2) For any dates $0 \leq t_1 < t_2 < \ldots < t_k \leq 1$, the changes $W(t_2)-W(t_1), W(t_3)-W(t_2), \ldots, W(t_k)-W(t_{k-1})$ are independent multivariate Gaussian with $W(s) - W(t) \sim N(0, s - t)$;

(3) For any given realization, $W(t)$ is continuous in t with probability 1.

Then, $W(1) \sim N(0, 1)$.

Proof II

If $u_t \sim i.i.d.(0, \sigma^2)$, we have $\frac{1}{\sqrt{T}} \sum_{t=1}^{T} u_t \xrightarrow{d} N(0, \sigma^2)$.

Theorem (Functional Central Limit Theorem)

$X_T(r)$ is a variable constructed from the sample mean of the first rth fraction of observations, $r \in 0, 1$, defined by

$$X_T(r) := 1/T \sum_{t=1}^{|Tr|^*} u_t = \begin{cases} 0 & \text{for } 0 \le r < 1/T \\ u_1/T & \text{for } 1/T \le r < 2/T \\ (u_1 + u_2)/T & \text{for } 2/T \le r < 3/T \\ \vdots & \\ (u_1 + u_2 + \cdots + u_T)/T & \text{for } r = 1 \end{cases}$$

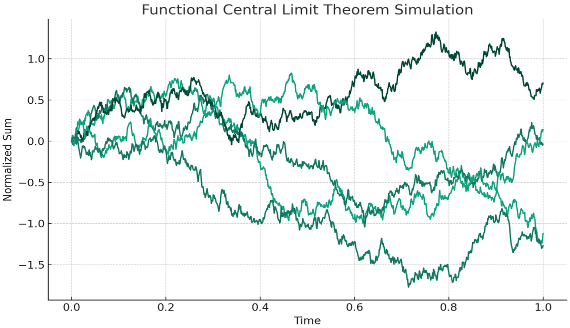
we have $\sqrt{T} \cdot X_T(\cdot)/\sigma \xrightarrow{d} W(\cdot)$.

Code Snippet - Simulation of FCLT

```python
import numpy as np
import matplotlib.pyplot as plt

def simulate_FCLT(num_paths=5, num_steps=1000):
    # Time grid
    t = np.linspace(0, 1, num_steps)

    for _ in range(num_paths):
        # Generate i.i.d. standard normal random variables
        increments = np.random.normal(0, 1, num_steps)

        # Calculate normalized partial sums
        partial_sums = np.cumsum(increments) / np.sqrt(num_steps)

        # Plot the path
        plt.plot(t, partial_sums)

    plt.title("Functional Central Limit Theorem Simulation")
    plt.xlabel("Time")
    plt.ylabel("Normalized Sum")
    plt.grid(True)
    plt.show()

# Simulate and plot the FCLT
simulate_FCLT()
```

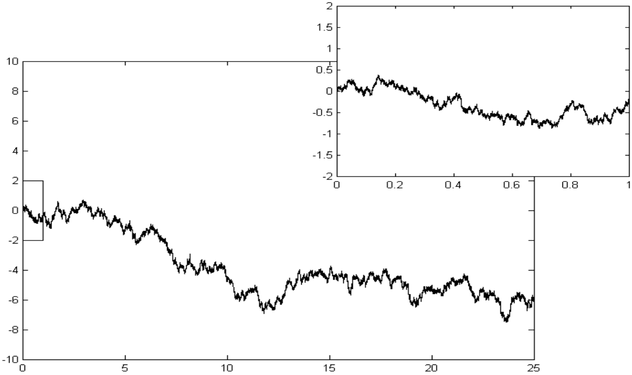
Simulation of FCLT

The codes above simply simulated five different realizations of the process $\sqrt{T} \cdot X_T(\cdot)/\sigma \xrightarrow{d} W(\cdot)$. The results are displayed below. It's easily to see that these paths are very similar to Winer process, providing visual evidence for functional central limit theorem.



Simulation of FCLT

The following picture is a standard Winer process downloaded from Wikipedia. It's easy to see that this process is very similar to above paths.



Proof III

Proof of FCLT:

$$\sqrt{T} \cdot X_T(r) = \frac{\sqrt{T}}{T} \sum_{t=1}^{|Tr|^*} u_t = \frac{\sqrt{|Tr|^*}}{\sqrt{T}} \cdot \frac{1}{\sqrt{|Tr|^*}} \sum_{t=1}^{|Tr|^*} u_t$$

Obviously, $(1/\sqrt{|Tr|^*}) \sum_{t=1}^{|Tr|^*} u_t \xrightarrow{d} N(0, \sigma^2)$ and $\sqrt{|Tr|^*}/\sqrt{T} \xrightarrow{d} \sqrt{r}$. So,

$$\sqrt{T} \cdot X_T(r) \xrightarrow{d} N(0, r\sigma^2)$$

namely, $\sqrt{T} \cdot X_T(r)/\sigma \xrightarrow{d} N(0, r)$, $\sqrt{T} \cdot X_T(\cdot)/\sigma \xrightarrow{d} W(\cdot)$.

Specifically, when $r = 1$, the function $X_T(r)$ is just the sample mean $X_T(1) = (1/T) \sum_{t=1}^{T} u_t$ and

$$\sqrt{T} X_T(1) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} u_t \xrightarrow{d} N(0, \sigma^2) \sim \sigma W(1)$$

Proof IV

Theorem (Continuous Mapping Theorem)

If $S_T(\cdot) \xrightarrow{d} S(\cdot)$ and $g(\cdot)$ is a continuous functional, then $g(S_T(\cdot)) \xrightarrow{d} g(S(\cdot))$.

For example, consider the function $S_T(\cdot)$ whose value at r is given by

$$S_T(r) = (\sqrt{T} X_T(r))^2$$

Since $\sqrt{T} \cdot X_T(\cdot) \xrightarrow{d} \sigma W(\cdot)$, we have

$$S_T(\cdot) \xrightarrow{d} \sigma^2 (W(\cdot))^2$$

Proof V

Under $\mathrm{H}_0 : \rho = 1$, consider the test statistics τ based on AR(1) regression:

$$\tau = \frac{\hat{\rho} - 1}{se(\hat{\rho})}$$

due to:

$$T(\hat{\rho} - 1) = \frac{T^{-1} \sum_{t=1}^{T} y_{t-1} u_t}{T^{-2} \sum_{t=1}^{T} y_{t-1}^2}$$

and

$$[T \cdot se(\hat{\rho})]^2 = \frac{\hat{\sigma}^2}{T^{-2} \sum_{t=1}^{T} y_{t-1}^2}$$

Proof VI

According to functional central limit theorem and continuous mapping theorem, it is easily to obtain $T^{-1}\sum_{t=1}^{T} y_{t-1}u_t \xrightarrow{d} (1/2)\sigma^2[[W(1)]^2-1]$.

Proof:

$$y_t^2 = (y_{t-1}+u_t)^2 = y_{t-1}^2 + 2y_{t-1}u_t + u_i^2$$

$$y_{t-1}u_t = (1/2)\{y_t^2 - y_{t-1}^2 - u_t^2\}$$

$$\sum_{i=1}^{T} y_{i-1}u_i = (1/2)\{y_T^2 - y_0^2\} - (1/2)\sum_{i=1}^{T} u_i^2$$

$$(1/T)\sum_{i=1}^{T} y_{i-1}u_i = (1/2)\cdot(1/T)y_T^2 - (1/2)\cdot(1/T)\sum_{i=1}^{T} u_i^2$$

Proof VII

This can be written

$$T^{-1}\sum_{i=1}^{T} y_{i-1}u_i = (1/2)S_T(1) - (1/2)(1/T)\sum_{i=1}^{T} u_i^2$$

Since $(1/T)\Sigma_{i=1}^{T} u_i^2 \rightarrow \sigma^2$. and $S_T(1) \xrightarrow{L} \sigma^2[W(1)]^2$, the proof is closed.
And $T^{-2}\sum_{t=1}^{T} y_{t-1}^2 \xrightarrow{d} \sigma^2 \cdot \int_0^1 [W(r)]^2 dr$

Proof VIII

We have defined $S_T(r) = T \cdot (X_T(r))^2$, then

$$S_T(r) = 1/T\Big(\sum_{t=1}^{|Tr|^*} u_t\Big)^2 = \begin{cases} 0 & \textit{for } 0 \le r < 1/T \\ y_1^2/T & \textit{for } 1/T \le r < 2/T \\ y_2^2/T & \textit{for } 2/T \le r < 3/T \\ \vdots & \\ y_T^2/T & \textit{for } r = 1 \end{cases}$$

$\int_0^1 S_T(r)dr = y_1^2/T^2 + y_2^2/T^2 + \cdots + y_{T-1}^2/T^2 = T^{-2}\sum_{t=1}^T y_{t-1}^2$

Since $S_T(r) \equiv [\sqrt{T} \cdot X_T(r)]^2$m we have $S_T(r) \xrightarrow{d} \sigma^2(W(r))^2$, so

$$T^{-2}\sum_{t=1}^{T} y_{t-1}^2 \xrightarrow{d} \sigma^2 \int_0^1 (W(r))^2 dr$$

Proof IX

So, $\hat{\rho}$ is characterized by

$$T(\hat{\rho}-1)=\frac{T^{-1}\sum_{t=1}^{T}y_{t-1}u_t}{T^{-2}\sum_{t=1}^{T}y_{t-1}^2}\xrightarrow{d}\frac{\int_0^1 W(r)dW}{\int_0^1[W(r)]^2dr}$$

On the other hand, $\hat{\sigma}^2=\frac{\sum_{t=1}^{T}\hat{u}_t^2}{T-1}\xrightarrow{P}\sigma^2$, and

$$[T\cdot se(\hat{\rho})]^2=\frac{\hat{\sigma}^2}{T^{-2}\sum_{t=1}^{T}y_{t-1}^2}\xrightarrow{d}\frac{1}{\int_0^1[W(r)]^2dr}$$

Totally there will be

$$\tau=\frac{T(\hat{\rho}-1)}{\sqrt{[T\cdot se(\hat{\rho})]^2}}=\frac{T^{-1}\sum_{t=1}^{T}y_{t-1}u_t}{[T^{-2}\sum_{t=1}^{T}y_{t-1}^2]^{1/2}[\hat{\sigma}^2]^{1/2}}\xrightarrow{d}\frac{\int_0^1 W(r)dW}{[\int_0^1[W(r)]^2dr]^{1/2}}$$

Proof X

Theorem (Convergence Theorem)

Suppose that y_t follows a random walk without drift $y_t = y_{t-1} + u_t$, where $y_0 = 0$, $u_t \sim IID(0, \sigma^2)$, then

(1) $T^{-1/2} \sum_{t=1}^{T} u_t \xrightarrow{d} \sigma \cdot W(1)$

(2) $T^{-1} \sum_{t=1}^{T} y_{t-1} u_t \xrightarrow{d} \frac{1}{2}\sigma^2 \{[W(1)]^2 - 1\}$

(3) $T^{-3/2} \sum_{t=1}^{T} t u_t \xrightarrow{d} \sigma \{W(1) - \int_0^1 W(r) dr\}$

(4) $T^{-3/2} \sum_{t=1}^{T} y_{t-1} \xrightarrow{d} \sigma \cdot \int_0^1 W(r) dr$

(5) $T^{-2} \sum_{t=1}^{T} y_{t-1}^2 \xrightarrow{d} \sigma^2 \cdot \int_0^1 [W(r)]^2 dr$

(6) $T^{-5/2} \sum_{t=1}^{T} t y_{t-1} \xrightarrow{d} \sigma \cdot \int_0^1 r W(r) dr$

(7) $T^{-3} \sum_{t=1}^{T} t y_{t-1}^2 \xrightarrow{d} \sigma^2 \cdot \int_0^1 r [W(r)]^2 dr$

(8) $T^{-(v+1)} \sum_{t=1}^{T} t^v \xrightarrow{d} 1/(v+1)$ *for* $v = 0, 1, \ldots$.

Limit Distribution of Three DF-test Forms II

- (2) With only drift term in the regression

$$T(\hat{\rho}-1) \xrightarrow{d} \frac{\frac{1}{2}(W(1)^2-1) - W(1)\int_0^1 W(r)dr}{\int_0^1 [W(r)]^2 dr - (\int_0^1 W(r)dr)^2}$$

Test statistics

$$\tau = \frac{T(\hat{\rho}-1)}{(T^2 \cdot se(\hat{\rho})^2)^{1/2}} \xrightarrow{d} \frac{\frac{1}{2}(W(1)^2-1) - W(1)\int_0^1 W(r)dr}{[\int_0^1 [W(r)]^2 dr - (\int_0^1 W(r)dr)^2]^{1/2}}$$

Proof I

Form (2) can be expressed in matrix notation as

$$y = X\beta + u = (1 \ y_{t-1}) \begin{pmatrix} \alpha \\ \rho \end{pmatrix} + u_t$$

then

$$\begin{bmatrix} \hat{\alpha} - \alpha \\ \hat{\rho} - \rho \end{bmatrix} = (X^{'}X)^{-1}X^{'}u$$

$$= \frac{1}{\sum y_{t-1}^2 \cdot \sum 1 - (\sum y_{t-1})^2} \begin{bmatrix} \sum y_{t-1}^2 & -\sum y_{t-1} \\ -\sum y_{t-1} & \sum 1 \end{bmatrix} \cdot \begin{bmatrix} \sum u_t \\ \sum y_{t-1} u_t \end{bmatrix}$$

Proof II

So,

$$T(\hat{\rho}-1)=\frac{T\cdot T^{-2}\sum y_{t-1}u_t - T^{-3/2}\sum y_{t-1}\cdot T^{-1/2}\sum u_t}{T\cdot T^{-3}\sum y_{t-1}^2-(T^{-3/2}\sum y_{t-1})^2}$$

On the other hand, $\hat{\sigma}^2=\frac{\sum_{t=1}^{T}\hat{u}_t^2}{T-1}\xrightarrow{P}\sigma^2$, and

$$[T\cdot se(\hat{\rho})]^2=\frac{T^2\hat{\sigma}^2\sum 1}{\sum y_{t-1}^2\cdot\sum 1-(\sum y_{t-1})^2}$$

Totally there will be

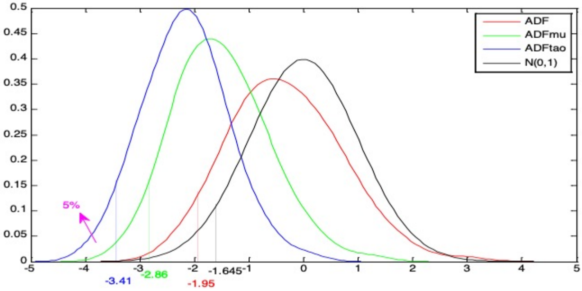
$$\tau=\frac{T(\hat{\rho}-1)}{\sqrt{[T\cdot se(\hat{\rho})]^2}}\xrightarrow{d}\frac{\frac{1}{2}(W(1)^2-1)-W(1)\int_0^1 W(r)dr}{[\int_0^1[W(r)]^2dr-(\int_0^1 W(r)dr)^2]^{1/2}}$$

Limit Distribution of Three DF-test Forms III

- (3) With drift and trend terms in the regression

$$T(\hat{\rho} - 1) \xrightarrow{d} \text{Non-standard Distribution}$$

Simulation of Limit Distribution and Critical Value



Red line: $\Delta y_t = \theta y_{t-1} + u_t$

Green line: $\Delta y_t = \alpha + \theta y_{t-1} + u_t$

Blue line: $\Delta y_t = \alpha + \beta_t \theta y_{t-1} + u_t$

Black line: standard normal distribution

ADF-test and Its Three Forms

When the error terms have autocorrelations, DF-test is unsuitable and needs to be revised. By expanding testing forms, we can get ADF-test (Augmented Dickey-Fuller).

Three forms of ADF-test corresponding to three unit root forms:

- (1) Without drift or trend term in the regression
  $$\triangle y_t = \delta y_{t-1} + \sum_{i=1}^{p} \phi_i \cdot \triangle y_{t-i} + u_t$$
- (2) With only drift term in the regression
  $$\triangle y_t = \beta_0 + \delta y_{t-1} + \sum_{i=1}^{p} \phi_i \cdot \triangle y_{t-i} + u_t$$
- (3) With both drift and trend terms in the regression
  $$\triangle y_t = \beta_0 + \beta_1 t + \delta y_{t-1} + \sum_{i=1}^{p} \phi_i \cdot \triangle y_{t-i} + u_t$$

Two Targets of the Test

- Size: Rejecting H0 under H0, that is, making the type III error. In unit root test, it is the probability of the statistics is less than the critical value, α, when the real data generating process(DGP) is a unit root process.
- Power: The probability of rejecting H0 under H1, $1-\beta$. In unit root test, it is the probability of the statistics is less than the critical value when the true DGP is "(trend-)stationary".

the disadvantage of ADF

- By applying rules to determine lag terms: AIC or BIC or ...
- The more Lag Terms included, the less autocorrelations the residual will have
- The more Lag Terms included, the more Degrees of Freedom it will lose

Idea of PP-test

- PP-test has three forms as well. However, regression with only drift term is considered in most cases.

$$y_t = \beta_0 + \rho y_{t-1} + u_t$$

- $H_0 : \rho = 1; \;\; H_1 : \rho < 1.$
- Phillips and Perron regard u_t as a error term which has complicated structures even including both autocorrelation and heteroscedasticity.
- So the key point lies in how to get a consistent and robust estimator.

Limit Distribution of PP-test Statistics

Test statistics

$$T(\hat{\rho}-1)-\frac{1}{2}[\frac{T^2 \cdot se(\hat{\rho})^2}{\hat{\sigma}^2}](\hat{\lambda}^2-\hat{\sigma}^2) \xrightarrow{d} \frac{\frac{1}{2}(W(1)^2-1)-W(1)\int_0^1 W(r)dr}{\int_0^1 [W(r)]^2 dr - (\int_0^1 W(r)dr)^2}$$

Where $\hat{\sigma}^2 = (T-2)^{-1}\sum_{t=1}^{T}(y_t - \hat{\beta}_0 - \hat{\rho}y_{t-1})^2 = \gamma_0$ denotes a consistent OLS estimator of σ^2 (error variance), and $\hat{\lambda}^2$ denotes a consistent N-W estimator of λ^2 (long-run variance).

Long-run Variance

- Long-run Variance

$$\lambda^2 = \sum_{j=-\infty}^{\infty} \gamma_j = \sum_{j=-\infty}^{\infty} Cov(u_t u_{t-j}) = 2\sum_{j=1}^{\infty} \gamma_j + \gamma_0$$

- The difference between σ^2 and λ^2

$$\sigma^2 = \lim_{T\to\infty} \frac{\sum_{t=1}^{T} E(u_t^2)}{T}$$

$$\lambda^2 = \lim_{T\to\infty} E[\frac{(\sum_{t=1}^{T} u_t)^2}{T}]$$

A Consistent Long-run Variance Estimator

- Newey & West use kernel function to estimate long-run variance

$$\hat{\lambda}^2 = \hat{\gamma}_0^2 + 2\sum_{j=1}^{q} \kappa_j \hat{\gamma}_j^2, \quad \hat{\gamma}_j^2 = T^{-1} \sum_{t=j+1}^{T} \hat{u}_t \hat{u}_{t-j}$$

Where kernel function is $\kappa_j = 1 - j/(q+1)$.

The idea of this estimator is to assume that the autocovariance (or correlation coefficient) of the error term after the q-th order is zero, so the long-run variance only needs to weight the average of the autocovariance in the previous q-order autocorrelation.

Comparison of DF(ADF)-test and PP-test

ADF test: $\tau = \frac{T(\hat{\rho}-1)}{\sqrt{[T\cdot se(\hat{\rho})]^2}} = \frac{T^{-1}\sum_{t=1}^{T} y_{t-1}u_t}{[T^{-2}\sum_{t=1}^{T} y_{t-1}^2]^{1/2}\hat{\sigma}^2]^{1/2}} \xrightarrow{d} \frac{\int_0^1 W(r)dW}{[\int_0^1 [W(r)]^2 dr]^{1/2}}$

PP test: $T(\hat{\rho}-1)-\frac{1}{2}[\frac{T^2\cdot se(\hat{\rho})^2}{\hat{\sigma}^2}](\hat{\lambda}^2-\hat{\sigma}^2) \xrightarrow{d} \frac{\frac{1}{2}(W(1)^2-1)-W(1)\int_0^1 W(r)dr}{\int_0^1 [W(r)]^2 dr-(\int_0^1 W(r)dr)^2}$

Long-run variance:

$$\lambda^2 = \sum_{j=-\infty}^{\infty} \gamma_j = \sum_{j=-\infty}^{\infty} Cov(u_t u_{t-j}) = 2\sum_{j=1}^{\infty} \gamma_j + \gamma_0$$

- These tests have the same limit distribution and critical value.
- ADF-test and DF-test have the same testing statistics but different testing models.
- DF-test and PP-test have the same testing model but different testing statistics.
- DF-test and ADF-test have better finite sample properties; PP-test has better large sample properties.

KPSS-test

- More often, KPSS(Kwiatkowski-Phillips-Schmidt-Shin) is used to test trend-stationary.
- Unlike the previous tests, KPSS-test sets no unit root case as its null hypothesis.
- The robust estimator of error variance in KPSS-test is also obtained by Newey-West's nonparametric methods.
- It is necessary to do KPSS-test when the data movement conforms to linear trend.
- Finally, KPSS-test is a right-tailed test.

Two Forms of KPSS-test

- (1) Without time trend term

$$y_t = \beta_0 + \mu_t + u_t, u_t \sim I(0)$$

$$\mu_t = \mu_{t-1} + \varepsilon_t, \varepsilon_t \sim IID(0, \sigma_\varepsilon^2)$$

- (2) With time trend term (common used)

$$y_t = \beta_0 + \beta_1 t + \mu_t + u_t, u_t \sim I(0)$$

$$\mu_t = \mu_{t-1} + \varepsilon_t, \varepsilon_t \sim IID(0, \sigma_\varepsilon^2)$$

- Hypothesis

$$H_0 : \sigma_\varepsilon^2 = 0 \implies y_t \sim I(0)$$

$$H_1 : \sigma_\varepsilon^2 > 0 \implies y_t \sim I(1)$$

Limit Distribution of KPSS-test Statistics I

Test statistics

$$\text{KPSS} = \frac{(\sum_{t=1}^{T} E_t^2)/T}{T\hat{\lambda}^2}$$

Where, $E_t = \sum_{i=1}^{t} \hat{u}_i, \quad t = 1, 2, \ldots, T$

$$\hat{\lambda}^2 = \hat{\lambda}_0^2 + 2\sum_{j=1}^{q} \kappa_j \hat{\lambda}_j^2$$

$$\hat{\lambda}_j^2 = T^{-1}\sum_{t=j+1}^{T} \hat{u}_t \hat{u}_{t-j}$$

$$\kappa_j = 1 - j/(q+1).$$

Limit Distribution of KPSS-test Statistics II

- Case I: without time trend term

$$\text{KPSS} \xrightarrow{d} \int_0^1 [W(r) - rW(1)]dr$$

- Case II: with time trend term

$$\text{KPSS} \xrightarrow{d} \int_0^1 [W(r) + r(2-3r)W(1) + 6r(r^2-1)\int_0^1 W(s)ds]dr$$

A Demo to Perform Unit Root Tests

We performed aforementioned unit root tests, namely ADF, PP, and KPSS tests, on CPI, PPI, and their first difference series. The results are reported in the following table.

	ADF	PP	KPSS
CPI	-3.26*	-2.54	0.18*
PPI	-1.66	-1.99	0.11
ΔCPI	-7.31***	-4.64***	0.03
ΔPPI	-5.57***	-8.40***	0.10

* The critical value of ADF and PP tests with trend and drift are -4.04, -3.45, -3.15 under the significance level of 1%, 5%, 10%. The critical value of KPSS test are 0.216, 0.146, 0.119 under the significance level of 1%, 5%, 10%.

Our analysis involved various tests, each yielding distinct results. We collectively opted for the type incorporating both drift and trend.

A Demo to Perform Unit Root Tests

- Concerning the CPI and PPI, the ADF and PP test indicated a failure to reject the null hypothesis, suggesting that these series are unit root processes. However, the two test results for the first-difference series of CPI and PPI allowed us to reject the null hypothesis.
- For both CPI and PPI, the KPSS test results contrasted with the ADF and PP tests. The KPSS test could only reject the null hypothesis at a 10% significance level of CPI series, but it suggested that PPI are stationary processes. Additionally, the KPSS test results for the first-difference series also failed to reject the null hypothesis, i.e. the series is stationary.

Structural Breaks

Supposing that the data exits a structural change, the DGP is:

$$y_t = \alpha_1 + \alpha_2 DU_t + \delta_1 t + \delta_2 D(TB)_t t + \gamma y_{t-1} + \varepsilon_t$$

where $DU_t = \begin{cases} 1 & t > T_{1,B} \\ 0 & otherwise \end{cases}$, $D(TB)_t = \begin{cases} 1 & t > T_{2,B} \\ 0 & otherwise \end{cases}$.
Using the third form of ADF-test to test whether this time series is stationary or not, we can obtain that size=0.0477, power=0.7272.

Perron Phenomenon (Perron (1989)):
The existence of structural breaks which has a permanent effect will lead to power loss, that is a non rejection of the unit root hypothesis even though it is not true.

Structural Breaks

Related literature:

- Perron and Rodriguez (2003) extended the GLS-detrending M-tests to the case where there is a changed in the trend function.
- Perron and Zhu (2005) analyzed the consistency and rate of convergence of the estimator of the break fraction by minimizing the total sum of squared residuals.
- Harris et al. (2009) proofed the unit root test based on the estimator of the unknown break fraction has an asymptotic null distribution that is the same as if the break fraction was known.
- Rodrigues (2013) showed that recursive adjustment of a deterministic component enhanced effectively the power of unit root tests in the structural break context

Structural Breaks

Estimation of the breakpoint:

- Assuming a known breakpoint, $T_B = \lambda T = T_b^0$.
- The break is assumed to be present but unknown. It can be estimated endogenously in two ways:
  - the estimated breakpoint is selected as the value that minimizes the null unit root test statistic over some breakpoint range. Then:

$$\tau(\lambda^{(1)}) = \inf_{\lambda \in \Gamma} \tau(\lambda)$$

    where Γ is a closed subset of $(0, 1)$.
  - the breakpoint is estimated based on the minimum sum of squared residuals (SSR) obtained from a regression on a set of deterministic variables:

$$\lambda^{(2)} = \arg\min_{\lambda \in \Gamma} SSR_i = \arg\min_{\lambda \in \Gamma} \sum_{t=1}^{T} \tilde{y}_{it}^2$$

.

Time-varying Variance

Assumption

Let $\varepsilon_t := C(L)\sigma_t \upsilon_t$, the lag polynomial satisfies $C(z) \neq 0$ for all $|z| \leq 1$ and $\Sigma_{i=0}^{\infty} i|C_i| < \infty$, and where $\upsilon_t \sim IID(0,1)$ with $E|\upsilon_t|^r < K < \infty$ for some $r \geq 4$. The volatility term $\sigma_t = \omega(t/T)$, where $\omega(\cdot)$ is non-stochastic and strictly positive. For $t < 0$, $\sigma_t \leq \sigma^\star < \infty$.

This covers an extremely wide class about volatility process, such as a single abrupt change, multiple volatility shifts, polynomial trending volatility and smooth-transition variance breaks and so on.

Time-varying Variance

The related literature:

- Cavaliere and Taylor. (2008) showed that under time-varying variance, the variation of variance in the sample being tested is perfectly replicated by using wild bootstrap.
- Cavaliere et al. (2011) focused on the impact of nonstationary volatility on the break fraction estimator and associated trend break unit root tests of Harris et al. (2009).
- Smeekes and Taylor (2012) designed the OLS- and GLS-detrended wild bootstrap DF unit root tests under time-varying variance, which is robust over uncertainty about the form of the deterministic trend.

Time-varying Variance

A quantity that will play a key role in what follows is given by the following function in $\mathscr{C}$, known as the variance profile of the process:

$$\eta(s) := (\int_0^1 \omega(r)^2 dr)^{-1} \int_0^s \omega(r)^2 dr$$

Defining $\overline{\omega}^2 := \int_0^1 \omega(r)^2 dr$ can be interpreted as the asymptotic average innovation variance, and the variance-transformed Brownian motion $W^\eta(s) := W(\eta(s)) = \int_0^s dW(\eta(s))$ where $W(\cdot)$ is a standard Brownian motion on [0,1], that is, a Brownian motion under a modification of the time domain. Following Cavaliere and Taylor (2007), we have

$$T^{-1/2} \sum_{t=1}^{\lfloor sT \rfloor} \varepsilon_t \xrightarrow{d} \overline{\omega} C(1) W^\eta(s)$$

Wild Bootstrap

The following steps constitute our proposed bootstrap algorithm:

- Generate the bootstrap residuals $\varepsilon_t^b = \zeta_t \hat{\varepsilon}_t$, where $\hat{\varepsilon}_t$ is obtained by differencing the partial GLS detrended data, that is, $\hat{\varepsilon}_t := y_t^\star - y_{t-1}^\star$. And ζ_t is an independent N(0,1) sequence, that satisfies $E(\zeta_t) = 0$ and $E(\zeta_t^2) = 1$. We choose ζ_t to be the two-point distribution: $P(\zeta_t = 1) = P(\zeta_t = -1) = 0.5$.
- The bootstrap sample is then obtained by constructing the partial sum process $y_t^b = z_t \gamma + \sum_{j=1}^{t} \varepsilon_j^b$.
- Based on y_t^b, construct the wild bootstrap unit root test statistics, τ^B.
- Calculate the bootstrap p-value: $p^b = 1 - G^b(\tau)$, where $G^b(\cdot)$ denotes the cumulative distribution function of τ^B, and τ is the statistic based on the original data, y_t.

Dual Test

- Generally, real economic variables are either stationary processes or unit root processes. It's a conventionality that unit root test refers to just the left-tailed test.

$$H_0: \ \rho = 1$$
$$H_1: \ \rho < 1$$

- When $\rho > 1$ for AR(1), $\{y_t\}$ is defined as Explosive Process.
- Price series with bubbles is usually an explosive process.
- Unlike unit root, explosive test is essentially a right-tailed test.

$$H_0: \ \rho = 1$$
$$H_1: \ \rho > 1$$

Moderate Deviations Process

Phillips (2007) tried to use Moderate Deviations Process to approach explosive process:

$$y_t = \rho y_{t-1} + u_t, \qquad t = 1, \ldots, T$$

$$\rho = \rho_T = 1 + \frac{c}{K_T}, \; K_T = o(T)$$

- The test statistics will obey Standard Cauchy Distribution:

$$\frac{\rho^n}{\rho^2 - 1}(\hat{\rho} - \rho) \xrightarrow{L} C(0, 1)$$

- u_t is a sequence of independent and identically distributed random variables with $Eu_t = 0$ and $Eu_t^2 = \sigma^2 < \infty$.

Bubble Test

Testing for the presence of explosive behavior in periodically collapsing bubble process:

- Evans (1991) argued right-tailed unit root tests have little power to detect periodically collapsing bubble.
- Phillips, Wu, and Yu (2011) proposed a sup augmented Dickey-Fuller (SADF) statistic to test for the presence of explosive behavior in periodically collapsing bubble process.

$$SADF(r_0) = \sup_{r_2 \in [r_0, 1]} ADF_0^{r_2}$$

- Homm and Breitung (2012) used the supremum of a set of backward recursive Chow test.

Bubble Test

- Phillips et al. (2015a, b, PSY) extended the test to the case where there may be multiple bubbles in the data and compared the estimation of the bubble origination and termination dates among the test procedures mentioned above.

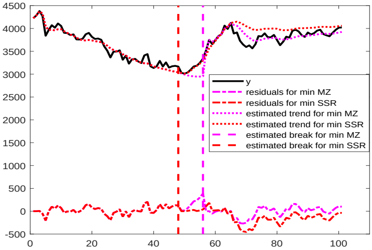
$$GSADF(r_0) = \sup_{r_2 \in [r_0, 1], r_1 \in [0, r_2]} ADF_{r_1}^{r_2}$$

- Harvey et al. (2016) introduced the non-stationary volatility in the context of the PWY test and proposed a wild bootstrap SADF test.

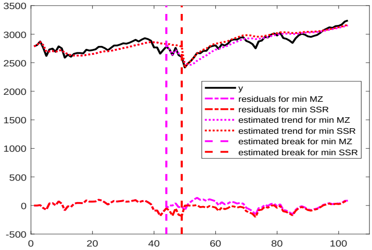
The Sino-US Trade Conflict in 2018

- In 2018, the trade conflict between China and the US is undoubtedly one of the most concerned events in the world economy.
- In this section, we apply the proposed test to the stock markets in China and the US. We analyze the trend of price movements represented by HS300 and S&P 500 respectively.
- The weekly data from January 2018 to December 2019 used in this paper are obtained from the Wind Economic Database.

The Sino-US Trade Conflict in 2018



(a) China: HS300 index



(b) US: S&P500 index

The above figure plotted the pattern of HS300 index in China and S&P500 index in the US which used the black rule, respectively. The purple dashed rules below denote the residuals of these indices with the objective of minimal MZ, while the red dashed rules denotes with the objective of minimal SSR. The purple and red spots mean our estimated trend with the objective of minimal MZ and SSR

The Sino-US Trade Conflict in 2018

Consider this influence, we specify the model as:

$$y_t = \gamma_1 + \gamma_2 t + \gamma_2 \mathbf{I}_{\{t > T_b^0\}} + \gamma_3 \mathbf{I}_{\{t > T_b^0\}}(t - T_b^0) + u_t, \qquad u_t = \rho u_{t-1} + \varepsilon_t$$

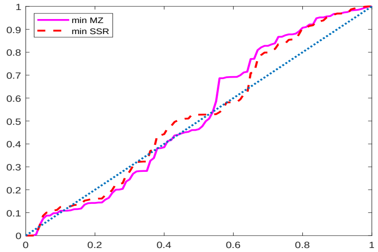
To assess the non-stationary volatility of stock prices, following Cavaliere and Taylor (2007), we estimate the variance profile, $\eta(s)$, by

$$\hat{\eta}(s) = \frac{\sum_{t=1}^{\lfloor sT \rfloor} \hat{\upsilon}_t^2 + (sT - \lfloor sT \rfloor)\hat{\upsilon}_{\lfloor sT \rfloor + 1}^2}{\sum_{t=1}^{T} \hat{\upsilon}_t^2}$$

where $\hat{\upsilon}_t$ is the residual from the ADF regression.

The Sino-US Trade Conflict in 2018

Figure 1 presents the estimated variance profiles. Clearly they show substantial deviations from the 45° line that represents a constant variance process.



(c) China (d) US

Figure: Estimated variance profiles.

The Sino-US Trade Conflict in 2018

- The breakpoints estimated by minimizing the MZ statistic are January 25, 2019 for China and November 02, 2018 for the US, while those estimated by minimizing SSR are November 30, 2018 for China and December 07, 2018 for the US.
- The results based on the minimum of SSR are closer to the day (December 02, 2018) when Sino-US trade negotiations made optimistic progress.
- Because all p-values are larger than 0.05, we conclude that both China and US stock indexes are unit root processes with a structural break.

Thank you !