OVERVIEW OF THE UNIT ROOT TEST

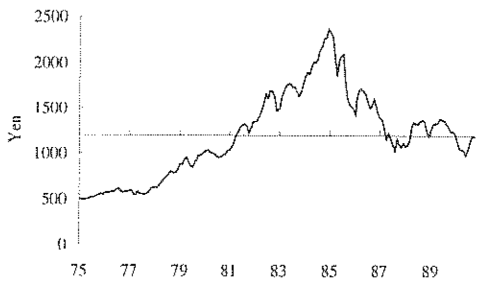
Kimio Morimune, Kyoto University

1 Univariate Series

1.1 Eyeball test

Here, we examine how often the series cross the x-axis. For example, the Nikkei500 series crosses the x-axis only seven times in fifteen years in Figure 1. A series with this property is often considered to be a candidate for a unit root series.

Fig 1. Nikkei500



The first differenced series

When the original series shows the property of small number of crossing, the first difference of the series, $\Delta X_t = X_t - X_{t-1}$, is plotted. As it is shown in Table 2, the first difference crosses the x-axis as often as a normal stationary series does.

Fig 2. The first difference of Nikkei500
$\Delta X(t)=X(t)-X(t-1)$

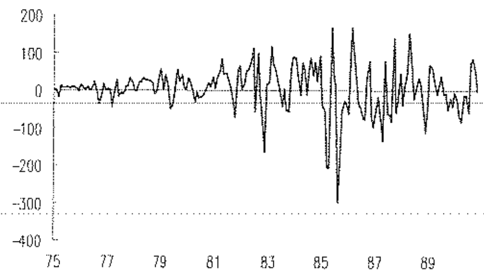


Table 1 AR(3) Estimation of Nikkei500

	Estimated Coeff	Standard Error	t value	P-value
Intercept	8.77	5.584	1.6	0.12
X(-1)	1.37	0.059	23.0	0.00
X(-2)	-0.49	0.097	-5.1	0.00
X(-3)	0.12	0.059	2.1	0.04
R^2	0.99	modified R^2	0.99	
DW statistic	2.02	Durbin's h	-0.16	0.88

Nikkei500 is the dependent variable. The number of observations is 283 from 1972M4 to 1995M10.

1.2 Estimating the Autoregressions

Autoregressive equation is of the form

(1) $$X_t = \alpha + \beta t + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \phi_3 X_{t-3} + \phi_4 X_{t-4} + u_t .$$

Table 1 shows the estimated coefficients and values of the test statistics. In the unit root analysis, the AR equation is transformed so that the test statistic is easily calculated. The lagged variables of order greater than or equal to 2 in the equation 1 are transformed into the difference, and the regression equation is now

(2) $$X_t = \alpha + \beta t + \phi X_{t-1} + \gamma_1 \Delta X_{t-1} + \gamma_2 \Delta X_{t-2} + \gamma_3 \Delta X_{t-3} + u_t \, . \; \phi = \phi_1 + \phi_2 + \phi_3 + \phi_4 \, .$$

The null hypothesis of the unit root is $H_0 : \phi = 1$. This results from the associated polynomial $f(x) = x^{4} - \phi_1 x^{2} - \phi_2 x^{3} - \phi_3 x^{4}$. ϕ is 1 so that one of the zeros is 1. The test statistic for testing $H_0 : \phi = 1$ is the t-ratio of the ϕ coefficient. The same test statistic is calculated by the t-ratio of the ϕ coefficient in the auxiliary regression

(3) $$\Delta X_t = \alpha + \beta t + \rho X_{t-1} + \gamma_1 \Delta X_{t-1} + \gamma_2 \Delta X_{t-2} + \gamma_3 \Delta X_{t-3} + u_t \, , \; \rho = \phi - 1 \, .$$

Table 2 gives estimated values of coefficients. The coefficient of X(-1) is close to 0, and its significance must be tested. The t-ratio is used in the test, however, the null distribution is different from the standard normal distribution. This will be explained later.[1]

Table 2 Transformed AR(3) Estimation of Nikkei500 (ΔX is the dependent variable)

	Estimated Coeff	Standard Error	t value	P-value
Intercept	8.77	5.584	1.6	0.12
X(-1)	-0.01	0.005	-1.3	???
ΔX(-1)	0.37	0.059	6.3	0.00
ΔX(-2)	-0.12	0.059	-2.1	0.04
R^2	0.13	modified R^2	0.12	
DW statistic	2.02	Durbin's h		

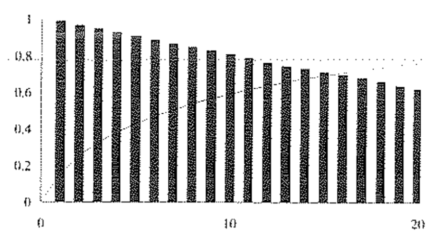
ΔNikkei500 is the dependent variable. The number of observations is 283 from 1972M4 to 1995M10.

1.3 AC of Stationary Series.

In the Box-Jenkins approach, the **sample autocorrelation** (AC) function of a series is always calculated and used to determine the lag order of the moving average process. The sample AC is usually found to be insignificant except for some lower orders. In the Nikkei500 series, the sample AC values gradually diminishes, but the speed of decrease is slow. It requires 24 months before the AC value reaches 1/2. (If the data goes back to 1972, it takes 46 months to reach 1/2.) This slow decrease in the AC function is regarded as a typical characteristic of the unit root series.

Once the first difference of the original series is taken, the sample AC is insignificant (very small in value) after the first order.

Fig 3. Sample auto-correlation (AC) function of the Nikkei500 series



1.4 Random Walk Process

The standard random walk with drift is defined as

(4) $$\Delta X_t = X_t - X_{t-1} = \mu + \varepsilon_t, t = 1, 2, \cdots, T$$

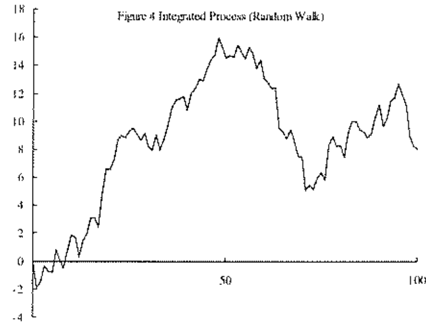
where the error term ε_t is a white noise. The drift μ is a fixed constant. By summing the series from the initial value,

(5) $$X_t = X_0 + \eta_{t-1} + t\mu + \varepsilon_t, \eta_{t-1} = \sum_{i=1,t-1} \varepsilon_i$$

The drift becomes a deterministic trend in the original series. The second term is the sum of white noise. When the drift is not included in (4), (5) consists of the initial value and the standard random walk. The initial value

[1] Critical values are -3.4 (5%) and -4.0 (1%).

stays always in the original series, and setting X_0 and μ to 0, the random walk is characterized by Figure 4. (When the variance of the white noise is set to unity, the white noise series is distributed around 0. The probability of crossing 2 and -2 is five percent.) The random walk series rarely crosses the x-axis. Once the series is away from the x-axis, the probability of crossing x-axis (mean reverting) is zero.



1.5 ARIMA Process

Assume the error term of (4), $\Delta X_t = \mu + u_t$, has the second order autoregressive form such as $u_t = \phi_1 u_{t-1} + \phi_2 u_{t-2} + \phi_3 u_{t-3} + \varepsilon_t$ where ε_t is a white noise. Substituting (4) into the autoregression, it follows an ARIMA (2,1,0) process such as

(6) $$\Delta X_t = \mu + \phi_1 \Delta X_{t-1} + \phi_2 \Delta X_{t-2} + \phi_3 \Delta X_{t-3} + \varepsilon_t .$$

This is called the Cochrane-Orcutt transformation in econometrics. 'I' in ARIMA stands for the integrated process. The order of integration is 1 which is denoted as I(1).

If the error term has a third order autoregression such as $u_t = \phi_1 u_{t-1} + \phi_2 u_{t-2} + \phi_3 u_{t-3} + \varepsilon_t$ and the regression is $\Delta X_t = \mu + u_t$, the regression is transformed as

(7) $$\Delta X_t = \alpha + \beta t + \rho X_{t-1} + \phi_1' \Delta X_{t-1} + \phi_2' \Delta X_{t-2} + \varepsilon_t, \quad \rho = 1 - \phi_1 - \phi_2 - \phi_3$$

where the level variable is X_{t-1}. This regression is the same as (5) when β and ρ are 0. An alternative transformation is

$$\Delta X_t = \alpha + \beta t + \phi_1'' \Delta X_{t-1} + \phi_2'' \Delta X_{t-2} + \rho X_{t-3} + \varepsilon_t, \quad \rho = 1 - \phi_1 - \phi_2 - \phi_3$$

where the level variable is X_{t-3}. From the second transformation, it may be easy to see that the lag order of the ARIMA regression is shorter by one under the null hypothesis.

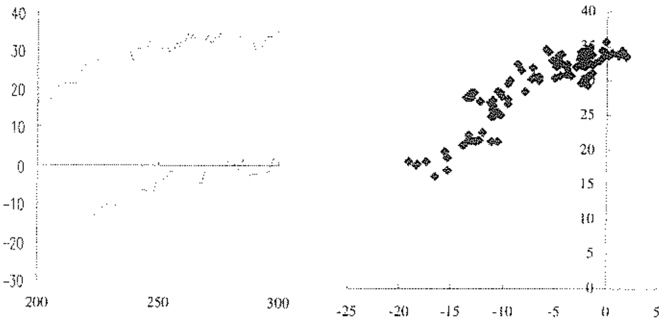
1.6 Spurious Regression

Assume there are two I(1) variables which are formed by independently distributed random variables such as $x_t = \sum_{i=1,t} w_i$ and $y_t = \sum_{i=1,t} v_i$. Since w_t and v_t are independent, x_t and y_t are also independently distributed. However, as Figure 5 shows, the scatter diagram between the two variables has a spurious linear relationship. This linear relationship is called the spurious regression and is caused by the common stochastic trend in the both variables. The linear relationship is estimated as

$$y_t = 34.7\,(95) + 0.83\,(20)\,x_t$$

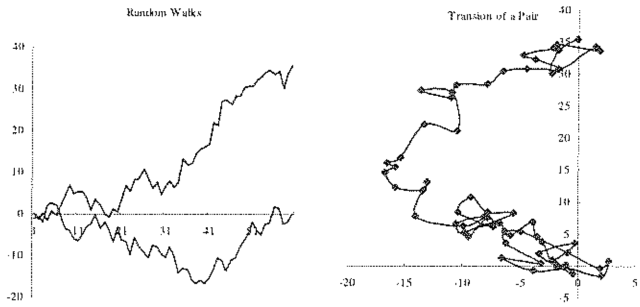
where the values in parentheses are the t-ratios. If we do not know both variables are random walks, this regression could be interpreted as highly significant. It is necessary to test the unit root in each variable to avoid the spurious regressions.

Figure 5. Two uncorrelated Random Walks (left) and the Scattered Diagram (right)



In Figure 5, the random walks are plotted from the 200th to 300th observations. If all the data are used in the plot, the scatter diagram gives a different impression from Figure 5. If it is possible to record data from the beginning of the history, we may not have the problem of spurious regressions, but we usually do not know when the history started.

Fig 6 Transition of two Random Walks



Every Fifth Pair is Plotted.

1.7 Common Factor Approach to the Unit Root Test

The unit root test proposed by Dickey-Fuller (1981) is explained below. The regression under the alternative hypothesis is

(8) $(X_t - \alpha - \beta t) = u_t,\ u_t = \phi u_{t-1} + \varepsilon_t.$

Under the null hypothesis $\phi =$., the regression equation is

(10) $\Delta X_t = \beta + \varepsilon_t.$

The test is performed by estimating the regression equation $(X_t - \alpha - \beta t) = u_t$ by least squares. Writing the residual as $\hat{u}_t$, the test statistic is calculated as the t-ratio of the $\rho = \phi -$. coefficient in

(11) $\Delta \hat{u}_t = \rho \hat{u}_{t-1} + \text{error}$

This t-ratio is denoted τ_τ, and it has been proved that τ_τ weakly converges to a simple function of the de-meaned and de-trended Brownian motion $\tilde{B}_X(r)$. This convergence is written as

(12) $$\tau_\tau \Rightarrow \frac{\int_0^1 \tilde{B}_X(r) d\tilde{B}_X(r)}{\sqrt{\int_0^1 \tilde{B}_X(r)^2 dr}}.$$

The critical value of the test statistic is calculated by simulation. (The 5% point is -3.4, and the 1% point is -4.0.)

1.8 Dickey-Fuller Method

The same test statistic can be calculated as the t-ratio of the ρ coefficient in the equation

(13) $\Delta y_t = \alpha + \beta t + \rho y_{t-1} + \varepsilon_t$

The equation (13) should be taken as an auxiliary regression to calculate the t-ratio by one step. This equation never gives a DGP (Data Generating Process) under the null hypothesis $H_0: \rho = 0$. It requires not only $\rho = 0$ but also $\beta = 0$. Dickey-Fuller based the unit root tests on (13), then they also proposed to test $H_0: \rho = \beta = 0$ by the F-ratio statistic. If (13) is derived from (9) carefully, it is

$$\Delta X_t = \rho X_{t-1} + (\alpha - \alpha\phi + \beta\phi) + \beta(1-\phi)t + \varepsilon_t.$$

Then $\phi = 1$ automatically excludes the deterministic trend.

1.9 Sample AC Approach

As it was explained by Figure 3, the sample AC is an intuitive measure to identify the unit root. In fact, the difference between the sample AC and 1 can be used as the test statistic. This statistic is the same as the least square estimator of ρ in (10). The weak convergence is given as follows and a Table of the critical values is available. This test is easy to use if a software for the regression analyses is capable of calculating the sample AC function of regression residuals.

(14) $$T(AC(1)-1) = T\frac{\Sigma \hat{u}_{t-1}\Delta\hat{u}_t}{\Sigma \hat{u}_{t-1}^2} \left(\Rightarrow \frac{\int_0^1 \tilde{B}_x(r)d\tilde{B}_x(r)}{\int_0^1 \tilde{B}_x(r)^2 dr}\right).$$

1.10 Augmented Dickey-Fuller Test

If the regression is (8) but the error term has a higher order autoregression such as $u_t = \phi_1 u_{t-1} + \phi_2 u_{t-2} + \cdots + \phi_p u_{t-p}$, the regression equation (10) for the test is replaced by

(15) $\Delta\hat{u}_t = \rho\hat{u}_{t-1} + \rho_1\Delta\hat{u}_{t-1} + \cdots + \rho_p\Delta\hat{u}_{t-p} + \text{error}$, $\rho = \phi_1 + \cdots + \phi_p - 1$

or $\Delta\hat{u}_t = \rho\hat{u}_{t-1} + \rho_1\Delta\hat{u}_{t-1} + \cdots + \rho_p\Delta\hat{u}_{t-p} + \text{error}$, equivalently. The test statistic is the t-ratio of ρ in either equation, and the critical values are the same as those of (12). The same test statistic can be calculated from the following auxiliary regression.

(16) $\Delta X_t = \alpha + \beta t + \rho X_{t-1} + \rho_1\Delta X_{t-1} + \cdots + \rho_{p-1}\Delta X_{t-p+1} + \varepsilon_t$, $\rho = \phi_1 + \cdots + \phi_p - 1$

The same test follows from the regression

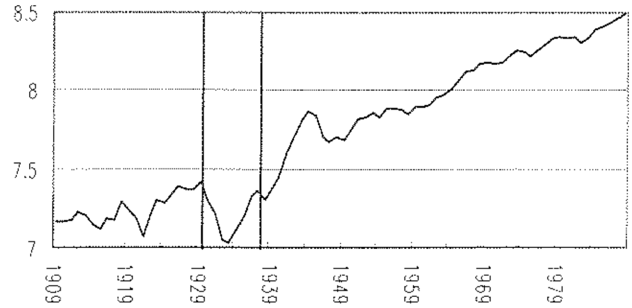
$$\Delta X_t = \alpha + \beta t + \rho_1\Delta X_{t-1} + \cdots + \rho_{p-1}\Delta X_{t-p+1} + \rho X_{t-p} + \varepsilon_t$$

as it was the case in (15). Interpretation of the test from the viewpoint of the sample autocorrelation or the sample partial autocorrelation does not follow in the higher order autoregression.

1.11 Discontinuous Trend Unit Root Test

Many of the US macro series are found to have the unit root by Nelson and Plosser (1982). See the last column in Table 3. Perron and Vogelsang extended the test so that it allows for a discontinuous trend. It was found that many of the US macro series were stationary by applying the Perron test. See the second to the last column in Table 3. The deterministic trend breaks at a particular break point in the sample period. It is naturally difficult to determine the break point but also it was criticized that the break point is chosen after data inspection. Morimune and Nakagawa (1999) analyzed that the misspecified break point causes bias or explosion in the test statistic asymptotically and proposed to use the break interval instead of the break point in applying the Perron test. Many US macro series turned out to be non-stationary again. See the third from the last column in Table 3.

Fig 7: GNP per capita



Break Interval: 30-37.

Table 3:JUMPED BREAK INTERVAL TEST

	Model	Jumped Interval	ρ coefficient	t ratio	Diagnostics	Diagnostic by Perron test	Diagnostic by ADF test
CPI	A	30-42	-0.03	(-1.59)	DS	DS	DS
Common Stock Price	C	30-36	-0.29	(-3.70)	DS	TS	DS
Money Stock	A	30-36	-0.14	(-3.10)	DS	TS	DS
GNP deflator	A	30-36	-0.14	(-3.30)	DS	TS	DS
Real Wages	C	30-36	-0.33	(-3.62)	DS	TS	DS
Interest Rate	A	30-44	0.15	(1.91)	DS	DS	TS
Real GNP per capita	A	30-37	-0.45	(-3.75)	DS	TS	DS
Nominal GNP	A	30-36	-0.26	(-2.38)	DS	TS	DS
Real GNP	A	30-36	-0.50	(-3.79)	TS	TS	DS

DS implies difference stationary, ie., a differenced sereis is stationary.
TS implies trend stationary, ie. a stationary series with trend.
A: Regression with intercept break only.
C: Regression with intercept and trend breaks.
Nelson and Plosser found these US macro series are mostly DS. Perron found them mostly TS by adding a break in t
Most of these series are found DS by using break intervals.

1.12 Over and Under Specification

If the DGP is an ARIMA series but the autoregression is estimated, the coefficient ρ of the additional term is estimated as

(17) $$\Delta y_t = \alpha + \beta t + \rho y_{t-1} + \rho_1 \Delta y_{t-1} + \cdots + \rho_p \Delta y_{t-p} + \varepsilon_t \ , \ \hat{\rho} = \frac{\sum \tilde{X}_{t-1} \Delta \tilde{X}_t}{\sum \tilde{X}_{t-1}^2}$$

$\tilde{X}_{t-1}$ is the de-meaned and de-trended series, and $\hat{\rho}$ is super consistent with 0, then the effect of the additional term in calculating the regressed value is of order $O(1/\sqrt{T})$. If the DGP is the stationary autoregression and the ARIMA regression is estimated, then the estimated coefficients are inconsistent. All classical results remain the same in the unit root analyses.

2 Vector Autoregression

2.1 Cointegration and the Super Consistency

The regression equation is of the form $y_t = \beta x_t + \varepsilon_t$ where ε_t is a white noise with variance σ^2. The regressor x_t is $\Sigma_{i=1,t} v_i$ where v_t is a white noise with variance σ_ε^2 and, for simplicity, v_t and ε_t are independently distributed. This implies both x_t and y_t are random walks but the linear relation $y_t - \beta x_t = \varepsilon_t$ is stationary. This stationary relationship among the integrated variables is called cointegration. Cointegrated relationships produce the consistent least square estimation which is totally different from the spurious regression case. The normalized difference of the least square estimator from the true coefficient is

(18) $$T(\hat{\beta} - \beta) = \frac{T}{\sum_{t=1,n} x_t^2} \sum_{t=1,n} x_t \varepsilon = \frac{1}{\int_0^1 B_x(r)^2 dr} \int_0^1 B_x(r) dB_\varepsilon(r) \ . \ \frac{1}{\sqrt{T}} \sum_{i=1,t} \varepsilon_i \Rightarrow B_\varepsilon(r) \ .$$

Since the speed of convergence in the cointegrated regression is faster than that in the classical regression, the least square estimator is said to be super consistent. Further, the t-ratio of the β coefficient is asymptotically distributed as normal. When v_t and ε_t are contemporaneously correlated, the auxiliary regression

$$y_t = \beta x_t + \gamma \Delta x_t + \varepsilon_t$$

is used in the estimation to avoid a higher order bias. This is called the dynamic regression by Phillips and Loretan (1991).

2.2 Multiple Cointegrations and Identification

Even if the stationary relationship includes more than two integrated variables, super consistency is still valid. However, if the second relation exists but is unknown, then the least square estimator is inconsistent. For example, the following two linear relationships are assumed among the three integrated variables.

(19) $$y_t = \beta x_t + \gamma z_t + u_t ,$$

(20) $$x_t = \delta z_t + w_t .$$

The least square estimator applied to the first equation is degenerate since the two right-hand side variables are cointegrated.

If the second equation is known to exist, all coefficients are identified and can be consistently estimated. This is proved as follows. Cointegration (20) implies

(21) $$x_t - \hat{\delta} z_t = \hat{w}_t \cong w_t$$

where $\hat{\delta}$ is super consistent and w_t is stationary. Using (21) to replace x_t in (19), it follows

(22) $$y_t = \kappa z_t + \beta \hat{w}_t + u_t , \quad \kappa = (\gamma + \beta \hat{\delta})$$

which is a cointegrated relation between y_t and z_t. Then

(23) $$y_t - \hat{\kappa} z_t = \hat{v}_t \cong \beta \hat{w}_t + u_t$$

where $\hat{\kappa}$ is super consistent and $\hat{v}_t$ is stationary. β can be consistently estimated by the least square estimator associated with the stationary regression (23) as long as w_t and u_t are independently distributed. Parameters are identified by using both the integrated and the stationary relations.

2.3 The Error Correction Model (ECM)

There are two relations between two variables. One is cointegration

(24) $$y_t = \beta x_t + u_t \quad u_t = \rho u_{t-1} + \varepsilon_t , \quad |\rho| < 1,$$

where ε_t is a white noise with variance σ^2, and a spurious regression

(25) $$\Delta y_t = \gamma \Delta x_t + v_t ,$$

where v_t is a white noise with variance σ_v^2. These two relations form a vector series as

$$\begin{pmatrix} -\beta & 1 \\ -\gamma & 1 \end{pmatrix} \begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} -\rho\beta & \rho \\ -\gamma & 1 \end{pmatrix} \begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_t \\ v_t \end{pmatrix}.$$

The VAR representation of the form $\mathbf{z}_t = \mathbf{A}_1 \mathbf{z}_{t-1} + \varepsilon_t$ is

(26) $$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \frac{1}{\gamma - \beta} \begin{pmatrix} \gamma - \beta\rho & \rho - 1 \\ \beta\gamma(1-\rho) & \gamma\rho - \beta \end{pmatrix} \begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix} + \frac{1}{\gamma - \beta} \begin{pmatrix} 1 & -1 \\ \gamma & -\beta \end{pmatrix} \begin{pmatrix} \varepsilon_t \\ v_t \end{pmatrix},$$

and the ECM of the form $\mathbf{z}_t = \mathbf{A}_1 \mathbf{z}_{t-1} + \varepsilon_t$ is

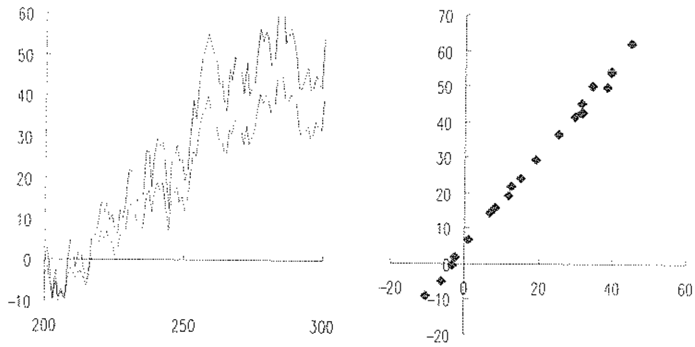
$$(27) \qquad \Delta\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \frac{\rho-1}{\gamma-\beta}\begin{pmatrix} 1 \\ \gamma \end{pmatrix}\begin{pmatrix} -\beta & 1 \end{pmatrix}\begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix} + \frac{1}{\gamma-\beta}\begin{pmatrix} 1 & -1 \\ \gamma & -\beta \end{pmatrix}\begin{pmatrix} \varepsilon_t \\ v_t \end{pmatrix}.$$

2.4 Estimator

The least square estimator is consistent. The least square estimator does not require any knowledge on the rank of cointegration (the number of cointegrations) in the coefficient matrix.

Maximum likelihood estimator was developed by Johansen (1995) which is super consistent. However, the rank of cointegration must be determined first. The Johansen method further requires knowledge on how the constant vector is specified, and how the deterministic trend term is specified in the ECM.

Fig 8: Cointegrated Series and the Scatter Diagram



Equation (27): β=0.8. γ=0.5. ρ=0.5. Error variance =1.

REFERENCE

Banerjee, A., Dolado, J. J., Galbraith, J. W., and Hendry, D. F. (1993), *Co-Integration, Error Correction, and The Econometric Analysis of Non-Stationary Data,* Oxford University Press.

Dickey, D. A. and Fuller, W. A. (1981), "Likelihood ratio statistics for autoregressive time series with a unit root," *Econometrica*, 49, 1057-72.

Engle, R. F., and Granger, C. W. J. (1987), "Co-integration and Error Correction: Representation, Estimation and Testing," *Econometrica,* 55, 251-276.

Hall, P., and Heyde, C. C. (1980), *Martingale Limit Theory and Applications,* Academic Press.

Hatanaka, Michio (1997), *Time series based Econometrics*, Oxford University Press.

Johansen, S. (1995), *Likelihood-based inference in cointegrated vector autoregressive models,* Oxford University Press,

Morimune, K. and Nakagawa, M. (1999), "The discontinuous trend unit root test when the break point is misspecified," *Mathematics and Computers in Simulation*, Vol.48, pp. 417-427, 1999.

Nelson, C. R., and Plosser, C. I. (1982). "Trends and Random Walks in Macro-economic Time Series," *Journal and Monetary Economics*, 27, 3-27

Perron, P., and Vogelsang, T. J. (1993), "Erratum for the great crash, the oil price shock, and the unit root hypothesis," *Econometrica*, 61, 248-249.

Phillips, P. C. B. (1986), "Understanding Spurious Regressions in Econometrics," *Journal of Econometrics*, 33, 311-340

Phillips, P. C. B. and Loretan, M. (1991), "Estimating long-run economic equilibria", *Review of economic studies*, 58,407-436.

Phillips, P. C. B. (1987), "Time Series Regression with a Unit Root," *Econometrica* 55, 277-301.