遠期匯率與即期匯率之隨機共整合分析

Re-examine the spot Exchange Rates and the Forward Exchange Rates by Stochastic Cointegration

研究生：林雅雯 撰
指導教授：李慶男博士

中華民國 九十三 年 七月
謝 辭

依山傍水好中 山，
只 嘆 光 陰 太 短 實；
小 女 不 才 多 貴 人，
心 懷 感 濃 道 不 完。
先 謝 李 教 授 慶 男，
計 量 涵 養 深 如 海；
口 試 委 員 與 院 長，
給 於 提 點 真 寶 貴。
文 亮 若 瑋 幫 大 忙，
同 窗 好 友 相 鼓 勵；
媽 咩 董 閃 亮 亮，
至 親 挚 友 長 相 伴。
大 小 所 秘 好 熱 心，
雅 雯 最 愛 經 濟 所。

08/04/2004
Abstract

There are gradually prosperous trades in foreign exchange markets, agents could hedge, speculate and arbitrage in markets. Market efficiency and whether future spot rates could be predicted by forward rates are worthy of investigate. Hakkio and Rush (1989) demonstrated that cointegration is a necessary condition for market efficiency hypothesis, so that the examination of cointegration to investigate the long-run relationship between the spot rates and forward rates is important. We consider a new method -- stochastic cointegration which contains heteroscedastic and stationary cointegration, to re-examine the relationships between spot and forward rates. The feature of stochastic cointegration is that the cointegrating residuals contain the integrated of order one process and heteroscedastic integrated process. However the special residuals would stochastically trendless over time, so that the spot rates and forward rates has long run equilibrium relationship. Conclusively, the future spot rates empirically are stochastic (and conventional) cointegrated with forward rates in Taiwan, Japan, and Singapore.

Keywords: stochastic cointegration, cointegration, spot, forward, exchange rates, market efficiency.
Contents

1 Introduction ................................................................. 1

2 Theoretical Framework and Literature Reviews ..................... 4
  2.1 Theoretical Framework .............................................. 4
  2.2 Literature Reviews ................................................. 7

3 Approach Analysis .......................................................... 15
  3.1 Unit Roots Tests .................................................. 15
    3.1.1 Dickey-Fuller Test .......................................... 16
    3.1.2 Phillips-Perron Test ......................................... 17
    3.1.3 Augmented Dickey-Fuller Test ............................. 18
  3.2 Conventional Cointegration ....................................... 20
    3.2.1 Engle and Granger Two Steps Method .................... 21
    3.2.2 Johansen Maximum Likelihood Method .................. 23
  3.3 Stochastic cointegration ......................................... 28
    3.3.1 The model .................................................. 29
    3.3.2 The Regression .............................................. 31
    3.3.3 Estimation .................................................. 32
    3.3.4 Hypothesis Tests and Test Statistics .................... 35
Chapter 1

Introduction

Since Fama (1970) provided the concept of market efficiency, the theory is broadly used in various financial assets markets. The original employment in foreign exchange market for market efficiency is addressed by Lenvich (1979), then this subject has been a popular studies from 1980’s till now. Under the foreign exchange market efficiency assumption, agents use all available information rationally, and they can not speculate any excess profits.

From 1970’s, Europe and U.S. successively adopted floating exchange system. However the sequential problem is the exchange rates appear to be more volatile. For the purpose of dealing with the volatility of exchange rates, the forward exchange markets have been established. There are prosperous transactions in liberalized and global economy, hence agents can hedge, speculate and arbitrage from foreign exchange markets. Thus, whether market is efficient rates becomes a hot topic. Hokkio (1989) demonstrated that cointegration is the necessary but not sufficient condition for market efficiency, and the cointegrating relationships between spot and forward exchange rates are interesting for economists.

The financial systems attained to maturity in some developing countries in Asia in which the exchange rates are flexible. Therefore, we consider three countries: Taiwan, Japan, and Singapore which carry out floating exchange rate sys-
tems for this study to examine the relationships between spot and forward rates.

Econometric methods are assistance in lots of analysis to determine linear economic relation, in which OLS is the most common and simplest method, in addition, it’s estimators are great. However, the stringent OLS assumptions leads to hardly conform to real economic problems. In particular, while economic variables are non-stationary, the employment of OLS generates spurious results which are provided by Granger and Newbold (1974).

Meese (1982) empirically demonstrated the spot and forward exchange rates are integrated of order one. And Engle and Granger (1987) addressed two-step cointegration procedure indicating the linear combinations of two $I(1)$ processes would lead to be stationary. Thus two-step cointegration method in the event is frequently employed to examine whether foreign market is efficient. However, the defects of two-step method is single equation, that will cause the cointegrating vector just to be one regardless of the number of variables. Moreover, the estimators is biased in finite sample, and the statistics do not have well-defined asymptotic distribution. so that Engle-Granger cointegrated method is inappropriate. Consequently, two-step method is replaced with Johansen’s maximum likelihood method (1991) for examining market efficiency hypothesis. Johansen procedure (1991) can not only estimate the cointegrating vectors but also test the number of cointegrating vectors. In addition, its multivariate structure would analysis the relationship among more than two variables. But the restraints regarding the residuals as Gaussian white noise process also produce inconvenience.

We wonder is that possible cointegration between exchange rates may present a nonlinear form, instead of conventional linear form. Harris, McCabe and Leybourne (2002) provided stochastic cointegration containing heteroscedastic and stationary cointegration allows the errors of regressions to be the nonlinear combination for $I(0)$ and heteroscedastic integrated process. Besides, HML model
do not stringently restrict residuals to be serially uncorrelated. So that it may be appropriate empirically. Thus, to consider this new method to re-examine the relationships between spot and forward rates.

The concept outlines the theoretical framework and literature reviews in chapter 2. The econometric analysis approaches and stochastic cointegration method are in chapter 3. Chapter 4 illustrates the empirical results for Taiwan, Japan and Singapore. The remaining chapter is conclusion and suggestion. Appendix A describes that the requirement of some asymptotic properties of the AIV estimator. Appendix B illustrates the trend figure of the spot and forward exchange rates for three countries: Taiwan, Japan, and Singapore. Note that in this paper, capital letter for $S_t$ and $F_t$ denote the observed exchange rates, and small letter for $s_t$ and $f_t$ denote the logarithm of exchange rates. Besides, variables in boldface are vectors or matrices, and in lightface are scalar. Currencies in this paper concluding the New Taiwan Dollar, the Japanese Yen and the Singapore Dollar are all to U.S Dollars.
Chapter 2

Theoretical Framework and Literature Reviews

2.1 Theoretical Framework

Fama (1970) defined that in a efficient market, prices at any point in time fully reflect all available information. That is, prices provide accurate signals for resource allocation. Therefore, firms can make production-investment decisions, and investors can choose among the securities that represent ownership of firms’ activities in a efficient market. i.e. none can get excess profits from a efficient market.

On the theory of market efficiency, however, prices do not fully reflect particular subsets of available information in the empirical research. There can be divided three categories depending on the nature of the information subset.

1. Weak form efficient market:
   the information subset is just historical prices.

2. Semi-strong efficient market:
   the information subset includes all obviously publicly available information.

3. Strong efficient market:
   investors or groups have monopolistic access to every information relevant for
price formation have recently appeared.

Since Fama (1970) provided the notions for market efficiency, this subject is generally applied in stock market. Levich (1979) originally employed market efficiency in foreign exchange market, and after that, there are amount of studies for foreign exchange market efficiency till now.

The assumption concerning market efficiency are the agents are risk neutral, the agents use all available information rationally, market is competitive, no information costs , and no transaction costs. Thus the foreign exchange efficient market indicates that the risk premium is zero, current prices should reflect all currently available information, expectations regarding future spot rates should be incorporated and reflected in forward rates, and the expected rate of return to speculation in the forward foreign exchange market conditioned on available information is zero. Thus the current forward rate is an unbiased predictor of the future spot rate, and the simple market efficiency hypothesis can be expressed as:

\[ f_{t,p} = E(s_{t+p}|\Omega_t) \]  

where \( f_{t,p} \) is the logarithm of the \( p \) – period forward rate determined at time \( t \), \( s_{t+p} \) is the logarithm of the future spot rates at time \( t + p \), and \( E(.)|\Omega_t \) is the mathematical expectation conditioned on the information set available to agents at time \( t \). The above equation has implied two assumptions: rational expectation hypothesis\(^1\) and zero risk premium hypothesis.\(^2\) Thus market efficiency contains

---

\(^1\)The market’s expectation is equal to the true expectation using all relevant information \( \Omega_t \):

\[ E(s_{t+k}|\Omega^m_t) = E(s_{t+k}|\Omega_t) \]

\(^2\)The forward rate represents the market’s expectation regarding the future spot rate conditional on the market’s information set \( \Omega^m_t \):

\[ f_{t,p} = E(s_{t+p}|\Omega^m_t) \]

---
these joint assumptions. And, the common form regresses $s_{t+p}$ on $f_{t,p}$:

$$s_{t+p} = \alpha + \beta f_{t,p} + u_{t+p}$$

(2.2)

While market is efficient, the residual $u_{t+p}$ should contain no information and therefore should be serially uncorrelated white noise process. The joint hypothesis for market efficiency, however, may be rejected in the situation of agents are risk adverse, the existence of information costs or transaction costs, then

$$f_{t,p} \neq \mathbb{E}(s_{t+p} | \Omega_t)$$

(2.3)

Cointegration is the necessary but not sufficient condition for market efficiency hypothesis, therefore the examination of cointegrating relationships between spot and forward rates becomes an important work Equation (2.2) is used in this paper for the stochastic cointegration test.
2.2 Literature Reviews

There is an enormous literature on testing whether the forward rate is an unbiased predictor of future spot rates. The earliest studies, e.g. Cornell (1977) and Levich (1979), regressed the log of the future spot rate on the log of the current forward rate. The results generally support the market efficiency hypothesis i.e. forward rate unbiasedness hypothesis (FRUH). Since the unit root property of exchange rates and the concern about the spurious regression were illustrated by Granger and Newbold (1974), the studies concentrated on the regression of the change in the log spot rate on the forward premium, e.g. Bilson (1981), the results reject the market efficiency hypothesis. Sequentially, the cointegration notion of Engle and Granger (1987) was broadly used in the latter studies, e.g. Hakkio and Rush (1989), Branhart and Szakmary (1991) and Naka and Whitney (1995). The cointegration relationship and tests become interesting points, and the empirical results are mixed and depend on the specification of cointegration models. However, the employment of cointegration for Crowder (1994) demonstrated that cointegration or lack of cointegration of spot exchange rates has nothing to do with foreign exchange market. We introduce some literatures in the remaining concept.

Hansen and Hodrick (1980) used the weekly spot rate and 3-month forward exchange rates from March 1970 to 1979 for seven currencies: the Canadian dollar, the Deutsche mark, the French franc, the U.K. pound, the Swiss franc, the Japanese yen, and the Italian lira to test the efficiency of the foreign exchange market. They employed two market forms:
First, to test the weak form of efficient market, and the regression of the forecast error on a constraint and two lagged forecast errors is showed as

\[ s_{i+13}^j - f_i^j = a_i + b_{i1}(s_i^j - f_{i-13}^i) + b_{i2}(s_{i-1}^j - f_{i-14}^i) + u_i^j \]

7
for \( i = 1, \ldots, 7 \) currencies.

Second, to test the semi-strong form efficient market, and the regression of the forecast error for a currency on lagged values of the own forecast error and four other currencies’ lagged forecast errors is showed as

\[
s_{i,t+13}^i - f_{i,t}^i = a_i + \sum_{j=1}^{5} b_{ji} (s_{t}^j - f_{t-13}^j) + u_{i,t}^i
\]

for \( i = 1, \ldots, 5 \) currencies. If market is efficient, all of the coefficients in above-mentioned two regressions are zero. Precisely, the null hypothesis is non-rejected.

The authors used data sampled interval is more finely than the forecast interval, that will cause overlapping problem and sequential serial correlation. Then, the conventional OLS regression would lead the estimator of coefficients to be inconsistent. Thus GLS is an appropriate approach to estimate the coefficients. To use GLS, however, requires the strict exogeneity for the independent variables, then \( f_{t,k} \) is useless in determining the optimal forecast for \( E(s_{t+k}) \). Therefore, Hansen and Hodrick made modified asymptotic covariance matrix, and employed OLS for estimation which is consistent but not fully efficient. To gain the asymptotic power of the tests, they increased the sample size of the data. The empirical result is presented that the Canadian dollar, the Deutsche mark, and the Swiss franc reject market efficiency hypothesis, moreover, the forecast error is autocorrelated. And it is probable that the asymptotic covariance matrix used for estimation and fluctuations in risk premiums cause the sample autocorrelations in forecast errors.

Bilson (1981) used 4-weekly spot rates and 1-month forward rates data from July 1974 to January 1980 for nine countries: Canada, U.K. France, Japan, Germany, Switzerland, Italy, Belgium and Netherlands. They considered the below regression of the spot rate of depreciation on the forward premium:

\[
S_t - S_{t-1} = \beta_0 + \beta_1 (F_{t-1} - S_{t-1}) + u_t
\]
to test the speculative efficiency. If the forward premium is an unbiased forecast, then $\beta_0 = 0, \beta_1 = 1$ and $E(u_t, u_{t-i}) = 0$ for $i \neq 0$. The results empirically indicated that pound, mark and Swiss franc do not reject the speculative efficiency. On the other hand, the Italian lira and the Dutch guilder reject that.

Frenkel (1981) used the monthly data from June 1973 to July 1979 for three currencies: the U.K. pound, the French franc, and the Deutsche mark to examine whether foreign exchange market is efficient. The exchange rate is a relative price of two assets and that expectations concerning the future course of events play a central role in affecting current exchange rates. If the foreign exchange market is efficient, current prices should reflect all currently available information, and expectations concerning future exchange rates should be incorporated and reflected in forward exchange rates. He considered the underlying regression:

$$s_t = a + bf_{t-1} + u_t$$

If the market for foreign exchange is efficient, then prices reflect all relevant available information, so that the residuals $u_t$ should contain no information and therefore should be serially uncorrelated. Further, the forward exchange rate is an unbiased forecast of the future spot exchange rate, then the joint hypothesis: $a = 0$ and $b = 1$ should be non-rejected. The empirical results of OLS, the foreign exchange market has been broadly consistent with the efficient market hypothesis.

Sanderson (1984) used the monthly spot rate and 1-month forward rate from February 1972 to April 1979 for three currencies: the U.K. pound, the Deutsch market, and the Canadian dollar. Under the rational expectations assumption, the regression specification is

$$f_t = \beta_0 + \beta_1 f_{t-n} + \varepsilon_t$$
If the joint hypothesis of market efficiency and the absence of a risk premium holds then $\beta_0 = 0, \beta_1 = 1$, $E(\varepsilon_i, \varepsilon_j) i \neq j$. The test result for OLS is consistent with market efficiency hypothesis. Furthermore, the author added three variables which is relevant in determining the exchange rate: the ratio of domestic to foreign price levels $(P/P^*)$, money supplies $(M/M^*)$ and output levels $(Y/Y^*)$ respectively to proceed market efficiency tests again. And the three regression is revealed as:

$$(s_t - f_{t-1}) = \beta_1 + \sum_{i=2}^{7} \beta_i (M/M^*)_{t-i} + u_t$$

$$(s_t - f_{t-1}) = \beta_1 + \sum_{i=2}^{7} \beta_i (Y/Y^*)_{t-i} + u_t$$

$$(s_t - f_{t-1}) = \beta_1 + \sum_{i=2}^{7} \beta_i (P/P^*)_{t-i} + u_t$$

The empirical consequence is also favourable to market efficiency hypothesis.

Hakkio and Rush (1989) used monthly data from July 1975 to October 1986 on the spot and forward rates for the British pound and German mark. If $S_{t+1}$ and $F_t$ are cointegrated, they can be written as an error correcting regression:

$$(S_{t-1} - S_t) = a(S_t - dF_{t-1} + b(F_t - F_{t-1})) + e_t$$

And market efficiency holds which requires $-a = b = d = 1$ with

$$S_{t+1} = F_t + e_t$$

That means the forward rate is an unbiased predictor of the future spot rate.

First, Hakkio and Rush examined whether the two spot rates and two forward rates were nonstationary. The consequence is yes. Second, they examine whether the two spot rates and two forward rates are cointegrated, respectively. Via
the demonstration of market efficiency from Granger(1986), the price of an asset incorporates all available information, has the important implication that prices from two efficient markets for different assets cannot be cointegrated. Then the result from seven tests suggests the two spot rates and two forward rates are not cointegrated. This indicates that both the German and UK spot and forward markets are efficient. Next, they examined whether the spot and forward rate are cointegrated from the same country. As the result, cointegration relationship is accepted. However cointegration is necessary but not sufficient for market efficiency. Thus, they finally used error correlation regression to examine whether market is efficient. The conclusion is that both German and UK markets are inconsistent the summarization of market efficiency.

Lai and Lai (1991) considered monthly spot rates and forward rates from July 1973 to December 1989 for five major currencies: the British pound, Deutsche mark, Swiss franc, Canadian dollar and Japanese yen against U.S. dollar. The basic equation is expressed as

\[ S_t = a + bF_{t-1,t} + u_t \]

And efficiency market is defined that \( a = 0 \) and \( b = 1 \), which is referred to as the unbiasedness hypothesis. First, to check whether each series of the spot rate and forward rate is a unit root by the augmented Dickey-Fuller test (ADF) and Phillips-Perron test (PP). The results are apparently that each series are unit root process, and the first difference of each series is stationary. Sequentially, to examine the cointegration relationship between \( S_t \) and \( F_{t-1,t} \). However, Engle and Granger approach could induce the inconsistence with the market efficiency hypothesis. Therefore using Johansen cointegration technique to examine the simple market efficiency hypothesis. Consequentially, the evidences suggest that
$S_t$ and $F_{t-1,t}$ are cointegrated, which implies the equilibrium relationship is

$$S_t - bF_{t-1,t} - a = 0$$

for every countries. Then to test the null hypothesis $b = 1$ and $a = 0, b = 1$, respectively. The result is found not favorable to the joint hypothesis of market efficiency and no-risk premium.

Branhart and Szakmary(1991) used daily spot rate and 1-month forward rate data from January 2, 1974 to November 30, 1988 for four currencies: the U.K. pound, the Canadian dollar, the Deutsche mark, and the Japanese yen. First, they employed DF test and ADF test to check whether spot rate and forward rate for each currency are nonstationary series. There are two kinds of regressive specifications to be tested, respectively:

1. level specification:

$$s_t = \mu + \theta f_{t-1} + e_t$$

2. present change specification:

$$s_t - s_{t-1} = \alpha + \beta (f_{t-1} - s_{t-1}) + \varepsilon_t$$

then the evidence is conclusively showed that two series for all currencies are non-stationary. Sequentially, to test whether each pairs of spot and forward rate are cointegrated, and the empirical consequence is presented that cointegrated relationship exists which means the relationship can also expressed as an error correlated model. However, the error correction form rejects market efficiency hypothesis empirically. Furthermore, the author use Chow test and Wald chi-square test to examine the stability of ECM coefficients, and result in negative coefficients with time-varying properties. That perhaps is the reason of rejection for market efficiency hypothesis.
Naka and Whitney (1995) used monthly data of one month forward and spot exchange rate from January 1974 to April 1991 for the seven major currencies: the U.K. pound, the Canadian dollar, the Deutsche mark, the French franc, the Italian lira, the Japanese yen, and the Swiss franc to test UBFH. The conventional regression for testing the unbiased forward rate hypothesis (UBFH) is

\[ s_t = \alpha + \beta f_{t-1} + e_t \]

The previously empirical result, however, suggests spot and forward rate are \( I(1) \) process and cointegrated, thus the OLS estimator of \( \beta \) is super-consistent, and the change in spot rate can be modeled by an error correlation model (ECM). Furthermore, Naka and Whitney identified parameters in the ECM that correspond to the parameters of the levels specification. Therefore, the link between the parameters of ECM and of levels specification is fully examined. Under the three assumptions:
1. the spot and forward rates are cointegrated,
2. first differences of forward rates are stationary,
3. first order autocorrelation is allowed.
then error correlation model is exhibited as:

\[ s_t - s_{t-1} = (1 - \rho)\alpha + (1 - \rho)(\beta f_{t-2} - s_{t-1}) + \beta(f_{t-1} - f_{t-2}) + \nu \]

For seven major currencies, conclusively, the empirical evidence is non-rejected the unbiased forward hypothesis.

Zivot (2000) used monthly spot rates and 1-month forward rates from January 1976 to June 1978 for three countries: U.K., Japan, and Canada. The author argued that the cointegration model between \( s_{t+1} \) and \( f_t \) is inappropriate. Because using a triangular cointegrated representation for \( (s_{t+1}, f_t)' \) leads that the OLS estimate of the coefficient on \( f_t \) in the levels regression is downward biased even
if market efficiency hypothesis is true. Therefore the author replaced $s_t$ into $s_{t+1}$ in the specification of cointegration model to examine the relationship in foreign exchange market. The simple cointegration model between $s_t$ and $f_t$ implies The simple cointegration model between $s_{t+1}$ and $f_t$, furthermore, the cointegrated form $(s_t, f_t)'$ would get the correct statistical inference.
Chapter 3

Approach Analysis

3.1 Unit Roots Tests

Stationarity has weak form and strong form, and the most commonly used form of stationarity in modeling time series is weakly-stationary or covariance-stationary. If the process $y_t$ satisfies:

$$E(y_t) = \mu$$
$$E(y_t - \mu)(y_{t-j} - \mu) = \gamma_j$$

for all $t$ and any $j$. The mean $\mu_t$ and the autocovariance $\gamma_{jt}$ for the process $y_t$ do not depend on the date $t$ as well. Then $\{y_t\}$ is said to be weakly-stationary process.

A nonstationary time series, on the contrary means that the shock is permanent and the variance of the nonstationary series is not consistent over time. The easiest way to analysis nonstationary time series is to make the series stationary by differencing. If nonstationary series needs to be differenced $k$ times to be stationary, then the series is said to be integrated of order $k$ which is denoted as $I(k)$.

While $k = 1$ means the process is to be stationary by differencing once which
is denoted as $I(1)$, it is also described as unit root process. There are three commonly used unit roots tests will be introduced.

### 3.1.1 Dickey-Fuller Test

Dickey and Fuller (1979) consider three autoregressive models that can be used to test the presence of the unit root:

\[
y_t = \rho y_{t-1} + u_t \tag{3.1}
\]

\[
y_t = \mu + \rho y_{t-1} + u_t \tag{3.2}
\]

\[
y_t = \mu + \delta t + \rho y_{t-1} + u_t \tag{3.3}
\]

where $y_0 = 0$, $u_t$ is i.i.d with mean zero and variance $\sigma^2$. The difference between the three regressions concerns the presence of the deterministic elements $\alpha$ and $\delta$. $\mu$ is the constant term, and $t$ is time trend.

Given $T$ observations $y_1, y_2, \ldots, y_T$, the OLS estimator for $\rho$ is $\hat{\rho}$ which is a consistent estimator. Consider model (3.1), (3.2) and (3.3), the estimator of $\rho$ is labeled as $\hat{\rho}$, $\hat{\rho}_\mu$ and $\hat{\rho}_\tau$, respectively. To test null hypothesis $\rho = 1$, the statistic for testing $H_0$ is labeled as $t$, $t_\mu$ and $t_\tau$, respectively.

Because of the true value $\rho = 1$, then the asymptotic distribution of $T(\hat{\rho} - 1)$, $T(\hat{\rho}_\mu - 1)$, $T(\hat{\rho}_\tau - 1)$, $t$, $t_\mu$, and $t_\tau$ is expressed as Brownian Motion processes. Noted that the assumption in model (3.2) is $\mu = 0$, The assumption in model (3.3) is $\delta = 0$, and the distribution of $\hat{\rho}_\tau$ and $\hat{r}_\tau$ are unaffected by the value $\delta$. If $\mu \neq 0$ for model (3.2) or $\delta \neq 0$ for model (3.3), the asymptotic distribution is $t_\mu$ and $t_\tau$ are normal.

Dickey and Fuller (1979) found the critical values in the Monte Carlo study. If the statistics are greater than the critical values, then the null hypothesis is accepted and we should conclude that the process is nonstationary.
3.1.2  Phillips-Perron Test

Phillips (1987) and Phillips and Perron (1988) released the stringent restrictions for the residual $u_t$ of AR(1) process which is iid($0, \sigma^2$). The assumption of the more general residual is showed that: $u_t$ is a zero-mean but otherwise heterogeneously distributed process satisfying certain restrictions on the serial dependence and higher moments.

The regression model is:

\[
y_t = \alpha + \rho y_{t-1} + u_t \tag{3.4}
\]

\[
u_t = \psi(L)\varepsilon_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}
\]

where $\sum_{j=0}^{\infty} j|\psi_j| < \infty$ and $\{\varepsilon_t\}$ is an i.i.d ($0, \sigma^2$) sequence, and finite fourth moment. Define

\[
\gamma_j \equiv \mathbb{E}(u_t u_{t-j}) = \sigma^2 \sum_{s=0}^{\infty} \psi_s \psi_{s+j} \quad \text{for } j = 0, 1, 2, \ldots
\]

\[
\lambda \equiv \sigma \sum_{j=0}^{\infty} \psi_j = \sigma \psi(1)
\]

\[
\xi \equiv u_1 + u_2 + \cdots + u_t \quad \text{for } t = 1, 2, \ldots, T
\]

with $\xi_0 = 0$.

This approach is to estimate equation (3.4) by OLS to calculate $\hat{\rho}_T$, and Phillips and Perron suggested that the adjustments are added to the OLS $t$ – statistic presented in (3.5) and $\hat{\rho}_T$ presented in (3.6), respectively,

\[
T(\hat{\rho}_T - 1) - \frac{1}{2}(T^2 \hat{\sigma}_{\hat{\rho}_T}^2 \div S_T^2)(\hat{\lambda}^2 - \hat{\gamma}_0) \tag{3.5}
\]

\[
(\hat{\gamma}_0 \hat{\lambda}^2)^{1/2}.T - \{\frac{1}{2}(\hat{\lambda}^2 - \hat{\gamma}_0)/\hat{\lambda}\} \times \{T.\hat{\sigma}_{\hat{\rho}_T}^2 \div s_T\} \tag{3.6}
\]
where $\hat{u}_t$ is the OLS estimator of the residual,
\[
\hat{\lambda}^2 = \hat{\gamma}_0 + 2 \sum_{j=1}^{q} [1 - j/(q + 1)] \hat{\gamma}_j
\]
\[
\hat{\gamma}_j = T^{-1} \sum_{t=j+1}^{T} \hat{u}_t \hat{u}_{t-j}
\]
$q$ is the lag truncation number, $s_T$ is the OLS estimate of the variance of $u_t$, and $\hat{\sigma}_{\hat{\rho}_T}$ is the OLS standard error of $\hat{\rho}_T$.

Under the null hypothesis $H_0 : \alpha = 0, \rho = 1; H_1 : \alpha \neq 0, |\rho| < 1$, the true process is $y_t = y_{t-1} + u_t$. Phillips and Perron provided the statistic which contains adjustment terms has the same asymptotic distribution with the statistic for Dickey-Fuller model. That is, the critical values of Dickey-Fuller model and PP model are equivalent. Therefore, the Dickey-Fuller study can be used in the PP approach as well.

### 3.1.3 Augmented Dickey-Fuller Test

Said and Dickey (1984) extended the regressive model of Dickey and Fuller (1979), the series of first differences are the general ARMA($p,q$) process under the null hypothesis, essentially the moving average term are approximated by including enough autoregressive terms, there are three models expressed:

\[
y_t = \rho y_{t-1} + \zeta_1 \Delta y_{t-1} + \zeta_2 \Delta y_{t-2} + \cdots + \zeta_{p-1} \Delta y_{t-p+1} + \varepsilon_t
\]  
(3.7)

\[
y_t = \alpha + \rho y_{t-1} + \zeta_1 \Delta y_{t-1} + \zeta_2 \Delta y_{t-2} + \cdots + \zeta_{p-1} \Delta y_{t-p+1} + \varepsilon_t
\]  
(3.8)

\[
y_t = \alpha + \delta t + \rho y_{t-1} + \zeta_1 \Delta y_{t-1} + \zeta_2 \Delta y_{t-2} + \cdots + \zeta_{p-1} \Delta y_{t-p+1} + \varepsilon_t
\]  
(3.9)

where $\alpha$ is constant term, $\delta$ is trend. $\{\varepsilon_t\}$ is the residual of AP($p$) process which is an i.i.d sequence with zero-mean, variance $\sigma^2$, and finite fourth moment. And
\[
\rho \equiv \phi_1 + \phi_2 + \cdots + \phi_p
\]
\[
\zeta_j \equiv -[\phi_{j+1} + \phi_{j+2} + \cdots + \phi_p] \quad \text{for } j = 1, 2, \cdots, p - 1
\]
The difference between the OLS estimate $\hat{\rho}$ and the hypothesized true value of unity is multiplied by the sample size and divided by $(1 - \hat{\zeta}_1 - \hat{\zeta}_2 - \cdots - \hat{\zeta}_{p-1})$, the resulting statistic has the same asymptotic distribution as the statistic of the DF tests. Hence, the critical value for ADF Test can be found in Dickey-Fuller study.
3.2 Conventional Cointegration

The use of straightforward linear regression was common practice to estimate equations involving nonstationary variables for a long time. Granger and Newbold (1974), however, provided it was inappropriate for testing hypotheses about the coefficients using standard statistical inference. Because standard statistical inference might cause spurious results. They considered the model:

\[ y_t = \beta x_t + u_t \]

and assumed that \( y_t \) and \( x_t \) are \( I(1) \) processes, \( y_t \) and \( x_t \) are not cointegrated, and the error term \( u_t \) is not \( I(0) \). Then if using OLS to estimate the coefficient \( \beta \) will lead to spurious regression. Therefore, economists suggested some solutions to avoid spurious regressions. For instance, the model is specified in differences instead of in levels, that means the differenced variables are usually stationary even if the original variables are not. Another approach is to remove a linear time trend from the variables and to specify the empirical relationship between them using de-trended variables.

Engle and Granger (1987) defined 'cointegration' is that: The components of the vector \( y_t \) are said to be cointegrated of order \( d, b \), denoted \( y_t \ CI(d, b) \), if all components of \( y_t \) are \( I(d) \), and there exists a vector \( a(\neq 0) \) so that

\[ z_t = a'y_t \quad I(d, b), b > 0 \]

The vector \( a \) is called the cointegration vector.

Engle and Granger (1987) summarized cointegration notions in the form of a proposition.

**Theorem.** Granger representation theorem:

Consider an \( (n \times 1) \) vector \( y_t \), where \( \Delta y_t \) has the Wold representation

\[ (1 - L)y_t = \delta + \Psi(L)\varepsilon_t \]

where \( \varepsilon_t \) is white noise with positive definite variance-covariance matrix and \( \{s\psi_s\}_{s=0}^\infty \) absolutely summable. Suppose that there are exactly \( h \) cointegrating
relations among the elements of \( y_t \). Then there exists an \((h \times n)\) matrix \( A' \) whose rows are linear independent such the \((h \times 1)\) vector \( z_t \) defined by:

\[
z_t = A'y_t
\]
is stationary. The matrix \( A' \) has the property that

\[
A'
\Psi(1) = 0
\]

If the process can be represented as the \( p \)-th order VAR in levels as

\[
y_t = \alpha + \Phi_1y_{t-1} + \Phi_2y_{t-2} + \cdots + \Phi_py_{t-p} + \varepsilon_t
\]

then there exists an \((n \times h)\) matrix \( B \) such that

\[
\Phi(1) = BA'
\]

and there further exist \((n \times n)\) matrices \( \zeta_1, \zeta_2, \cdots, \zeta_{p-1} \) such that

\[
\Delta y_t = \zeta_1\Delta y_{t-1} + \zeta_2\Delta y_{t-2} + \cdots + \zeta_{p-1}\Delta y_{t-p+1} + \alpha - Bz_{t-1} + \varepsilon_t
\]

Cointegration implies that there exist one or more long-run relationships among variables, and deviations from which tend to be eliminated over time.

**3.2.1 Engle and Granger Two Steps Method**

Consider the regression:

\[
y_t = \beta x_t + u_t \tag{3.11}
\]

where two variables \( y_t \) and \( x_t \) are both \( I(1) \).

Step 1. Employ OLS to estimate the residual \( \hat{u}_t \)

Then to estimate the parameter \( \beta \) by OLS, and get the OLS residual \( \hat{u}_t \). If there exists cointegration relationship, then \( u_t \) \( I(0) \). Since \( \hat{\beta} \) converges the true value \( \beta \) at the rate \( T \) instead of the usual rate \( \sqrt{T} \), the estimator \( \hat{\beta} \) is superconsistent.
Step 2. To determine whether cointegration relationship exists by testing whether $u_t$ is $I(1)$ process.

Test the null hypothesis that there is no cointegration relationship between $y_t$ and $x_t$ by testing the null hypothesis $u_t$ is $I(1)$. Engle and Granger (1987) proposed several useful approaches to test the null hypothesis $u_t$ is $I(1)$:

1. CRDW
2. DF
3. ADF
4. RVAR
5. ARVAR
6. UVAR
7. AUVAR

and conclude that ADF is better. The general ADF test specification is:

$$\Delta\hat{u}_t = \psi^*\hat{u}_{t-1} + \sum_{i=1}^{p-1} \psi^* \Delta\hat{u}_{t-i} + \mu + w_t$$  \hspace{1cm} (3.12)

where $w_t \ iid(0, \sigma^2)$. Via above-maintained two-step test, if the evidences is in favorable to existing cointegration relationship, then these variables can be also expressed as error correction model (ECM). The superconsistency of $\hat{\beta}$ assures that the two-steps estimators of the parameters of the ECM have the same asymptotic distribution as the ones that one would get if $\beta$ is known.

Engle and Granger two-step test (1987) are single equation methods which is simple procedure in practice, nevertheless, there exists some defects. One of the defects is although the $OLS$ estimate $\hat{\beta}$ is superconsistent, its asymptotic distribution depends on nuisance parameters arising from endogeneity of the regressor and serial correlation in errors. The second, while the sample size is not large enough, even if $\hat{\beta}$ is superconsistent, that would result in finite sample biased. Finally, the two-variable model restricts the number of cointegrating vector is
one. Hence this test would not be fit to more than two-variable model or to the coitegrating vectors numbers exceed in 1.

### 3.2.2 Johansen Maximum Likelihood Method

Johansen Maximum Likelihood Method is provided by Johansen (1988) and Johansen and Juselius (1990). This procedure applying maximum likelihood to the VAR model, and consider the relationships among more than two variables. Besides, that can also estimate and examine the numbers of cointegrating vectors.

Let $y_t$ denote an $(n \times 1)$ vector. The maintained hypothesis is that $y_t$ follows a VAR($p$) in levels, and all of the elements for $y_t$ are $I(1)$ process. In addition, the errors are Gaussian.

$$
y_t = \mu + \Pi_1 x_{t-1} + \Pi_2 x_{t-2} + \cdots + \Pi_p x_{t-p} + \varepsilon_t \quad t = 1, 2, \ldots, T
$$

where $\mu$ is constant term, and $\varepsilon_t$ is iid $N(0, \Omega)$. Moreover, VAR($p$) in levels can be written as:

$$
\Delta y_t = \mu + \zeta_1 \Delta y_{t-1} + \zeta_2 \Delta y_{t-2} + \cdots + \zeta_{p-1} \Delta y_{t-p+1} + \zeta y_{t-1} + \varepsilon_t \quad (3.13)
$$

where

$$
\zeta = -(I_n - \Pi_1 - \Pi_2 - \cdots - \Pi_p) = -\Pi(1)
$$

$$
\zeta_i = -(I_n - \Pi_1 - \Pi_2 - \cdots - \Pi_i) \quad i = 1, 2, \ldots, p - 1
$$

Suppose that each individual variable $y_{it}$ is $I(1)$, and $r$ linear combinations of $y_t$ are stationary. That implies $\zeta$ can be showed as

$$
\zeta = -\alpha \beta'
$$

where $\beta$ is the cointegrating matrices, and $\alpha$ is the adjustment coefficients for both $\alpha$ and $\beta$ ($r \times n$) matrices. The number of cointegrating relations relies on
the rank of $\zeta$, and the rank of $\zeta$:

1. $\text{rank}(\zeta) = n$,

$\zeta$ is full rank means that all components of $y_t$ is $I(1)$, that is $y_t$ is a stationary process.

2. $\text{rank}(\zeta) = 0$,

$\zeta$ is null matrix, that means there is no cointegration relationships.

3. $0 < \text{rank}(\zeta) = r < n$,

That means the variables for $y_t$ are cointegrated, and the number of cointegrating vectors is $r$.

Step 1: Calculate Auxiliary Regressions

Regress $\Delta y_t$ on $\Delta y_{t-1}, \ldots, \Delta y_{t-p+1}$, and get the residuals. For simply, let

$$ Z_{0t} = \Delta y_t $$

$$ Z_{1t} = 1, \Delta y_{t-1}, \ldots, \Delta y_{t-p+1} $$

$$ Z_{pt} = y_{t-p} $$

and let $\Gamma$ consist of the parameters $(\mu, \zeta_1, \ldots, \zeta_{p-1})$. Rewording equation (3.13) as

$$ Z_{0t} = \Gamma Z_{1t} + \zeta Z_{pt} + \varepsilon_t $$

To use OLS for estimation, then get the residuals: $R_{0t}$ and $R_{pt}$, and the residual sums of squares is:

$$ S_{ij} = M_{ij} - M_{i1} M_{11}^{-1} M_{1j} \quad (i, j = 0, p) $$
That gives the concentrated likelihood function:

\[ L(\alpha, \beta, \Omega) = |\Omega|^{-T/2} \exp\left[-\frac{1}{2} \sum_{t=1}^{T} \left( R_{0t} - \zeta R_{pt} \right)' \Omega^{-1} \left( R_{0t} - \zeta R_{pt} \right) \right] \tag{3.14} \]

Step 2: Calculate Canonical Correlations
To get the maximum of the likelihood function by solving the eigenvalues of

\[ |\lambda S_{pp} - S_{p0} S_{00}^{-1} S_{pp}| = 0 \]

We obtain the eigenvalues: \( \hat{\lambda}_1 > \hat{\lambda}_2 > \ldots > \hat{\lambda}_n > 0 \), and the normalized eigenvectors: \( \hat{V} = (\hat{v}_1, \hat{v}_2, \ldots, \hat{v}_n) \), where \( \hat{V}' S_{pp} \hat{V} = I \).

Step 3: Calculate Maximum Likelihood Estimates of Parameters
If cointegration relations exist, then the cointegrating vector \( \beta \) is constructed by the first \( r \) normalized eigenvectors, that is \( \hat{\beta} = (\hat{v}_1, \hat{v}_2, \ldots, \hat{v}_r) \), and \( \zeta = \alpha \beta' \).

Eq[3-2] can be rewritten as

\[ L(\alpha, \beta, \Omega) = |\Omega|^{-T/2} \exp\left[-\frac{1}{2} \sum_{t=1}^{T} \left( R_{0t} - \alpha \beta' R_{pt} \right)' \Omega^{-1} \left( R_{0t} - \alpha \beta' R_{pt} \right) \right] \]

This function is minimized for fixed \( \beta \) to give

\[ \hat{\alpha}(\hat{\beta}) = S_{00}^{-1} \hat{\beta} \left( \hat{\beta}' S_{kk} \hat{\beta} \right)^{-1} \]
\[ \hat{\Omega}(\hat{\beta}) = S_{00} - S_{0p} \left( \hat{\beta}' S_{kk} \hat{\beta} \right)^{-1} \hat{\beta}' S_{0k} \]
\[ \hat{\Gamma} = (M_{01} - \zeta M_{p1}) M_{11}^{-1} \]

where

\[ R_{0t} = Z_{0t} - M_{01} M_{11}^{-1} Z_{1t} \]
\[ R_{pt} = Z_{pt} - M_{p1} M_{11}^{-1} Z_{1t} \]
\[ M_{ij} = T^{-1} \sum_{t=1}^{T} Z_{it} Z_{jt}', \quad i, j = 0, 1, p \]

\( M_{ij} \) is defined as the product moment matrices.
To determine the number of cointegrating vectors, Johansen suggests two tests:

1. Trace Test

\[ H_0 : \text{Rank}(\zeta) \leq r, \text{i.e. there are at most } r \text{ cointegrating vectors.} \]
\[ H_1 : \text{Rank}(\zeta) > r \]

The test statistic is:

\[ \lambda_{\text{trace}} = -T \sum_{i=r+1}^{n} \ln(1 - \hat{\lambda}_i) \]

The statistic has a limit distribution which can be expressed in terms of a \((n-r)\)-dimensional Brownian motion, and the limit distribution is equivalent to the trace for the matrix \(Q\).\(^2\)

2. Maximum eigenvalues test

\[ H_0 : \text{there are } r \text{ cointegrating vectors} \]
\[ H_1 : \text{there are } r + 1 \text{ cointegrating vectors} \]

The test statistic is:

\[ \lambda_{\text{max}} = -T \ln(1 - \hat{\lambda}_{r+1}) \]

The test statistic is asymptotically distributed as the maximum eigenvalue of the matrix \(Q\).

There are, however, at least two problems for Johansen procedure:

1. the assumption for the errors are independent and normal.

\[ Q = [\int_0^1 W(r)dW(r)'][\int_0^1 W(r)W(r)'dr]^{-1}[\int_0^1 W(r)dW(r)'] \]
2. Johansen method produces more outliers than the other methods. By comparison, although OLS method is more robust, Johansen procedure is still the most commonly used system method in cointegration analysis, in spite of the defects.
3.3 Stochastic Cointegration

Engle-Granger (1987) indicated that a linear combination of individually $I(1)$ series to be $I(0)$, equivalently, the residual error in the linear is to be stationary. However, if we consider two respects: one is that as the regressors increase in magnitude, the residual variance would also increase; the other is that the variance of the error process might change over time due to other factors. Then Engle-Granger (EG) cointegration is not sufficiently general to cover all nonstationary economic models. Thus Hansen (1992) provided a concept which allows the variance of the error to be asymptotically nonstationary in a regression. He considered a process $w_t = \sigma_t e_t$ which is called bi-integrated (BI) process, where $\sigma_t \equiv I(1)$ as the variance part and $e_t \equiv I(0)$ as the stationary part. The differences between BI process and $I(1)$ are: the former will tend to cross its mean value, and the latter will wander around without tendency to return to any particular value.

The nonlinear regression model is defined as heteroscedastic cointegration (HCI) which can be expressed:

$$y_t = \beta_0 + \beta'_1 x_t + w_t$$

$$x_t = x_{t-1} + e_{3t}$$

$$w_t = \sigma_t e_{1t}$$

$$\sigma_t = \sigma_{t-1} + e_{2t}$$

where the $n \times 1$ vectors $x_t$ is $I(1)$, the error $w_t$ is a BI process, $\sigma_t$ is scale $I(1)$ process. To compare EG cointegration and HCI with some respects: The first point is the error term. The error of EG cointegration reveals a $I(0)$ stationary process, and it presents a BI process for HCI. Secondly, the error’s variance with no stochastic trend is constant in EG specification; and the error’s variance for HCI is potentially unbounded which due to the second moments of the BI error
process growing linearly over time. Finally, both the variance of regressors for EG cointegration and HCI possess stochastic trend and grow linearly over time.

Harris, McCabe and Leybourne (2002) consider nonlinear and nonstationary time series, which is termed as stochastically integrated processes, consists of heteroscedastic and $I(1)$ integration. And they propose stochastic cointegration, finding the interrelationships between stochastically integrated processes, contains conventional EG cointegration (Engle and Granger 1987) and heteroscedastic cointegration (Hansen 1992). The regressions are stochastically cointegrating, and the regression errors appear to be $I(0)$ stationary process plus heteroscedastically integrated process.\(^3\) In the sense that variables could still trend together over the long run with some conditions. Stochastic cointegration is much weaker than EG cointegration is the former requires only $I(1)$ is absent, and the latter requires the presence of $I(0)$ instead. Besides, HML model release the regression of Hansen (1992) which imposes asymmetry condition, that is only the regressand could to be HI process, and all regressors require to be $I(1)$. Thus, HML model imposing no asymmetry condition, so that it is possible both the regressand and some (or all) of the regressors are heteroscedastically integrated processes.

### 3.3.1 The Model

Assume that the observable vector time series $z_t$ satisfied:

\[
\begin{align*}
    z_t &= \mu + \Pi_t w_t + \epsilon_t \\
    w_t &= w_{t-1} + \eta_t \\
    \Pi_t &= \Pi + V_t
\end{align*}
\]  

for $t = 1, \ldots, T$. Where $z_t, w_t, \epsilon_t, \eta_t, \mu$ are $m \times 1$ vectors, and $\Pi_t, \Pi, V_t$ are $m \times m$ matrices. The disturbances $\epsilon_t, \eta_t$ and $V_t$ are mean zero stationary processes.

\(^3\)The heteroscedastically integrated process abbreviated by HI process, that is equivalent to BI process in Hansen (1992)).
which may be correlated. $w_t$ is a vector intergrated process with $w_0 = \eta_0$, and $\mu$ is a vector of constants. $\Pi_t$ is random and produces shocks nonlinearly into $z_t$. Rewriting $z_t$ as:

$$z_t = \mu + \Pi w_t + \varepsilon_t + V_t w_t 
$$

(3.16)

$z_t$ contains a integrated process $\Pi w_t$ and a new shock term $\varepsilon_t + V_t w_t$ which has a linear component $\varepsilon_t$ and a nonlinear component $V_t w_t$.

Let $e_i$ be an $m \times 1$ vector with 1 in its $i$th position and 0 elsewhere, so that

$$e_i' z_t = e_i' \mu + e_i' \Pi w_t + (e_i' V_t w_t + e_i' \varepsilon_t) 
$$

(3.17)

If $e_i' \Pi \neq 0$, then $z_{it}$, the $i$'th element of the vector $z_t$, is said to be stochastic integrated. In addition, if $e_i E(V_t V_t') e_i > 0$, $z_{it}$ is said to be heteroscedastically integrated; if $e_i V_t = 0$, $z_{it}$ is simply I(1). In the event, a stochastically integrated variable contains I(1) and heteroscedastic integration.

If there exists a vector that makes the linear combination for above model meaningful and makes it defined stochastic cointegration, then such a vector is termed stochastically integrated vector. Let $c$ is a nonzero $m \times 1$ vector and consider

$$c' z_t = c' \mu + c' \Pi w_t + (c' V_t w_t + c' \varepsilon_t) 
$$

(3.18)

If $c' \Pi = 0$ then the variables in $z_t$ are said to be stochastically cointegrated; otherwise they are not stochastically cointegrated.

Given $c' \Pi = 0$, if $E(c' V_t V_t' c) = 0$, then $c' z_t = c' (\mu + \varepsilon_t)$ is stationary referred to stationary cointegration. If, in addition, $V_t = 0$, the variables are all integrated and cointegrated in Engle-Granger sense. If $E(c' V_t V_t' c) > 0$, then the variables in $z_t$ are said to be heteroscedastically cointegrated.
Assumption LP (linear process assumption): Let
\[
\zeta_t = \begin{bmatrix} \text{vec}(V_t) \\ \eta_t \\ \epsilon_t \end{bmatrix}
\]
and be generated by the vector linear process \( \zeta_t = \sum_{j=0}^{\infty} C_j \xi_{t-j} \), where

1. \( \sum_{j=0}^{\infty} j^2 ||C_j||^2 < \infty \) with \( C_0 \) having full rank,
2. \( \xi_t \) is a martingale difference sequence with respect to \( \mathcal{F}_t \), the \( \sigma \)-field generated by \( \{ \zeta_s, s \leq t \} \).
3. \( \mathbb{E}(\xi_t \xi'_t | \mathcal{F}_{t-1}) = I_{m^2+2m} \),
4. For all \( i \) and \( t \), \( \mathbb{E}(\xi_{it}^8) < B^* \) i.e., the eighth moment is uniformly bounded.

Under Assumption LP, \( \epsilon_t + V_t w_t \) is stochastically trendless.\(^4\) And \( c' z_t = c'(\epsilon_t + V_t w_t) \) is stochastically trendless as well.

### 3.3.2 The Regression

The regression formulation is expressed as that: Partition \( z_t \) into two parts as \( z_t = (y_t, x'_t) \), where \( y_t \) is a scalar and \( x_t \) is a \( (m-1) \times 1 \) vectors. Then equation (1) becomes
\[
\begin{pmatrix} y_t \\ x'_t \end{pmatrix} = \begin{pmatrix} \mu_y \\ \mu_x \end{pmatrix} + \begin{pmatrix} \pi_{yt}' \\ \Pi_{xt}' \end{pmatrix} w_t + \begin{pmatrix} \epsilon_{yt} \\ \epsilon_{xt} \end{pmatrix},
\]
\[
\begin{pmatrix} \pi_{yt}' \\ \Pi_{xt}' \end{pmatrix} = \begin{pmatrix} \pi_{yt}' \\ \Pi_x' \end{pmatrix} + \begin{pmatrix} v_{yt}' \\ V_{xt}' \end{pmatrix}.
\]

\(^4\)Trendless is similar to the concept of a mixingale and the associated notion of asymptotically unpredictability. A behavior of the process up to time \( t \) has negligible effect on its behavior into the infinite future.
Defining \( c = (1, -\beta)' \) and \( \alpha = \mu_y - \beta' \mu_x \) gives the regression formulation

\[
y_t = \alpha + x'_t \beta + u_t \tag{3.20}
\]

\[
u_t = c' \Pi w_t + c' V_t w_t + c' \varepsilon_t
\]

with

\[
c' \Pi = \pi'_t - \beta' \Pi_x,
\]

\[
c' V_t = v'_y t - \beta' V_x t,
\]

\[
c' \varepsilon_t = \varepsilon_y t - \beta' \varepsilon_x t
\]

The error term \( u_t \) is composed of a stochastic trend \( c' \Pi w_t \), a heteroscedastic component \( c' V_t w_t \) and stationary term \( c' \varepsilon_t \). When \( c' \Pi = 0 \) i.e. \( u_t \) contains no stochastic trend component and \( \beta \) is referred to as the stochastically cointegrating vector. To assume that \( \text{rank}(\Pi_x) = m - 1 \) which ensures that no further sub-stochastically cointegrating relationships exist among the \( x_t \).

Note that a constant will be always included in the regression, and the error term \( u_t \) need not have zero so that \( \alpha \) is not an intercept anymore.

### 3.3.3 Estimation

OLS estimator is proved to be inconsistent in the HML model, then they proposed a valid method to solve the inconsistency problem. That is 'asymptotic instrument variables', abbreviated by AIV. Defining \( X_t = (1, x'_t) \), with \( k = k(T) > 0 \) is used as an instrument, the AIV estimator of \( b = (\alpha, \beta)' \) is \( \hat{b}_k = (\hat{\alpha}_k, \hat{\beta}_k)' \) expressed by

\[
\hat{b}_k = \left( \sum_{t=k+1}^{T} X_{t-k} X'_t \right)^{-1} \sum_{t=k+1}^{T} X_{t-k} y_t \tag{3.21}
\]

and the AIV residuals is \( \hat{u}_t = y_t - \hat{\alpha}_k - x'_t \hat{\beta}_k \). \( \text{cov}(\text{vec}(\Pi_{x,t-k}), V'_t C) \) is not a zero matrix under Assumption LP, if \( k \to \infty \) and it will converge to 0 for processes.
satisfying Assumption LP. Therefore, the condition \( k = O(T^{1/2}) \) is required. That is, the AIV estimator is consistent if we let \( k \to \infty \) as \( T \to \infty \).

Because stochastic cointegration contains stationary cointegration and heteroscedastic cointegration, then the estimation should be separated into two parts.

**Part 1.**
If \( E(c'V_tV_t'c) = 0 \) (stationary cointegration)
Assume model () with \( c'\Pi = 0 \), Assumption LP hold, and let \( k = O(T^{1/2}) \) then

\[
\hat{\alpha}_k - \alpha = O_p(T^{-1/2}),
\]

\[
\hat{\beta}_k - \beta = O_p(T^{-1}).
\]

Then, \( \hat{\alpha}_k \) and \( \hat{\beta}_k \) are consistent regardless of whether \( v_{yt}, V_{xt} = 0 \) or whether \( V_{xt} \) and \( c'\varepsilon_t \) are correlated or uncorrelated.

**Part 2.**
If \( E(c'V_tV_t'c) > 0 \) (heteroscedastic cointegration)
Assume model () with \( c'\Pi = 0 \), Assumption LP holds, and let \( k = O(T^{1/2}) \), then

\[
\left( \begin{array}{c}
\hat{\alpha}_k - \alpha \\
T^{1/2}(\hat{\beta}_k - \beta)
\end{array} \right) \overset{d}{\to} 
\left( \begin{array}{cc}
\Pi_x \int_0^1 W(s)ds & \Pi_x \int_0^1 W(s)W(s)'ds \Pi_x' \\
1 & \int_0^1 W(s)'dU(s) + \text{tr}(\Lambda_0) \\
\int_0^1 [W(s) \otimes I_{m-1} \otimes W(s)']dH(s) + \text{tr}(\Lambda_1)\Pi_x \int_0^1 W(s)ds
\end{array} \right)^{-1}
\]

where \( \Lambda_j = \sum_{i=j}^{\infty} E(V_t'c\eta_{t-i}) \) for \( j = 0, 1 \).

However the asymptotic distribution contains nuisance parameters \( \Lambda_0, \Lambda_1 \) and there exists the correlation between these Brownian motions, which imply that the asymptotic distribution is not mixed normal. For solving these problems,
Exogeneity condition is imposed.

**Exogeneity Condition:**

\[
E(\text{vec}(V_t)\eta'_{t-j}) = 0, \quad j = 0, \pm 1, \pm 2, \ldots \tag{3.22}
\]

Then

\[
\left( \begin{array}{c}
\hat{\alpha}_k - \alpha \\
T^{1/2}(\hat{\beta}_k - \beta)
\end{array} \right) \xrightarrow{d} \left( \begin{array}{c}
1 \\
\Pi_x \int_0^1 W(s)'d\Pi_x' \\
\Pi_x \int_0^1 W(s)W(s)'d\Pi_x' \\
\int_0^1 W(s)'dU(s) \\
\int_0^1 [W(s) \otimes I_m] \otimes W(s)'dH(s)
\end{array} \right)^{-1}
\]

where \( W \) and \((U, H')'\) are independent Brownian motions.\(^5\)

It seems perfect that the asymptotic distribution of \((\hat{\alpha}_k, \hat{\beta}_k)'\) is already normal. But autocorrelation and heteroscedasticity are still present in the regression error \( u_t = c'V_tw_t + c'\varepsilon_t \). To eliminate autocorrelation and heteroscedasticity, there are some necessary extra notation and assumption. Let \( X_t = (1, x_t')' \) and define

\[
\Sigma_k = T\left( \sum_{t=k+1}^T X_{t-k}X_t' \right)^{-1} \hat{\Omega}(X_{t-k}\hat{u}_t)(\sum_{t=k+1}^T X_tX_t')^{-1}
\]

where \( \hat{u}_t = y_t - \hat{\alpha}_k - x_t'\hat{\beta}_k \) are the AIV regression residuals. \( \hat{\Omega}(.) \) is the estimated long-run variance matrix.\(^6\)

---

5 see Appendix A

6 We use a HAC type covariance matrix (see Andrews, 1991). Let \( a_t \) is any vector time series,

\[
\hat{\Omega}(a_t) = \Gamma_0(a_t) + \sum_{j=1}^l \lambda(j) \hat{\Gamma}_j(a_t) + \hat{\Gamma}_j(a_t)'
\]

\[
\hat{\Gamma}_j(a_t) = T^{-1} \sum_{t=j+1}^T a_t a_{t-j}'
\]

\( \lambda(\cdot) \) is one of the usual admissible windows, with the lag truncation parameter \( l \).
Assumption KN (kernel and lag length)
1. $\lambda(0) = 1$,
2. $0 \leq \lambda(x) \leq 1$ for $-1 < x < 1$, $\lambda(x) = 0$ for $|x| \geq 1$,
3. $\lambda(x)$ is the even and continuous on $[0, 1]$,
4. $l \to \infty$ as $T \to \infty$ such that $l = o(k)$ and $l < k$.

The consequence is to assume model (4) with $c'\Pi = 0$, ASSUMPTION LP and KN and the exogeneity condition hold. Let $k = O(T^{1/2})$, then

$$
\hat{\Sigma}_k^{-1/2} \begin{pmatrix} \hat{\alpha}_k - \alpha \\ \hat{\beta}_k - \beta \end{pmatrix} \overset{d}{\to} N(0, I_m)
$$

where

$$
\hat{\Sigma}_k^{-1/2} = T^{-1/2}( \sum_{t=k+1}^{T} X_{t-k}X_t' )^{-1}\hat{\Omega}(X_{t-k}\hat{u}_t)^{1/2} 
$$

(3.23)

### 3.3.4 Hypothesis Tests and Test Statistics

To test $H^0 : \beta_i = \beta_i^*$ against $H^1 : \beta_i \neq \beta_i^*$, the $t$-statistic is

$$
t = \frac{\hat{\beta}_{ki} - \beta_i^*}{\sqrt{\left(\hat{\Sigma}_k^{-1}\right)_{ii}}}
$$

where $\sqrt{\left(\hat{\Sigma}_k^{-1}\right)_{ii}}$ means the element positions of the covariance matrix, $i = 1, 2$.

Also the $t$-statistic of $\alpha$ has the same formula. And under $H^0$, the $t$-statistic for $\alpha$ and $\beta$ are both asymptotically standard normal.

The null hypothesis of stochastic cointegration against the alternative of no stochastic cointegration is expressed by the model (3.20) as

$$
H^0 : c'\Pi = 0 \\
H^1 : c'\Pi \neq 0
$$
Under $H^0$, The null hypothesis of stationary cointegration against the heteroscedastic alternative is

$$H^0_0: E(c'V_tV'_t c) = 0$$
$$H^0_1: E(c'V_tV'_t c) > 0$$

**Testing $H^0$ Against $H^1$**

To test stochastic cointegration against non-stochastic cointegration, the point is to test whether $c'\Pi = 0$ in

$$u_t = c'\Pi w_t + c'V_tw_t + c'\varepsilon_t$$

Consider

$$S_{nc} = \sum_{t=k+1}^{T} u_t u_{t-k}$$

which is specified to eliminate the nuisance parameters that result from the autocorrelation and from the presence of $c'V_t w_t$. Where the lag $k$ is allowed to increase with $T$, and becomes large eliminates any correlation between $u_t$ and $u_{t-k}$ under $H^0$. Under the alternative, because of the I(1) term $c'\Pi w_t$ allowing $k$ to grow does not eliminate correlation between $u_t$ and $u_{t-k}$.

To find the asymptotic distribution of the test statistic:

Assume the model (4), Assumption LP and KN hold, and $k = O(T^{1/2})$, then

(1) under $H^0$:

$$\hat{S}_{nc} = \frac{T^{-1/2} \sum_{t=k+1}^{T} \hat{u}_t\hat{u}_{t-k}}{\Omega(\hat{u}_t\hat{u}_{t-k})} \mathop{\to}^{d} N(0, 1)$$

(2) under $H^1$: the distribution of $|\hat{S}_{nc}|$ diverges as $T \to \infty$
Testing $H_0^0$ Against $H_1^0$

To partition the composite $H^0$ into the null of stationary cointegration against the heteroscedastic alternative, we get

$$u_t = c'V_t w_t + c'\varepsilon_t$$

The null hypothesis is $E(c'V_tV'_tc) = 0$ against the alternative $E(c'V_tV'_tc) > 0$. Consider the statistic

$$S_{hc} = \sum_{t=1}^{T} tu_t^2$$

To find the asymptotic distribution of the test statistic:

Assume the model(4), Assumption LP and KN hold, then

(1) under $H_0^0$:

$$\hat{S}_{hc} = (1/12)^{1/2} \frac{T^{-3/2} \sum_{t=1}^{T} t(\hat{u}_t^2 - \hat{\sigma}_u^2)}{\hat{\Omega}(\hat{u}_t^2 - \hat{\sigma}_t^2)} \overset{d}{\rightarrow} N(0, 1)$$

(2) under $H_1^0$; the distribution of $|\hat{S}_{hc}|$ diverges as $T \rightarrow \infty$

Where $\hat{\sigma}_u^2 = T^{-1} \sum_{t=1}^{T} \hat{u}_t^2$.

Test the null of $I(1)$ against the alternative of HI for any series

Consider any individual process $\{y_t\}$, If $\hat{S}_t$ is calculated using $\Delta y_t - \hat{\Delta}y$ in place of $\hat{u}_t$, then the statistic is denoted as $\hat{S}_{hi}$.

(1) under $H_0$, if the process is $I(1)$:

$$\hat{S}_{hi} \overset{d}{\rightarrow} N(0, 1)$$

(2) under $H_1$, if the process is HI: $|\hat{S}_{hi}|$ diverges.
Chapter 4

Empirical Results

4.1 Data Description

The daily data is from January, 1, 1999 to December, 31, 2003 for three currencies: the New Taiwanese dollar, the Japanese yen, and the Singapore dollar obtained from AREMOS database constructed by TAIWAN ECONOMIC DATA CENTER. \( S_t \) and \( F_{t,k} \) are the levels of the spot exchange rate and the \( k \) - period forward exchange rate determined at time \( t \). We employ the logarithm of all exchange rates which are denoted as \( s_t \) and \( f_{t,k} \) to avoid Siegel’s Paradox.\(^1\) And the trend of the observed spot and forward exchange rates for the three countries are illustrated in Appendix B.

\(^1\)Siegel’s Paradox means that while market efficiency holds, that following two equations must hold simultaneously.

\[
F_t = \mathbb{E}_t(S_{t+1}) \tag{1}
\]

\[
\frac{1}{F_t} = \mathbb{E}_t\left(\frac{1}{S_{t+1}}\right) \tag{2}
\]

However, according to Jensen’s inequality:

\[
\frac{1}{F_t} < \mathbb{E}_t\left(\frac{1}{S_{t+1}}\right) \tag{3}
\]

equation (2) and (3) conflict.
4.2 Unit Roots Tests

Meese and Singleton (1982) addressed that whether or not levels of differences of exchange rates are assumed to be stationary can be lead to substantially different conclusions. And there are strong empirical evidences in three industrialized countries showing that the presence of unit roots processes exists in the exchange rates in levels.

We therefore must confirm spot rates and forward rates are stationary or non-stationary before proceeding analysis which is a regular procedure for market efficiency hypothesis test. It is important because if the exchange rates are non-stationary processes, the analysis must adopt cointegration methods to avoid inducing spurious regression.

The direct methods is the use of unit roots tests to examine time series properties for these variables. There are some common methods for unit roots tests depicted in the former chapter. Since the market efficiency involves with rational expectations which may cause forecast errors, and some literatures, e.g. Dwyer (1992), indicate that forecast errors would emerge autocorrelations and heterogeneity. Besides, the exchange rates are generally influenced with more than one lagged periods. For these purposes, The employment of two different model specification: Augmented Dickey-Fuller Test (ADF) and Phillips-Perron Test (PP) examine the time series properties for the exchange rates. If ADF test which is parametric method and PP test which is semi-parametric method are all in favorable to the $I(1)$ properties of spot rates and forward rates, then the spot and forward rates for three currencies are exactly unit root processes. The principle about choosing the lag truncation parameters for ADF test is to take
the minimum AIC value,\(^2\) and the lag truncation parameters for PP test is to adopt Schwert (1987)'s suggestion.\(^3\)

The empirical results are exhibited in Table 4.1, 4.2, 4.3 and 4.4 can not reject the null hypothesis, that is integrated of order one properties, at 1%, 5% or 10% levels. Either spot rates or forward rates of each currencies, via PP and ADF tests, have strong evidences to present unit root processes in levels, and the existence of stationary processes in differences.

\(^2\)AIC is suggested by Akaike to choose the lagged length \(p\) to minimize, and the formula is:

\[
AIC(p) = \ln |\Sigma_{\epsilon}(p)| + \frac{2pn^2}{T}
\]

where \(n\) is numbers of variables, \(T\) is sample size, and \(p\) is the lagged length parameter which must be estimated.

\(^3\)Schwert (1987) suggested that:

\[
\text{lag length} = \text{INT}[4(n/100)^{1/4}]
\]
### Table 4.1: ADF test for spot exchange rates

<table>
<thead>
<tr>
<th>Currencies</th>
<th>Lag Length</th>
<th>$t$ – Statistics</th>
<th>Lag Length</th>
<th>$t$ – Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Constant</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NT/Dollars</td>
<td>5</td>
<td>-1.08</td>
<td>4</td>
<td>-14.41*</td>
</tr>
<tr>
<td>JY/Dollars</td>
<td>8</td>
<td>-1.53</td>
<td>7</td>
<td>-12.26*</td>
</tr>
<tr>
<td>SD/Dollars</td>
<td>4</td>
<td>-2.01</td>
<td>3</td>
<td>-16.77*</td>
</tr>
<tr>
<td><strong>Constant and Trend</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NT/Dollars</td>
<td>5</td>
<td>-1.44</td>
<td>4</td>
<td>-14.14*</td>
</tr>
<tr>
<td>JY/Dollars</td>
<td>8</td>
<td>-1.32</td>
<td>7</td>
<td>-12.28*</td>
</tr>
<tr>
<td>SD/Dollars</td>
<td>4</td>
<td>-1.70</td>
<td>3</td>
<td>-16.82*</td>
</tr>
</tbody>
</table>

Note: *, **, and *** are significant at 1%, 5%, 10% levels, respectively, to reject $H_0$. The critical values for $t$ – statistics is suggested by Dicky and Fuller studies. The null means series is $I(1)$. The lag length $p$ is to choose the minimum value of $AIC$, $AIC(p) = \ln |\hat{\Sigma}(p)| + \frac{2mp^2}{T}$.

### Table 4.2: ADF test for forward exchange rates

<table>
<thead>
<tr>
<th>Currencies</th>
<th>Lag Length</th>
<th>$t$ – Statistics</th>
<th>Lag Length</th>
<th>$t$ – Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Constant</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NT/Dollars</td>
<td>5</td>
<td>-1.08</td>
<td>4</td>
<td>-14.45*</td>
</tr>
<tr>
<td>JY/Dollars</td>
<td>6</td>
<td>-1.53</td>
<td>5</td>
<td>-14.68*</td>
</tr>
<tr>
<td>SD/Dollars</td>
<td>4</td>
<td>-2.13</td>
<td>3</td>
<td>-16.63*</td>
</tr>
<tr>
<td><strong>Constant and Trend</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NT/Dollars</td>
<td>5</td>
<td>-1.44</td>
<td>4</td>
<td>-14.45*</td>
</tr>
<tr>
<td>JY/Dollars</td>
<td>6</td>
<td>-1.27</td>
<td>5</td>
<td>-14.71*</td>
</tr>
<tr>
<td>SD/Dollars</td>
<td>4</td>
<td>-1.76</td>
<td>3</td>
<td>-16.68*</td>
</tr>
</tbody>
</table>

Note: *, **, and *** are significant at 1%, 5%, 10% levels, respectively, to reject $H_0$. The critical values for $t$ – statistics is suggested by Dicky and Fuller studies. The null means series is $I(1)$. The lag length $p$ is to choose the minimum value of $AIC$, $AIC(p) = \ln |\hat{\Sigma}(p)| + \frac{2mp^2}{T}$. 
### Table 4.3: PP test for spot exchange rates

<table>
<thead>
<tr>
<th>Currencies</th>
<th>Lag Length</th>
<th>$t$ - Statistics</th>
<th>Lag Length</th>
<th>$t$ - Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Constant</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NT/Dollars</td>
<td>7</td>
<td>-0.99</td>
<td>6</td>
<td>-39.73*</td>
</tr>
<tr>
<td>JY/Dollars</td>
<td>7</td>
<td>-1.48</td>
<td>6</td>
<td>-36.26*</td>
</tr>
<tr>
<td>SD/Dollars</td>
<td>7</td>
<td>-2.027</td>
<td>6</td>
<td>-39.68*</td>
</tr>
<tr>
<td><strong>Constant and Trend</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NT/Dollars</td>
<td>7</td>
<td>-1.38</td>
<td>6</td>
<td>-39.71*</td>
</tr>
<tr>
<td>JY/Dollars</td>
<td>7</td>
<td>-1.26</td>
<td>6</td>
<td>-36.28*</td>
</tr>
<tr>
<td>SD/Dollars</td>
<td>7</td>
<td>-1.82</td>
<td>6</td>
<td>-39.73*</td>
</tr>
</tbody>
</table>

Note: *, **, and *** are significant at 1%, 5%, 10% levels, respectively, to reject $H_0$. The critical values for $t$ - statistics is suggested by Dicky and Fuller studies. The null means series is $I(1)$. And lag length = INT[$4(n/100)^{1/4}$].

### Table 4.4: PP test for forward exchange rates

<table>
<thead>
<tr>
<th>Currencies</th>
<th>Lag Length</th>
<th>$t$ - Statistics</th>
<th>Lag Length</th>
<th>$t$ - Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Constant</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NT/Dollars</td>
<td>7</td>
<td>-1.01</td>
<td>6</td>
<td>-40.07*</td>
</tr>
<tr>
<td>JY/Dollars</td>
<td>7</td>
<td>-1.52</td>
<td>6</td>
<td>-39.64*</td>
</tr>
<tr>
<td>SD/Dollars</td>
<td>7</td>
<td>-2.29</td>
<td>6</td>
<td>-39.91*</td>
</tr>
<tr>
<td><strong>Constant and Trend</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NT/Dollars</td>
<td>7</td>
<td>-1.34</td>
<td>6</td>
<td>-40.05*</td>
</tr>
<tr>
<td>JY/Dollars</td>
<td>7</td>
<td>-1.31</td>
<td>6</td>
<td>-37.66*</td>
</tr>
<tr>
<td>SD/Dollars</td>
<td>7</td>
<td>-1.85</td>
<td>6</td>
<td>-39.96*</td>
</tr>
</tbody>
</table>

Note: *, **, and *** are significant at 1%, 5%, 10% levels, respectively, to reject $H_0$. The critical values for $t$ - statistics is suggested by Dicky and Fuller studies. The null means series is $I(1)$. And lag length = INT[$4(n/100)^{1/4}$].
4.3 Johansen Maximum Likelihood Method Test

The evidences from previous section represent $s_t$ and $f_{t,30}$ are both $I(1)$ processes, the sequential motivation is to analysis the long-run relationship between these two variables. There has been amount of studies testing market efficiency hypothesis by Engle and Granger two-step method. Tow-step method, however, has some problems, in particular, in finite samples, the estimated residuals will appear more stationary than the true value, and the Dickey-Fuller critical values will be numerically too small, then leading to reject it’s unit roots properties.

The use of Johansen maximum likelihood method will avoid above-mentioned problems, it can examine the numbers of cointegrating vectors as well. The specification just has two variables, so that there exists at most one cointegrating vector.

The results in table 4.6 and 4.7 reveal one cointegrating vector between the log of future spot rates and the log of one-month forward rates in Taiwan, Japan and Singapore.

From Johansen method, the cointegrating coefficients are also estimated. And the relationship between future spot rates and forward rates in three currencies are described as:

$$s_{t+30}^{NW} = -0.0315 + 1.0003f_{t,30} + u_{t+30}$$
$$s_{t+30}^{JY} = 1.2154 + 0.9915f_{t,30} + u_{t+30}$$
$$s_{t+30}^{SP} = 0.0192 + 0.6612f_{t,30} + u_{t+30}$$
Table 4.5: Normalized Cointegrating Coefficients

<table>
<thead>
<tr>
<th>Currencies</th>
<th>( s_{t+30} )</th>
<th>( f_{t,30} )</th>
<th>constant term</th>
</tr>
</thead>
<tbody>
<tr>
<td>NT/Dollars</td>
<td>1</td>
<td>1.0003</td>
<td>-0.0315</td>
</tr>
<tr>
<td>JP/Dollars</td>
<td>1</td>
<td>0.9915</td>
<td>1.2154</td>
</tr>
<tr>
<td>SG/Dollars</td>
<td>1</td>
<td>0.9912</td>
<td>0.0192</td>
</tr>
</tbody>
</table>
About trace test:

\[
\begin{align*}
H_0 : r \leq 0 & \quad \text{the null form means the cointegrating vector } \leq 0 \\
H_1 : r > 1 & \quad \text{the alternative means the cointegrating vector } > 1 \\
H_0 : r \leq 1 & \quad \text{the null form means the cointegrating vector } \leq 1 \\
H_1 : r > 2 & \quad \text{the alternative means the cointegrating vector } > 1
\end{align*}
\]

Table 4.6: Trace test

<table>
<thead>
<tr>
<th></th>
<th>TW/Dollar</th>
<th>JP/Dollar</th>
<th>SG/Dollar</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Eigenvalue</td>
<td>Trace statistic</td>
<td>at 5% critical value</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TW/Dollar</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H_0 : r \leq 0</td>
<td>0.0542</td>
<td>71.33*</td>
<td>15.41</td>
</tr>
<tr>
<td>H_1 : r &gt; 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H_0 : r \leq 1</td>
<td>0.0011</td>
<td>1.38</td>
<td>3.76</td>
</tr>
<tr>
<td>H_1 : r &gt; 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JP/Dollar</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H_0 : r \leq 0</td>
<td>0.0454</td>
<td>60.37*</td>
<td>15.41</td>
</tr>
<tr>
<td>H_1 : r &gt; 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H_0 : r \leq 1</td>
<td>0.0026</td>
<td>3.19</td>
<td>3.76</td>
</tr>
<tr>
<td>H_1 : r &gt; 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SG/Dollar</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H_0 : r \leq 0</td>
<td>0.0376</td>
<td>50.93*</td>
<td>15.41</td>
</tr>
<tr>
<td>H_1 : r &gt; 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H_0 : r \leq 1</td>
<td>0.0029</td>
<td>3.65</td>
<td>3.76</td>
</tr>
<tr>
<td>H_1 : r &gt; 2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: * and ** indicate to reject H_0 at 1% and 5% significance levels, respectively. The critical values are suggested by Osterwald and Lenam (1996). And r means the numbers of cointegrating vectors.
About max-eigenvalue test:

\[ \begin{align*}
H_0 : r &= 0 \quad \text{the null means the cointegrating vector} = 0 \\
H_1 : r &= 1 \quad \text{the alternative form means the cointegrating vector} = 1 \\
\end{align*} \]

\[ \begin{align*}
H_0 : r &= 1 \quad \text{the null form means the cointegrating vector} = 1 \\
H_1 : r &= 2 \quad \text{the alternative means the cointegrating vector} = 2 \\
\end{align*} \]

<table>
<thead>
<tr>
<th>Table 4.7: Johansen Max-λ test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>TW/Dollar</strong></td>
</tr>
<tr>
<td>( H_0 : r = 0 )</td>
</tr>
<tr>
<td>( H_1 : r = 1 )</td>
</tr>
<tr>
<td>( H_0 : r = 1 )</td>
</tr>
<tr>
<td>( H_1 : r = 2 )</td>
</tr>
<tr>
<td><strong>JP/Dollar</strong></td>
</tr>
<tr>
<td>( H_0 : r = 0 )</td>
</tr>
<tr>
<td>( H_1 : r = 1 )</td>
</tr>
<tr>
<td>( H_0 : r = 1 )</td>
</tr>
<tr>
<td>( H_1 : r = 2 )</td>
</tr>
<tr>
<td><strong>SG/Dollar</strong></td>
</tr>
<tr>
<td>( H_0 : r = 0 )</td>
</tr>
<tr>
<td>( H_1 : r = 1 )</td>
</tr>
<tr>
<td>( H_0 : r = 1 )</td>
</tr>
<tr>
<td>( H_1 : r = 2 )</td>
</tr>
</tbody>
</table>

Note: * and ** indicate to reject \( H_0 \) at 1\% and 5\% significance levels, respectively. The critical values are suggested by Osterwald and Lenam (1996). And \( r \) means the numbers of cointegrating vectors.
4.4 Stochastic Cointegration Empirical Results

Stochastic cointegration i.e. HML model is a nonlinear cointegration method which allows the regression errors to be $I(0)$ plus HI process. Because of it’s more general property, We consider this procedure to re-examine the relationship in Taiwan, Japan and Singapore foreign exchange markets. While simple market efficiency hypothesis holds which means:

$$s_{t+p} = \alpha + \beta f_{t,p} + u_{t+p}$$

where $\alpha = 0$, $\beta = 1$, and the error term $u_{t+p}$ consisting of no information at time $t$ presents white noise process. However the data sampled interval is more finely that the forecast interval, that will cause overlapping problem. If to estimate the above equation by OLS sequentially leading the error term to be serially correlated. Consider HML model (3.20) which permit serial correlation for errors. In addition, Hakkio (1989) demonstrated that cointegration is the necessary condition for market efficiency hypothesis. Thus this model is employed. HML regresses the specification by asymptotic instrument variables method. The AIV estimators of $\alpha$ and $\beta$ are $\hat{\alpha}_k$ and $\hat{\beta}_k$, respectively. The estimation results are in table (4.8). The HAC type covariance matrix has been mentioned in previous chapter.\footnote{\textsuperscript{4}}

\[ \hat{\Sigma}_k = T \left( \sum_{t=k+1}^{T} \mathbf{X}_{t-k} \mathbf{X}'_{t} \right)^{-1} \hat{\Omega} \left( \sum_{t=k+1}^{T} \mathbf{X}_{t} \mathbf{X}'_{t} \right)^{-1} \]

where $\hat{u}_t = y_t - \hat{\alpha}_k - x'_t \hat{\beta}_k$ are the AIV regression residuals. $\hat{\Omega}(.)$ is the estimated long-run variance metric. Let $\mathbf{a}_t$ is any vector time series,

\[ \hat{\Omega}(\mathbf{a}_t) = \hat{\Gamma}_0(\mathbf{a}_t) + \sum_{j=1}^{l} \lambda_j (\nabla_j(\hat{\Gamma}_j(\mathbf{a}_t) + \hat{\Gamma}_j(\mathbf{a}_t)')) \]

\[ \hat{\Gamma}_j(\mathbf{a}_t) = T^{-1} \sum_{t=j+1}^{T} \mathbf{a}_t \mathbf{a}'_{t-j} \]
Note that $k = \text{Int}[T^{1/2}]$, the lag truncation parameter $l = \text{Int}[12(T/100)^{1/4}]$, and the Bartlett window is:

$$\lambda(x) = \begin{cases} 
1 - x, & 0 \leq x \leq 1 \\
0, & x > 1 
\end{cases}$$

Table 4.8: Stochastic cointegration test

<table>
<thead>
<tr>
<th>currencies</th>
<th>$\hat{\alpha}_k$</th>
<th>$\hat{\beta}_k$</th>
<th>$\hat{\alpha}_k$</th>
<th>$\hat{\beta}_k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>NT/Dollars</td>
<td>-0.0019</td>
<td>1</td>
<td>7.7534D-12</td>
<td>2.1449D-12</td>
</tr>
<tr>
<td>JY/Dollars</td>
<td>0.9024</td>
<td>0.8105</td>
<td>0.8051</td>
<td>0.1683</td>
</tr>
<tr>
<td>SD/Dollars</td>
<td>0.0874</td>
<td>0.8462</td>
<td>0.0083</td>
<td>0.0112</td>
</tr>
</tbody>
</table>

Note: The instrument $k = \text{INT}[T^{1/2}] = 35$.

To test the coefficients whether conform to the simple market efficiency sequentially. The result presents in table (4.9).

Table 4.9: Tests Statistics

<table>
<thead>
<tr>
<th>currencies</th>
<th>Statistics for $H_0 : \alpha = 0$</th>
<th>Statistics for $H_0 : \beta = 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>NT/Dollars</td>
<td>1.5186D-14</td>
<td>0</td>
</tr>
<tr>
<td>JY/Dollars</td>
<td>0.7566</td>
<td>0.0319</td>
</tr>
<tr>
<td>SD/Dollars</td>
<td>0.0007</td>
<td>0.0017</td>
</tr>
</tbody>
</table>

Note: * and ** are significant at 5% and 1% levels to reject $H_0$.

The critical values are suggested in standard normal study.

$\lambda(.)$ is one of the usual admissible windows, with the lag truncation parameter $l$. 

48
The comparison between table (4.8) and (4.9), visually, the value $\hat{\alpha}_k$ of the Japanese Yen is closely to 1, however, it’s statistic do not reject the null $H_0 : \hat{\alpha}_k = 0$. That is the puzzle for HML model, $t$ - tests based on $\hat{\alpha}_k$ will be inconsistent under the alternative, and hence will lack power.

The null hypothesis of stochastic cointegration against the alternative of no stochastic cointegration is expressed by

$$H^0 : c' \Pi = 0$$
$$H^1 : c' \Pi \neq 0$$

(1) under $H^0$:

$$\hat{S}_{nc} = \frac{T^{-1/2} \sum_{t=k+1}^{T} \hat{u}_t\hat{u}_{t-k}}{\Omega(\hat{u}_t\hat{u}_{t-k})} \overset{d}{\rightarrow} N(0, 1)$$

(2) under $H^1$: the distribution of $|\hat{S}_{nc}|$ diverges as $T \to \infty$

Under $H^0$, the null hypothesis of stationary cointegration against the heteroscedastic alternative is

$$H^0_0 : E(c' V_t V'_t c) = 0$$
$$H^0_1 : E(c' V_t V'_t c) > 0$$

(1) under $H^0_0$:

$$\hat{S}_{hc} = (1/12)(1/2) T^{-3/2} \sum_{t=1}^{T} t(\hat{u}_t^2 - \hat{\sigma}_u^2) \overset{d}{\rightarrow} N(0, 1)$$

(2) under $H^0_1$: the distribution of $|\hat{S}_{hc}|$ diverges as $T \to \infty$, where $\hat{\sigma}_u^2 = T^{-1} \sum_{t=1}^{T} \hat{u}_t^2$.

Consider any individual process $\{y_t\}$, if $\hat{S}_t$ is calculated using $\Delta y_t - \Delta y$ in place of $\hat{u}_t$, then the statistic is denoted as $\hat{S}_{hi}$.

(1)under $H_0$, if the process is $I(1)$:

$$\hat{S}_{hi} \overset{d}{\rightarrow} N(0, 1)$$
(2) Under $H_1$, if the process is HI:
$|\hat{S}_{hi}|$ diverges.

<table>
<thead>
<tr>
<th>currencies</th>
<th>$\hat{S}_{nc}$</th>
<th>$\hat{S}_{hc}$</th>
<th>$\hat{S}_{hi}(\text{regressands})$</th>
<th>$\hat{S}_{hi}(\text{regressors})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>NT/Dollars</td>
<td>1.1564</td>
<td>-0.0542</td>
<td>0.0247</td>
<td>-0.0262</td>
</tr>
<tr>
<td>JY/Dollars</td>
<td>1.9579</td>
<td>-0.0871</td>
<td>-0.1613</td>
<td>-0.4134</td>
</tr>
<tr>
<td>SD/Dollars</td>
<td>-0.8638</td>
<td>0.0238</td>
<td>0.0836</td>
<td>0.0718</td>
</tr>
</tbody>
</table>

Note: * and ** are significant at 5% and 1% levels to reject $H_0$. The critical values are suggested in standard normal study.

From Table 4.10, the empirical evidences show that spot rates and forward rates in Taiwan, Japan and Singapore present conventionally $I(1)$ processes by $\hat{S}_{hi}$ test. And the long run relationships for spot and forward rates are all stochastic cointegration for three cointegration. Because stochastic cointegration contains heteroscedastic and conventional cointegration, then the employment of statistic $\hat{S}_{hc}$ examines the null hypothesis, that is conventional cointegration, against to the alternative heteroscedastically cointegrated form. The evidence is in favorable to conventional cointegration at 5% or 1% significant levels. Table 4.10 demonstrates that the errors’ variance of cointegrating specifications for the three currencies is constant instead of growing linearly over time.
Chapter 5

Conclusion and Suggestion

Cointegration is necessary condition for market efficiency hypothesis, so that the examination of cointegration to investigate the long-run relationship between the spot rates and forward rates is important. Since there are gradually prosperous trades in foreign exchange markets, agents could hedge, speculate and arbitrage in markets. Thus market efficiency and whether future spot rates could be predicted by forward rates are worth to investigate.

The employment of stochastic cointegration concerning estimation and hypothesis test is a general method encompassing conventional and heteroscedastic cointegration, and is present to be nonlinear form. That is because of the residual for stochastic cointegration has hesteroscedastic integration and stationary process, and the variance of residual perhaps grows over time. However, Stochastic cointegration imposing assumption LP is valid, so that the complex residual will tend to be stochastically trendless, and has tendency to long-run equilibrium over time. Thus the use of this new procedure to re-examine the relationship between
the exchange rates for three currencies, and the result equations are:

\[
\begin{align*}
    s_{t+30}^{NT} &= -0.0019 + f_{t,30}^{NT} + u_{t+30}^{NT} \\
    s_{t+30}^{JP} &= 0.9024 + 0.8104 f_{t,30}^{JP} + u_{t+30}^{JP} \\
    s_{t+30}^{SD} &= 0.0874 + 0.8462 f_{t,30}^{SD} + u_{t+30}^{SD}
\end{align*}
\]

From \( t \) - tests, the AIV estimators of constant terms for three currencies is favorable to the null = 0, and the AIV estimators of slope terms for three currencies in favorable to the null = 0. There is a puzzle that the constant term of the Japanese yen is visually closely to 1, but via \( t \) - tests is favorable to 0. And the authors of MHL model explain the lack of power lead to inconsistency. Conclusively, the future spot rates empirically are stochastic (and conventional) coinegrated with forward rates in Taiwan, Japan, and Singapore.

Stochastic cointegration is a good method except the defect of testing the AIV estimators of constant terms, so that the model must be justified. Besides, all my studies concern zero risk premium, the time-varying premium, however, is encountered frequently in real world. Thus the subjects can be studied later.
Appendix A

The derivation of the asymptotic properties of the AIV estimator.

\[ T^{-1/2} \sum_{t=k+1}^{[Ts]} \begin{pmatrix} \zeta_t \\ \text{vec}(\zeta_t \zeta_{t-k}') \end{pmatrix} \xrightarrow{d} \begin{pmatrix} B_1(s) \\ B_2(s) \end{pmatrix} \]

where \( B_1 \) and \( B_2 \) are independent vector Brownian motions with covariance matrices \( \Omega_{11} \) and \( \Omega_{22} \), respectively.

\[ T^{-1/2} \sum_{t=k+1}^{[Ts]} \text{vec}(\zeta_t \zeta_{t-\tau}') \xrightarrow{d} \tau \otimes B_1(s) + B_2(s) \]

where \( \zeta'(\tau)_t = \tau + \zeta_t \), and \( \tau = (\text{vec}(\Pi)', 0'_{m \times 1}, 0'_{m \times 1})' \).

\[ T^{-1/2} \sum_{t=1}^{T} \eta_t \xrightarrow{d} S'_{\eta} B_1(s) \equiv W(s) \]

\[ T^{-1/2} \sum_{t=k+1}^{[Ts]} V'_t c \xrightarrow{d} S'_{v} B_1(s) \equiv U(s) \]

\[ T^{-1/2} \sum_{t=k+1}^{[Ts]} \text{vec}(V'_t c \text{vec}(\Pi_{x,t-k})) \xrightarrow{d} (S'_{vx} \otimes S'_{v})' [\tau \otimes B_1(s) + B_2(s)] \equiv H(s) \]

with \( \tau = (\text{vec}(\Pi)', 0'_{m \times 1}, 0'_{m \times 1})' \), and

\[ \eta_t = S'_{\eta} \zeta_t \]
\[ V'_t c = S'_{c} \zeta_t \]
\[ \epsilon_t = S'_{\epsilon} \zeta_t \]
\[ \text{vec}(V_{xt}) = S'_{vx} \zeta_t \]

where \( S_{\eta}, S_{c}, S_{\epsilon} \) and \( S_{vx} \) are nonstochastic matrices.
Appendix B

Exchange Rates for Three Countries

- the spot rates of Taiwan
- the forward rates of Taiwan
- the spot rates of Japan
- the forward rates of Japan
- the spot rates of Singapore
- the forward rates of Singapore
References


