CHAPTER 3

THEORETICAL FRAMEWORK AND METHODOLOGY

The first part of this chapter provides the theoretical concept of purchasing power parity and related concepts, such as the law of one price and the real exchange rate. Related econometric concepts are discussed in the second part. The third part summarizes the procedure for testing the validity of PPP. The discussion on data employed in this study will be provided in the last part.

3.1 Theoretical Background

3.1.1 Real Exchange Rate

The real exchange rate is defined by the nominal exchange rate multiplied by the ratio of national price levels, foreign price level divided by domestic price level (Taylor and Taylor, 2004). Let Q denote the real exchange rate, E the nominal exchange rate, P the domestic price level and P^* the foreign price level. The real exchange rate can be written as:

$$Q = E \times \frac{P^*}{P}. \tag{3.1}$$

The real exchange rate is a useful concept and simple mean to gauge the strength of currencies. In other words, the real exchange rate measures the purchasing power of a unit of foreign currency in the foreign economy relative to the purchasing power of an equivalent unit of domestic currency in the domestic economy.

From equation (3.1), taking logarithm on both sides of the equation yields the following equation:

$$q = e + p^* - p \tag{3.2}$$

where q, e, p^* and p denote logarithms of Q, E, P^* and P, respectively.

3.1.2 Law of One Price

According to Hallwood and MacDonald (2000), the *'Law of One Price'* (LOP) states that in competitive markets free of impediments to international trade, such as tariff barriers and transactions costs, no capital flows, the economies are operating at a full employment level, identical traded goods sold in different countries must sell for the same price when their prices are expressed in terms of the same currency. That is, for any good *i*:

$$P_i = EP_i^* \tag{3.3}$$

where P_i is the domestic currency price of good *i*,

P_i^* is the foreign currency price of good *i*, and

E is the nominal exchange rate, defined as the domestic currency price of foreign currency.

There is a sensible mechanism driving the LOP. Suppose for some reason that P_i is greater than EP_i^*; it would be profitable to buy good *i* in a foreign country, ship it to the home country and sell it at higher price. This process will continue until it is no longer profitable. Thus, it can be said that international arbitrage causes the price of every good to be equalized across countries when measured in a common currency.

3.1.3 Absolute and Relative Purchasing Power Parity

The absolute version of PPP is extended from the LOP in a way that the LOP applies to individual commodities (such as commodity *i*), while PPP applies to the general price level, which is a composite of the prices of all the commodities that enter into the reference basket. This can be expressed as:

$$\sum_{i=1}^{N} \alpha_i P_i = E \sum_{i=1}^{N} \alpha_i P_i^* \tag{3.4}$$

where α_i is weights that satisfies $\sum_{i=1}^{N} \alpha_i = 1$.

According to Obsfeld and Rogoff (1996), under the absolute version of PPP, the real exchange rate should equal 1, or at least have a tendency to return quickly to 1 when that long run ratio is disturbed for some reason. It can be said that the real exchange rate exhibits mean reversion to 1.

The relative version requires that the percentage change in exchange rates between two currencies over any period equals the difference between the percentage changes in national price levels. To simplify this, the nominal exchange rate and price ratio are moving together along the time, that is

$$\frac{\sum_{i=1}^{N}\alpha_i P_{it}}{\sum_{i=1}^{N}\alpha_i P_{i,t-1}} = \frac{E_t}{E_{t-1}} \cdot \frac{\sum_{i=1}^{N}\alpha_i P_{it}^{*}}{\sum_{i=1}^{N}\alpha_i P_{i,t-1}^{*}} . \tag{3.5}$$

It is notable that relative PPP pays attention to 'changes' in price and exchange rate, while absolute PPP focuses on 'levels' of price and exchange rate.

Unfortunately, the measures of consumer prices published by national statistical agencies are of little use in constructing a measure of absolute PPP. Since they are typically reported as indices relative to a base year, they only measure the rate of change of the price level from the base year, not its absolute value. In this regard, they can only be used in measuring relative PPP, or, equivalently, changes in real exchange rates. Another failing of standard published CPIs is that they typically involve somewhat different baskets of commodities across countries, though their constructions are usually similar enough that comparisons of changes are still useful.

3.1.4 Weak and Strong Purchasing Power Parity

Another classification of PPP is dividing it into weak type and strong type. According to Pedroni (2004) and Drine and Rault (2007), strong PPP restricts the cointegration coefficient between the nominal exchange rate and relative price levels to be 1.[1]

Weak PPP, however, requires less restriction in that two variables are cointegrated but the cointegrating vector can differ from unity. This version of the PPP hypothesis posits that although the nominal exchange rate and relative price levels may move together over long periods, there are reasons to think that in practice the movements may not be directly proportional, leading to cointegrating slopes

[1]Drine and Rault (2007) has also added that given that the cointegrating vector between the nominal exchange rate and the relative price level is unitary, strong PPP can be investigated by testing whether the real exchange rate is stationary or not.

different from 1. For example, the presence of such factors as international transportation costs, measurement errors, differences in price indices, and differential productivity shocks have been used to explain why the cointegrating slope may differ from unity under the weak version of PPP.

3.2 Related Econometric Concepts

3.2.1 Univariate Unit Root Test

There are various types of univariate unit root test. Two famous tests, Augmented Dickey-Fuller test and Phillips-Perron test are selected in this study. Details of each test are discussed in this section.

Augmented Dickey-Fuller (ADF) Test

One of the univariate versions of unit root test is Augmented Dickey-Fuller (ADF) test. To test the null hypothesis of unit root of y_t process, the following regression has to be estimated:

$$\Delta y_t = \alpha + \delta t + \rho y_{t-1} + \sum_{j=1}^{m} \beta_j \Delta y_{t-j} + u_t \tag{3.6}$$

where Δ is a difference operator, α, δ, ρ and β are parameters to be estimated, t is a time trend and u is an error term following white noise process. The testing hypothesizes are:

H_0 : $\rho = 0$ (non-stationary)

H_1 : $\rho < 0$ (stationary)

If $\rho < 0$ then y_t is stationary process; whereas, $\rho = 0$ implies a unit root of y_t, or y_t is non-stationary. In this step, choosing appropriate lag length, m, is an important issue. Including too many lags reduces the power of test to reject the null hypothesis of unit root due to a loss of degree of freedom. On the other hand, too few lag will not appropriately capture the actual error process, so that ρ and its standard error will not be well estimated (Ender, 1995). The suggestion is to start with a relatively long lag length and pare down the model using t-test or F-test until the lag is significantly different from zero. Another possible method is to use Schwarz Information Criterion (SIC) to select the optimal lag length automatically.

Phillips-Perron (PP) Test

It is worth noting that the distribution theory supporting the Dickey-Fuller test requires the disturbances to be statistically independent and have a constant variance. Phillips and Perron (1988) developed a generalization of the Dickey-Fuller by allowing the errors to be weakly independent and heterogeneously distributed. In this test, Newey-West correction for heteroscedasticity and serial correlation are employed to choose the optimal value of lag truncation.

For the case that the data is generate by a random walk:

$$y_t = y_{t-1} - u_t \qquad (3.7)$$

with u_t is identically and independently distributed with zero mean and variance σ^2. The regression model includes a constant term:

$$y_t = \alpha + \rho y_{t-1} - u_t , \qquad (3.8)$$

then the Phillips-Perron Z_t statistic is constructed as

$$Z_t = t_T \left(\frac{\gamma_0}{\lambda^2} \right)^{1/2} - \frac{T \hat{\sigma}_{\hat{\rho}_T} (\lambda^2 - \gamma_0)}{2 \lambda s_T} , \qquad (3.9)$$

where t_T is t-ratio of $\hat{\rho}_T$, γ_0 is a consistent estimate of the error variance σ^2, λ^2 is an estimator of the residual spectrum at frequency zero, $\hat{\sigma}_{\hat{\rho}_T}$ is the OLS standard error for $\hat{\rho}_T$ and s_T^2 is the OLS estimate of the variance of u_t. Phillips and Perron (1988) proved that under the null hypothesis of unit root ($\rho = 1$), the limiting distribution of Z_t statistic is the same as limiting distribution of ADF test based on *t*-statistic[2].

3.2.2 Cointegration Test

To test for long run relationship between two or more time series variables, the cointegration tests are introduced. In this study, Engle-Granger and Johansen tests for cointegration are selected. This section presents the details of each test.

---

[2]Critical values for the Phillips-Perron Z_t test can be obtained from Hamilton (1994).

<u>Engle-Granger Cointegration Test</u>

According to Engle and Granger (1987), the components of vector $x_t = (x_{1t}, x_{2t}, \ldots, x_{nt})'$ are said to be *cointegrated of order d, b*, denoted by $x_t \sim CI(d,b)$ if

(i) All components of x_t are integrated of order *d*.

(ii) There exists a vector $\beta = (\beta_1, \beta_2, \ldots, \beta_n)$ such that linear combination $\beta x_t = \beta_1 x_{1t} + \beta_2 x_{2t} + \ldots + \beta_n x_{nt}$ is integrated of order $(d-b)$, where $b > 0$.

The vector β is called the *cointegrating vector.*

From this definition, there are some crucial points to remark. Firstly, cointegration refers to a *linear* combination of non-stationary variables. It is possible that nonlinear long run relationships exist. However, only the linear cointegration can be tested by this technique. Secondly, all variables must be integrated of the same order. If the variables are integrated of different orders, they cannot be cointegrated. Thirdly, if x_t has n components, there may be at most $n-1$ cointegrating vectors. The number of cointegrating vectors is called the cointegrating rank of x_t. It should also note that the term 'cointegration' normally refers to the case in which variables are CI(1,1).

Engle and Granger (1987) proposed a straightforward test to investigate whether two I(1) variables are cointegrated of order CI(1,1). Before the two-step is applied, it is required to pretest the variables for their order of integration. If the variables are integrated of different orders, it is possible to conclude that they are not cointegrated. In this step, unit root test can take the role to identify if the series is stationary [~I(0)] or not [~I(1)].

Once the identical unit roots of two series are confirmed, the two-step test can be conducted. The first step is to estimate the long run equilibrium relationship in the following regression model.

$$x_t = \alpha + \beta y_t + u_t \qquad (3.10)$$

By applying OLS to estimate the model in equation (3.10), the series $\{\hat{u}_t\}$ is obtained. $\{\hat{u}_t\}$ is the series of the estimated residuals of the long run relationship. If these deviations from long run equilibrium are found to be stationary, the $\{x_t\}$ and $\{y_t\}$ are cointegrated. Therefore, in the second step, the cointegration can be examined by testing for stationarity of $\{\hat{u}_t\}$.

Engle and Granger (1987) proposed seven statistics to test for cointegration. Two statistics are selected in this study. The first statistic is Cointegrating Regression Durbin-Watson (CRDW) which is indeed the standard DW statistic of the cointegrating regression equation (3.10).

The testing hypothesizes are:

$H_0: \; CRDW \to 0 \;\; \equiv \;\; \{\hat{u}_t\} \sim \mathrm{I}(1)$ (no cointegration)

$H_1: \; CRDW \to 2 \;\; \equiv \;\; \{\hat{u}_t\} \sim \mathrm{I}(0)$ (cointegration)

The null hypothesis of CRDW approaching zero implies that cointegration is rejected. Otherwise, if the CRDW approaches to two, the long run relationship between two variables exists.

Another statistic selected is ADF statistic. The ADF statistic is used to verify whether $\{\hat{u}_t\}$ is stationary or not. There is no need to include an intercept term in ADF regression since $\{\hat{u}_t\}$ is a residual from a regression equation. Thus, the regression equation for ADF statistic is

$$\Delta\hat{u}_t = \rho\hat{u}_{t-1} + \sum_{j=1}^{m}\gamma_i\Delta\hat{u}_{t-j} + \varepsilon_t \qquad (3.11)$$

based on these hypothesizes

$H_0: \; \rho = 0 \;\; \equiv \;\; \{\hat{u}_t\} \sim \mathrm{I}(1)$ (no cointegration)

$H_1: \; \rho < 0 \;\; \equiv \;\; \{\hat{u}_t\} \sim \mathrm{I}(0)$ (cointegration)

The critical values of both statistics for the case of two variables are provided by Engle and Granger (1987) while those of multiple variables are available in Engle and Yoo (1987).

In order to reject the null hypothesis of no cointegration, equivalently, confirm the existence of cointegration, the ADF statistic has to have greater magnitude (in absolute value) than the critical values.

Johansen Multivariate Cointegration Test

Johansen (1988) and Johansen and Juselius (1990) proposed another approach to test for cointegration. While Engle-Granger test relies on testing the residuals of equilibrium regression, this method is based mainly on relationship between the rank of matrix and its characteristic roots. Moreover, this process can determine the number of cointegrating vectors among interested variables.

The Johansen techniques applied maximum likelihood methods for the analysis of cointegration in Gaussian vector autoregressive (VAR) models which allow for constant term and seasonal dummies. For simplicity, the deterministic part is excluded from VAR model. It should also note that this method is designed to handle I(1) and I(0) variables. If some of the series are I(2), standard Johansen approach cannot be applied.

Consider the following VAR model of order *k*:

$$X_t = \Pi_1 X_{t-1} + \Pi_2 X_{t-2} + \ldots + \Pi_k X_{t-k} + u_t \quad (t = 1,\ldots,T) \qquad (3.12)$$

where X_t is $(n \times 1)$ vector of variables containing in the VAR model and $X_{-k+1},\ldots,X_0$ are fixed

Π_i is $(n \times n)$ matrix of coefficients

u_t is $(n \times 1)$ vector of residuals and $u_1,\ldots,u_t$ are independent Gaussian variables with mean zero and variance matrix Λ.

The model in equation (3.12) can be rewritten as

$$\Delta X_t = \Gamma_1 \Delta X_{t-1} + \ldots + \Gamma_{k-1} \Delta X_{t-k+1} + \Pi X_{t-k} + u_t \qquad (3.13)$$

where $\Gamma_i = -I + \Pi_1 + \Pi_2 + \ldots + \Pi_i$, ($i = 1,\ldots,k-1$)

$\Pi = -I + \Pi_1 + \Pi_2 + \ldots + \Pi_k$.

Equation (3.13) is expressed as a traditional first difference VAR model except for the term ΠX_{t-k}. This is the key point of this method. The matrix Π is the long run impact matrix. The main purpose of this method is to investigate whether the matrix Π contains any information about the long run relationships between the variables in the data vector X_t. This can be achieved by considering the rank of matrix Π. Since Π is $(n \times n)$ matrix, the rank of Π can be at most *n*. There are three possible cases:

(a) If rank(Π) = *n*, i.e. the matrix Π has full rank, the vector process X_t is stationary.

(b) If rank(Π) = 0, i.e. the matrix Π is null, the equation (3.13) is the usual VAR model in first difference. The process X_t is non-stationary and there is no cointegrating vector.

(c) If rank(Π) = *r*, $0 < r < n$, then there are *r* cointegrating vectors and equation (3.13) can be interpreted as an error correction model. Moreover, there are $n \times r$ matrices α and β such that $\Pi = \alpha\beta'$ where α is a matrix of error correction coefficients (speed of adjustment) and β is a matrix of cointegrating vectors (long run coefficients). The cointegrating vectors β have the property that $\beta' X$ is stationary even though X_t itself is non-stationary.

Details of procedure can be found in the original papers. The concept of finding the rank of Π is described here. Rank of an $(n \times n)$ matrix is equal to the number of linearly independent eigenvectors of the matrix. Therefore, to obtain the rank of Π, we need to find all characteristic roots of Π. If rank of Π is zero, all these characteristic roots are zero and this implies no cointegration. Otherwise, if there exist non zero roots, the roots can be ordered such that $\lambda_1 > \lambda_2 > ... > \lambda_n$. In practice, we can obtain only the estimates of Π ($\hat{\Pi}$) and the characteristic roots ($\hat{\lambda}_i$).

The following two statistics are proposed to test for the number of cointegrating vectors, *r*.

(i) Trace Test (λ_{trace})

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

H_0 : There are at most *r* cointegrating vectors.

H_1 : There are more than *r* cointegrating vectors.

(ii) Maximal Eigenvalue Test ($\lambda_{\max}$)

$$\lambda_{\max}(r, r+1) = -T \ln(1 - \hat{\lambda}_{r+1})$$

H_0 : Number of cointegrating vector is *r*.

H_1 : Number of cointegrating vector is *r+1*.

If $\hat{\lambda}_i = 0$, $\ln(1 - \hat{\lambda}_i) = \ln(1) = 0$. Therefore, it is obvious that λ_{trace} equals zero when all $\hat{\lambda}_i = 0$. If $\hat{\lambda}_i \neq 0$, $\ln(1 - \hat{\lambda}_i)$ is negative and hence λ_{trace} is positive. Like λ_{trace}, the value of $\lambda_{\max}$ is small if the estimated characteristic root, $\hat{\lambda}_{r+1}$, is closed to zero. The critical values for these statistics are provided in Johansen and Juselius (1990).

3.2.3 Panel Unit Root Test

Various versions of panel unit root test are developed to encounter the low power of univariate unit root tests. Levin, Lin and Chu (2002) and Im, Pesaran and Shin (2003) tests are used in this study.

Levin, Lin and Chu (LLC) test

Levin and Lin (1993) and Levin *et al.* (2002) developed a panel unit root test that has more power than univariate unit root tests by imposing the same first order autoregressive coefficient (ρ) on all series, but allowing for individual (industries, regions or countries) specific intercepts (α_{0i}) and time trends (α_{1i}). Furthermore, not only the error variance is permitted to arbitrarily vary across individuals, but also the pattern of high-order serial correlation.

To determine whether the process y_{it} is unit root or stationary for each individual in the panel, the following assumptions are required.

(a) Assume that y_{it} is generated by one of the following models:

Model 1: $\Delta y_{it} = \rho y_{i,t-1} + u_{it}$

Model 2: $\Delta y_{it} = \alpha_{0i} + \rho y_{i,t-1} + u_{it}$

Model 3: $\Delta y_{it} = \alpha_{0i} + \alpha_{1i} t + \rho y_{i,t-1} + u_{it}$, where $-2 < \rho \leq 0$ for $i = 1,...,N$.

(b) u_{it} is independently distributed across individuals and follow a stationary invertible ARMA process for each individual,

$$u_{it} = \sum_{j=1}^{\infty} \theta_{ij} u_{i,t-j} + \varepsilon_{it}$$

(c) $E(u_{it}^4) < \infty$; $E(\varepsilon_{it}^2) \geq B_\varepsilon > 0$; and $E(u_{it}^2) + 2\sum_{j=1}^{\infty} E(u_{it} u_{i,t-j}) < B_u < \infty$ for all $i = 1,...,N$ and $t = 1,...,T$.

For the model specification, Levin *et al.* (2002) have noted that the omission of a deterministic element (eg. an intercept or time trend) when it is indeed present results in inconsistency of test. In contrast, the inclusion of irrelevant deterministic element reduces the power of the test. Therefore, model 2 and model 3 will be used in this study.

This method tests the null hypothesis that all individuals in panel have integrated time series (unit roots) versus the alternative hypothesis that each individual time series is stationary. In other words, the alternative hypothesis states that *all* series in the panel have to be stationary. This hypothesis will be relaxed later.

Levin *et al.* (2002) suggested a three-step procedure to implement the panel unit root test. Step 1 is to perform separate ADF regressions for each individual in the panel and generate two orthogonalized residuals. Ratio of long run to short run standard deviations for each individual is estimated in step 2. In the final step, the panel statistics, t-statistic (t_ρ) and adjusted t-statistic (t_ρ^*) are computed. Though the t-statistic (t_ρ) for model 2 diverges to negative infinity, the adjusted t-statistic (t_ρ^*) asymptotically follows the standard normal distribution under the null hypothesis. It is also notable that the LLC test is particularly useful for panels of moderate size, between 10 and 250 individuals with 25-250 time series observations per individual. If the time series dimension of the panel is very large, the univariate unit root test will have sufficient power[3].

Im, Pesaran and Shin (IPS) test

The Im *et al.* (2003) test relaxes the assumption of identical first order autoregressive coefficients of Levin *et al.* (2002). This approach allows the heterogeneous of ρ or, equivalently, allows the real exchange rate of each country reverts to its mean at different rates.

IPS test evaluates the following model

$$\Delta y_{it} = \alpha_i + \rho_i y_{i,t-1} + u_{it}\ . \qquad (3.14)$$

based on these hypothesizes

H_0 : $\rho_i = 0$ for all $i = 1,...,N$

H_1 : $\rho_i < 0$ for some i, $i = 1,...,N$.

In contrast with the LLC test, the alternative hypothesis of IPS test allows for some (but not all) of the individual series to have unit root.

[3]In the case that the cross-section dimension is very large, Levin *et al.* (2002) suggested the existing panel data procedures of MaCurdy (1982), Hsiao (1986), Holtz-Eakin *et al.* (1988) and Breitung and Meyer (1991) as the appropriate tools.

Im *et al.* (2003) assume that u_{it} is independently and normally distributed with zero mean and finite heterogeneous variance, σ_i^2. Moreover, under the alternative hypothesis, the fraction of the individual processes that are stationary is also assumed to be non-zero, i.e. $\lim_{N\to\infty}(N_1/N)=\delta, 0<\delta\le 1$ since it is necessary for the consistency of the panel unit root tests.

This method pools N separate cross-section unit root test statistics to evaluate the following equation

$$\Delta y_{it} = \alpha_i + \rho_i y_{i,t-1} + \sum_{j=1}^{p_i} \beta_{ij} \Delta y_{i,t-j} + u_{it} \tag{3.15}$$

Let $\tilde{t}_{i,T}$ denote the *t*-statistics used to evaluate the null hypothesis of the unit root in the standard individual ADF. The *t*-bar statistic is calculated from

$$\tilde{t}\text{-}bar_{NT} = \frac{1}{N}\sum_{i=1}^{N} \tilde{t}_{iT} \ . \tag{3.16}$$

Following this setting, Im *et al.* (2003) showed that for $T>5$ and the individual statistics, $\tilde{t}_{i,T}$, $i=1,...,N$, are identically and independently distributed with finite second order moments, i.e. mean and variance, as $N\to\infty$, the standardized *t*-bar statistic ($Z_{\tilde{t}\text{-}bar}$), defined as

$$Z_{\tilde{t}\text{-}bar} = \frac{\sqrt{N}\left(\tilde{t}\text{-}bar_{NT} - E(\tilde{t}_T)\right)}{\sqrt{Var(\tilde{t}_T)}} \tag{3.17}$$

converges to a standard normal variate, N(0,1), under the null hypothesis.

For the case of serial correlation in u_{it}, Im *et al.* (2003) also derived the standardized *t*-bar statistic, demeaned-IPS statistic, denoted by $W_{t\text{-}bar}$.

$$W_{t\text{-}bar} = \frac{\sqrt{N}\left(t\text{-}bar_{NT} - N^{-1}\sum_{i=1}^{N} E[t_{iT}(p_i,0)\,|\,\rho_i=0]\right)}{\sqrt{N^{-1}\sum_{i=1}^{N} Var[t_{iT}(p_i,0)\,|\,\rho_i=0]}} \tag{3.18}$$

Under the null of nonstationary, $W_{t\text{-}bar}$ converges to a standard normal distribution where N and $T\to\infty$ such that $N/T\to k$, for a finite non-negative constant k.

Although $Z_{\tilde{t}\text{-}bar}$ and $W_{t\text{-}bar}$ are asymptotically equivalent, results from simulations indicated that $W_{t\text{-}bar}$ statistic performs much better than $Z_{\tilde{t}\text{-}bar}$ statistic. This is because $W_{t\text{-}bar}$ statistic takes the underlying ADF orders into account in computing the mean and the variance adjustment factors.

Lopez and Papell (2007) mentioned about the power of LLC and IPS tests. Given that the maintained hypothesis of homogeneity is correct, the LLC test has greater power than the IPS test if $\rho < 0$. Both tests are correctly sized if $\rho = 0$. On the contrary, if the maintained hypothesis of homogeneity is not correct, one possibility is that there is a mixed panel; some of the ρ's are less than 0 and some of the ρ's are equal to 0. In that case, the IPS test has greater power than the LLC test.

Combined Individual Tests (Fisher Test and Choi Test)

Instead of constructing panel test statistics by combining the test statistics, another approach is to combine the observed significant levels (*p*-values) from individual tests. There are several literature on this issue[4], however, the Fisher test based on the sum of the log-*p*-values has been widely recommended (Maddala and Wu, 1999).

Maddala and Wu (1999) used Fisher's result to propose an alternative approach to test for unit root and cointegration in panel data by combining tests from individual cross-sections to obtain a test statistic for the full panel. This test is very attractive due to its simplicity and its robustness to statistic choices, lag length and sample size.

Let p_i be the significance level (p-value) from any individual unit root test (or cointegration test) for cross-section *i*. It is assumed that the test statistics are continuous and the significance level p_i $(i = 1, 2, \ldots, N)$ are independent uniform (0,1) variables. Under the null of unit root for all cross-sections (or no cointegration relation in panel) and the crucial assumption of cross-sectional independence, the combination of *p*-values proposed by Maddala and Wu (1999) defined as

$$P_{MW} = -2\sum_{i=1}^{N}\log(p_i) \tag{3.19}$$

[4]According to Maddala and Wu (1999), this issue can be found in Tippett (1931) and Fisher (1932). Different tests are reviewed in Hedges and Olkin (1985).

Tippett (1931) suggested using the distribution of the smallest of the *p*-values, π_i. Furthermore, there have been several other suggestions about the *p*-values combinations and 16 of them are listed in Becker (1997).

has a Chi-square distribution with $2N$ degrees of freedom, when T tends to infinity and N is fixed.

Besides the P_{MW} statistic, Choi (2001) suggested a similar standardized statistic:

$$Z = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \Phi^{-1}(p_i) \tag{3.20}$$

where $\Phi^{-1}(\bullet)$ is the standard normal cumulative distribution function. Under the null hypothesis of unit root and the similar assumptions imposed on p_i, Z converges to a standard normal distribution.

Pesaran Test

Unlike the unit root tests of the first generation which assume that the individual time series in panel are cross-sectionally independently distributed, the test of Pesaran (2003) allows for cross-sectional dependence in series. Instead of dealing with cross-sectional dependence by demeaning the series like demeaned-IPS test, this test augments the standard DF or ADF regressions with the cross section averages of lagged levels and first-differences of the individual series. Then the standard panel unit root tests can be calculated based on the simple averages of the individual cross-sectionally augmented ADF statistics, denoted CADF.

Let y_{it} be the observation on the i^{th} cross section unit at time *t* and suppose that it is generated according to the following simple dynamic linear heterogeneous panel data model

$$y_{it} = (1-\phi_i)\mu_i + \phi_i y_{i,t-1} + u_{it}, \quad i = 1,\ldots,N;\ t = 1,\ldots,T \tag{3.21}$$

$$u_{it} = \gamma_i f_t + \varepsilon_{it} \tag{3.22}$$

where y_{i0}, the initial value, is given

f_t is the unobserved common effect

ε_{it} is the individual-specific.

Equation (3.21) and (3.22) can be written as

$$\Delta y_{it} = \alpha_i + \rho_i y_{i,t-1} + \gamma_i f_t + \varepsilon_{it} \tag{3.23}$$

where $\alpha_i = (1-\phi_i)\mu_i$, $\rho_i = -(1-\phi_i)$ and $\Delta y_{it} = y_{it} - y_{i,t-1}$.

The unit root hypothesizes to test are:

$H_0:\ \ \rho_i = 0$ for all *i = 1,...,N*

$H_1:\ \ \rho_i < 0$ for *i = 1,...,N₁*

$\rho_i = 0$ for *i* = $N_1+1, N_1+2, \ldots, N$.

Similar to the IPS test, the assumption that $\lim_{N\to\infty}(N_1/N) = \delta$, $0 < \delta \le 1$ is required for the consistency of the panel unit root test. Under a set of assumptions of Pesaran (2003) and if the residuals are serially uncorrelated, $\bar{y}_{t-1} = (1/N)\sum_{i=1}^{N} y_{i,t-1}$ and $\Delta\bar{y}_t = (1/N)\sum_{i=1}^{N}\Delta y_{i,t}$ are sufficient for asymptotically filtering out the effects of the unobserved common factor, f_t. Therefore, the unit root test can base on the t-statistic of the OLS estimate of ρ_i in the following cross-sectionally ADF, denoted CADF, regression:

$$\Delta y_{i,t} = \alpha_i + \rho_i y_{i,t-1} + c_i\bar{y}_{t-1} + d_i\Delta\bar{y}_t + u_{it} \qquad (3.24)$$

The distribution and critical values of the CADF statistic are derived and computed. Pesaran (2003) suggested that this CADF statistic can be applied for univariate unit root test when information on the cross section average, $\bar{y}_t$, is available.

Aside from CADF statistic, a truncated version of the CADF statistic, denoted CADF*, is suggested in order to avoid undue influence of extreme outcomes in case of small T samples. The value of CADF is given as:

$$CADF_{if}^{*} = \begin{cases} -K_1 & if \quad CADF_{if} \le -K_1 \\ CADF_{if} & if \quad -K_1 \le CADF_{if} \le K_2 \\ K_2 & if \quad CADF_{if} \ge K_2 \end{cases} \qquad (3.25)$$

where K_1 and K_2 are positive constants such that the probability that $CADF_{if}$ falls in the interval $[-K_1, K_2]$ is sufficiently large. The simulated values of K_1 and K_2 for each case are reported in Pesaran (2003).

The panel unit root test is now constructed based on the CADF or CADF* statistic. A cross-sectionally augmented IPS test (CIPS or CIPS*) based on the average of individual CADF or CADF* statistics are as followed.

$$CIPS = \frac{1}{N}\sum_{i=1}^{N} CADF_{if} \tag{3.26}$$

$$CIPS^{*} = \frac{1}{N}\sum_{i=1}^{N} CADF_{if}^{*} \tag{3.27}$$

For the case of serially correlated errors, it is shown that the individual CADF statistics have similar asymptotic distribution as in the serially uncorrelated case, given that the CADF regressions are augmented with the lagged changes of the individual series and the lagged changes of the cross section averages. Moreover, $CADF_i$ statistics not depend on factor loadings. However, they are asymptotically correlated due to their dependence on the common factor. Consequently, the standard central limit theorems do not apply to the CIPS statistic. Fortunately, it is shown that the limit distribution of the CIPS*, the truncated version of CIPS, exists and is free of nuisance parameter. Therefore, CIPS* statistic can be used to test the hypothesis of unit root in panel data. The critical values of CIPS and CIPS* statistics are given in Pesaran (2003).

3.2.4 Panel Cointegration Test

Several cointegration tests for panel data are developed to extend the use of time series cointegration test. Among these tests, the tests proposed by Pedroni (1997, 2004) and Kao (1999) are utilized in this study.

Pedroni Test

Pedroni (1997, 2004) developed several residual-based test statistics for heterogeneous panel cointegration, allowing individual specific fixed effects, deterministic trends, as well as individual specific slope coefficients. If the underlying data generating process (DGP) is assumed to permit individual members of the panel to differ in whether or not they are cointegrated, then the testing hypothesizes are as follow:

H_0 : "All of the individuals of the panel are not cointegrated."

H_1 : "A significant portion of the individuals are cointegrated."

The regression model is

$$y_{it} = \alpha_i + \delta_i t + \beta_i X_{it} + e_{it} \tag{3.28}$$

for $i = 1,\ldots,N$, $t = 1,\ldots,T$, where X_{it} is an m-dimensional column vector for each member i and β_i is an m-dimensional row vector for each member i. The variables y_{it} and X_{it} are assumed to be integrated of order one, denoted I(1), for each member i of the panel, and under the null of no cointegration the residual e_{it} will also be I(1).

Let the partitioned vector $z'_{it} \equiv (y_{it}, X'_{it})$ such that the true process z_{it} is generated as $Z_{it} = Z_{it-1} + \xi_{it}$, for $\xi'_{it} \equiv (\xi^{y}_{it}, \xi^{X'}_{it})$. The $(m+1)\times(m+1)$ asymptotic covariance matrix is given by $\Omega_i \equiv \lim_{T\to\infty} E\left[T^{-1}(\sum_{t=1}^{T}\xi_{it})(\sum_{t=1}^{T}\xi'_{it})\right]$. The following conditions are assumed to hold with regard to the time series dimension.

(a) The process $\xi'_{it} \equiv (\xi^{y}_{it}, \xi^{X'}_{it})$ satisfies $1/\sqrt{T}\sum_{t=1}^{[Tr]}\xi_{it} \Rightarrow B_i(\Omega_i)$, for each member i as $T \to \infty$, where $B_i(\Omega_i)$ is vector Brownian motion with asymptotic covariance Ω_i such that the $m \times m$ lower diagonal block $\Omega_{22i} > 0$. In other word, this condition states that the standard functional central limit theorem is assumed to hold individually for each member series as T grows large.

(b) The individual processes are assumed to be independent and identically distributed (i.i.d.) cross-sectionally, so that $E[\xi_{it}, \xi'_{js}] = 0$ for all $s,t,i \neq j$.

Following Pedroni (2004), the statistics are constructed by first estimating the hypothesized cointegrating relationship separately for each member of the panel and then pooling the resulting residuals when constructing the panel tests for the null of no cointegration.

Let $\tilde{e}_{it} = (\Delta\hat{e}_{it}, \hat{e}_{it-1})'$, $A_i = \sum_{t=1}^{T}\tilde{e}_{it}\tilde{e}'_{it}$ where $\hat{e}_{it}$ is estimated from a model based on the regression in equation (3.28). The test statistics are calculated as follows:

Panel-variance ratio (Panel-v), $Z_{\hat{v}_{NT}}$

$$Z_{\hat{v}_{NT}} \equiv \hat{L}_{11}^{2}\left(\sum_{i=1}^{N} A_{22i}\right)^{-1} \tag{3.29}$$

Panel-rho, $Z_{\hat{\rho}_{NT^{-1}}}$

$$Z_{\hat{\rho}_{NT^{-1}}} \equiv \left(\sum_{i=1}^{N} A_{22i} \right)^{-1} \sum_{i=1}^{N} \left(A_{21i} - T\hat{\lambda}_i \right) \tag{3.30}$$

Panel-t (parametric), or Panel-ADF, $Z_{\hat{t}_{NT}}$

$$Z_{\hat{t}_{NT}} \equiv \left(\tilde{\sigma}_{NT}^2 \sum_{i=1}^{N} A_{22i} \right)^{-1/2} \sum_{i=1}^{N} \left(A_{21i} - T\hat{\lambda}_i \right) \tag{3.31}$$

where $\hat{\mu}_{it} = \hat{e}_{it} - \hat{\rho}_i \hat{e}_{it-1}$, $\hat{\lambda}_i = T^{-1} \sum_{s=1}^{K} w_{sK} \sum_{t=s+1}^{T} \hat{\mu}_{it} \hat{\mu}_{i,t-s}$ for some choice of lag window $w_{sK} = 1 - s/(1-K)$, $\hat{s}_i^2 = T^{-1} \sum_{t=2}^{T} \hat{\mu}_{it}^2$, $\hat{\sigma}_i^2 = \hat{s}_i^2 + 2\hat{\lambda}_i$, $\tilde{\sigma}_{NT}^2 \equiv N^{-1} \sum_{i=1}^{N} \hat{\sigma}_i^2$, and $\hat{L}_{11}^2 = N^{-1} \sum_{i=1}^{N} \hat{L}_{11i}^2$ where $\hat{L}_{11i}^2 = \hat{\Omega}_{11i} - \hat{\Omega}'_{21i} \hat{\Omega}_{22i}^{-1} \hat{\Omega}_{21i}$ such that $\hat{\Omega}_i$ is a consistent estimator of Ω_i.

The asymptotic distributions of the statistics and their critical values are available in Pedroni (2004) and Pedroni (1999), respectively.

<u>Kao test</u>

The test of Kao (1999) follows the same basic approach as the Pedroni tests, but it assumes homogeneous coefficients on the first-stage regressors. Kao (1999) studied a spurious regression using Least-Square Dummy Variable (LSDV) in panel data and proposed a residual-based test for cointegration regression in panel data. He studied Dickey-Fuller (DF) test and an augmented Dickey-Fuller (ADF) test to test the null of no cointegration. Since his simulation results suggested that ADF test clearly dominates the DF tests when σ is large, only ADF test will be discussed here.

Let $y_{it} = \sum_{s=1}^{t} u_{is}$ and $x_{it} = \sum_{s=1}^{t} \varepsilon_{is}$ where u_{it} and ε_{it} are assumed to be independent across *i*; the spurious LSDV regression model is

$$y_{it} = \alpha_i + \beta X_{it} + e_{it} \tag{3.32}$$

for i = *1,…,N*, *t* = *1,…,T*.

The ADF test can be applied to the residuals using the following regression:

$$\hat{e}_{it} = \rho \hat{e}_{it-1} + \sum_{j=1}^{p} \varphi_j \Delta \hat{e}_{it-j} + \upsilon_{itp} \,, \tag{3.33}$$

where p is chosen so that the residual υ_{itp} are serially uncorrelated and $\hat{e}_{it}$ is the estimate of e_{it} from equation (3.32).

With the null hypothesis of no cointegration, the ADF test statistic can be constructed as

$$ADF = \frac{t_{ADF} + \sqrt{6N}\hat{\sigma}_{\upsilon} / 2\hat{\sigma}_{0u}}{\sqrt{\hat{\sigma}_{0\upsilon}^{2} / 2\hat{\sigma}_{\upsilon}^{2} + 3\hat{\sigma}_{\upsilon}^{2} / 10\hat{\sigma}_{0\upsilon}^{2}}} \,, \tag{3.34}$$

where t_{ADF} is t-statistic of ρ in (3.33). Kao (1999) has shown that the asymptotic distribution of ADF statistic will converge to a standard normal distribution by the sequential limit theory.

Fisher Test

Similar to unit root test, the idea of combining p-values of individual test can also be applied to the cointegration test of panel data. Fisher type test suggested by Maddala and Wu (1999) is applied to the test of Johansen under the cross-sectional independence assumption. The statistic and its distribution are the same as P_{MW} described in previous section.

3.2.5 Test of Parameter Constancy

Swamy (1970) proposed a procedure for estimating random coefficient regression which the parameters are allowed to vary over the cross-sectional units. This model allows both random intercept and slope parameters that vary around common means. The random parameters can be considered outcomes of a common mean plus an error term, representing a mean deviation for each individual.

The model to estimate is as follow:

$$\underset{(T\times 1)}{\underline{y}_i} = \underset{(T\times \Lambda)}{X_i} \; \underset{(\Lambda\times 1)}{\underline{\beta}_i} + \underset{(T\times 1)}{\underline{u}_i} \,. \tag{3.35}$$

Allowing slope coefficients to vary across i, equation (3.35) can be written as

$$\begin{bmatrix} \underline{y}_1 \\ \underline{y}_2 \\ \vdots \\ \underline{y}_N \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix} \underline{\bar{\beta}} + \begin{bmatrix} X_1 & 0 & \dots & 0 \\ 0 & X_2 & \dots & 0 \\ \vdots & \vdots & . & \vdots \\ 0 & 0 & \cdots & X_N \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_N \end{bmatrix} + \begin{bmatrix} \underline{u}_1 \\ \underline{u}_2 \\ \vdots \\ \underline{u}_N \end{bmatrix}, \qquad (3.36)$$

or more compactly as

$$\underline{y} = X\underline{\bar{\beta}} + D\underline{\delta} + \underline{u} \qquad (3.37)$$

Efficient estimators of equation (3.37) and several tests are proposed in the original paper. One useful test, which will be applied in this study, is the test of equality between fixed coefficient vectors in N relations with heteroskedastic disturbances. Swamy (1970) has suggested to test whether the coefficient vectors $\underline{\beta}_i$ ($i = 1,...,N$) are fixed and are all equal before estimating any model under his setting.

Based on these hypothesizes:

H_0 : $\underline{\beta}_1 = \underline{\beta}_2 = ... = \underline{\beta}_N = \underline{\beta}$

H_1 : Otherwise,

the homogeneity statistic can be constructed as follow:

$$H_\beta = \sum_{i=1}^{N} \frac{(\underline{b}_i - \hat{\underline{\beta}})' X_i' X_i (\underline{b}_i - \hat{\underline{\beta}})}{s_{ii}} \qquad (3.38)$$

where $\hat{\underline{\beta}} = \left[\sum_{i=1}^{N} \frac{X_i' X_i}{s_{ii}} \right]^{-1} \sum_{i=1}^{N} \frac{X_i' X_i}{s_{ii}} \underline{b}_i$; $\underline{b}_i = (X_i' X_i)^{-1} X_i' \underset{\sim}{y}_i$; $s_{ii} = \frac{\underline{y}_i' M_i \underline{y}_i}{T - \Lambda}$ is an unbiased estimator of σ_{ii} and $M_i = I - X_i (X_i' X_i)^{-1} X_i'$. Under the null hypothesis, the asymptotic distribution of H_β is χ^2 with $\Lambda(N-1)$ d.f. as $T \to \infty$ and N is fixed.

3.3 Model and Test Procedure

For each group of countries, the tests will be conducted based on time series approach and panel approach. For each approach, there are two main tasks. The first task is to verify the relative version of PPP by testing the unit root of the real exchange rate. If the real exchange rate is found to be stationary, relative PPP is confirmed. The second task is to investigate the long run relationship of nominal exchange rate and price levels via the cointegration test. If the nominal exchange rate and price ratio are cointegrated, the weak version of PPP is confirmed. Subsequently, it is crucial to further investigate whether the cointegrating coefficients are 1s. If that is the case, not only weak PPP, but strong PPP is also asserted. Before the cointegration testing can be performed, however, it is required that the nominal exchange rate and price levels are integrated of the same order. Therefore, the preliminary analysis on the series of nominal exchange rate and price levels are conducted by testing for the unit root of all series in both level form and first difference form. Once the same order of cointegrated are confirmed, the cointegration test can be applied to test for PPP.

Table 3.1

Summary on Process of Testing for PPP for a Group of Countries

Time Series Approach	Panel Data Approach
1. Stationary tests of RER: *ADF, PP* Unit root → reject relative PPP Stationary → relative PPP holds	1. Stationary tests of RER: *LLC, IPS, MW (Fisher), Choi, Demeaned-IPS, CIPS** Unit root → reject relative PPP Stationary → relative PPP holds
2. Cointegration test of NER, PR and CPI • Pre-test for order of integrated, I(d), of NER and PR (or CPI) ↓ • Test for cointegration: *Engle-Granger, Johansen* No cointegration → reject PPP ↓ • (if cointegrated) Check symmetry and proportionality conditions Both satisfied → Strong PPP holds Otherwise → Weak PPP holds	2. Cointegration test of NER and PR • Pre-test for order of integrated, I(d), of NER and PR ↓ • Test for cointegration: *Pedroni, Kao, Fisher* No cointegration → reject PPP ↓ • (if cointegrated) Check for unity Coefficients Yes → Strong PPP holds No → Only weak PPP holds

Note: RER stands for real exchange rate.
NER stands for nominal exchange rate.
PR stands for price ratio.

3.3.1 Testing for Evidence Supporting PPP via Time Series Analysis

Stationary Test of the Real Exchange Rate

Firstly, traditional time series approach is used to determine the existence of PPP between Thailand and trade partners. To verify the existence of relative PPP, it is necessary that the real exchange rate displays reversion toward its own mean, whether it is unity or not. Thus, one popular testing procedure is to examine the behavior of the real exchange rate whether it is stationary or not. Long run PPP is said to hold if the sequence of real exchange rate is stationary or, equivalently, real exchange rate does not contain unit root. Hence, the unit root test is applied to real exchange rate in this manner.

By employing the well-known univariate unit root tests, such as ADF test or PP test, of the real exchange rate, the existence of relative PPP can be examined. For instance, the stationarity of real exchange rate can be examined by ADF test using this equation:

$$\Delta q_t = \alpha + \delta t + \rho q_{t-1} + \sum_{j=1}^{m} \beta_j \Delta q_{t-j} + u_t \qquad (3.39)$$

where optimal lag length, *m*, is chosen by Schwarz Information Criterion (SIC). The PP statistic is also applied to test for stationarity of real exchange rates.

Cointegraion Test of Nominal Exchange Rate and Price Ratio

Aside from verifying the real exchange rates, PPP can also be verified by testing for cointegration between nominal exchange rate and price ratios. The cointegration relationship can be specified as follows:

Restricted Model: $\ln E_t = \alpha + \beta \ln(P_t / P_t^*) + u_t \qquad (3.40)$

Unrestricted Model: $\ln E_t = \alpha + \beta_1 \ln(P_t) + \beta_2 \ln(P_t^*) + u_t \qquad (3.41)$

In this study the Engle-Granger two-step method and Johansen multivariate cointegration are applied to each individual series. The cointegration implies a long run relationship between economic variables. In other words, cointegration between exchange rate and price levels confirms the existence of weak PPP. However, to obtain strong PPP, more conditions are required. Firstly, the estimated coefficients of price ratio (β) and domestic price index (β_1) should be positive, while the estimated coefficient of foreign price index (β_2) should be negative to satisfy the condition of symmetry. Secondly, β should approach positive unity or β_1 and β_2 should have equal magnitude to satisfy the condition of proportionality.

3.3.2 Testing for Evidence Supporting PPP via Panel Analysis

Panel Unit Root of Real Exchange Rate

As suggested by numerous amont of literature that univariate unit root test has low power, the panel version of unit root test will also be investigated. Since an individual unit root test might not reject the null hypothesis of unit root in real exchange rate, it is expected that the panel of real exchange rate may somehow exhibit more evidence in favor of PPP if it is indeed valid.

According to Levin *et al.* (2002), a panel version of the ADF test for the real exchange rate of country i at time t, q_{it}, is defined by the following equation:

$$\Delta q_{it} = \alpha_i + \rho q_{i,t-1} + \sum_{j=1}^{m} \beta_{ij} \Delta q_{i,t-j} + u_{it} \tag{3.42}$$

where $i = 1,...,N$ indexes the countries, $t = 1,...,T$ the time periods and $j = 1,...,m$ the number of lags. This approach jointly tests if all series in the panel follow a unit root process under the null hypothesis.

$H_0: \quad \rho = 0$ (All real exchange rates contain unit roots.)

$H_1: \quad \rho < 0$ (All real exchange rates are stationary.)

Estimating equation (3.42), the null hypothesis of a unit root will be rejected in favor of the alternative of level stationarity if ρ is significantly less than zero

Similarly, the IPS test of the real exchange rate estimates the following equation:

$$\Delta q_{it} = \alpha_i + \rho_i q_{i,t-1} + u_{it}, \tag{3.43}$$

under these hypothesizes:

$H_0: \rho_i = 0 \ \ \forall\, i = 1,...,N$ (All real exchange rates have unit roots.)

$H_1: \rho_i < 0 \ \ \exists\, i, \ i = 1,...,N.$ (At least one real exchange rate are stationary.)

Apart from LLC test and IPS test, Fisher-ADF test, Fisher-PP test, Choi-ADF test, Choi-PP test, demeaned-IPS test and Pesaran's CIPS* test will also be conducted to confirm the results.

Panel Cointegration Test

Similar to the time series approach, the weak and strong version of PPP can be investigated through the cointgration between the nominal exchange rate and price ratio. When two series are non-stationary and integrated of the same order and the spurious regression of these series yields the stationary process, these series are cointegrated. Regarding this fact, in order to test the cointegration between nominal exchange rate and price ratio, both series, e_{it} and $p_{it} - p_{it}^{*}$, must follow the same order of integrated, I(1).

Pedroni (2001), as extended from his previous work, examined PPP by panel cointegration tests using the following specification:

$$e_{it} = \alpha_i + \gamma_i (p_{it} - p_{it}^{*}) + \varepsilon_{it} \qquad (3.44)$$

Equation (3.44) indicates the relationship between nominal exchange rate (e_{it}) and price ratio ($p_{it} - p_{it}^{*}$) or, equivalently, nominal exchange rate (e_{it}) and price ratio ($p_{it} - p_{it}^{*}$), in logarithm forms, are cointegrated with slope γ_i. The slope coefficients, γ_i, are allowed to vary by individual because factors leading to a nonunit value of cointegrating slope coefficient can be expected to have different magnitudes across countries.

Equation (3.44) is used to test for cointegration between nominal exchange rate and price ratio by applying Pedroni test, Kao test and Fisher test. If the cointegration is found, the validity of weak PPP is confirmed. For strong PPP to hold, however, it also requires that $\gamma_i = 1$ for all *i = 1,...,N*. Therefore, to ensure the existence of strong PPP, testing whether all γ_i are identical must be conducted. One of the possible approaches is to test the constancy of coefficients proposed by Swamy (1970). Rejection of parameter constancy implies the statement "Cointegrating coefficients possess value of 1, $\gamma_i = 1$, for all *i = 1,...,N*" is false. Subsequently, strong PPP is not valid.

3.4 Data Sources and Definition

There are various choices of a real exchange rate deflator. Chinn (2000) has discussed about choosing a deflator for calculating the real exchange rate. Since consumer bundles might be more similar across countries than producer or wholesale bundles, CPI may provide a more consistent measurement of price levels and thus of real exchange rates. The disadvantage of CPI and WPI, however, is that they include non-traded items. For this purpose, CPI or WPI which cover highly tradable goods are presented. In addition, if the concerned countries are fundamentally exporting to the third country markets, then choosing export price index may be more suitable. Nonetheless, export unit value indices are infamously subject to measurement error and the composition of the bundles of exports are more likely to vary broadly across countries than the corresponding WPI or CPI bundles. In this study, the real exchange rate is constructed from CPI due to the simplicity and availability of data.

The series of bilateral exchange rates (nominal exchange rate) in terms of country's currency per US dollar and consumer price indices are extracted from the International Financial Statistics (IFS) database of the International Money Fund to construct the real exchange rates and price ratios. All series are taken quarterly from 1987Q1 to 2006Q4. CPI series are seasonally adjusted using U.S. Census Bureau's X12 seasonal adjustment.

For the case of China, the CPI series is not reported in terms of CPI index, but in terms of percentage change relative to the same period of the previous years. Therefore, China's CPI can be calculated using this series with CPI index in 2003 from UNESCAP as a reference year. Though this is a possible choice, it might also lead to some distortions in data series.

Since this study analyzes the PPP based mainly on Thailand, Thai baht will be used as the numeraire currency, equivalently, using Thailand as a foreign country. Thus, the nominal exchange rate is defined as price of each domestic currency per baht. The exchange rates reported in IFS, however, are price of each currency per US dollar. Therefore, the nominal exchange rate of currency A per Thai baht will be calculated as follows:

$$\text{Nominal Exchange Rate} = \frac{\text{currency A / US dollar}}{\text{Baht / US dollar}}.$$