# CHAPTER 4

# METHODOLOGY

Time series econometric model is utilized to examine the relationship between the variables as suggested in the model. Prior to analyzing the long-run cointegrated time series relationships and short-run dynamics, time series variables are tested to discover the order of integration of each variable in the model.

## 4.1 Unit Root Test

In estimating time-series models, the time-series properties of the data will have an important influence on the specification of the econometric model and on the choice of estimator. A large number of macroeconomic time-series are trended and therefore in most cases are nonstationary. The problem with nonstationary data is that the conventional OLS regression method can easily yields incorrect conclusions. Usually, the regression has a very high $R^2$, t-statistics that make the estimates significant, but the results may have no economic meaning whatsoever. The reason is that the OLS estimates may be inconsistent, and thus statistical inferences are invalid. The results obtained from this kind of regression are said to be spurious and these regressions are named spurious regressions (Granger and Newbold, 1974). Testing for stationarity is, therefore, crucial for analyzing time series model.

By definition, it is required that a stationary process have a constant mean and variance over time. In addition, the covariance between any two time periods is constant and time-invariant. Formally, a stochastic process $x_t$ is said to be stationary if the following conditions are satisfied for all values of *t* and *t-s*:

$$E(x_t) = E(x_{t-s}) = \mu$$

$$E\left[(x_t - \mu)^2\right] = E\left[(x_{t-s} - \mu)^2\right] = \sigma_x^2 \quad : \quad \text{var}(x_t) = \text{var}(x_{t-s}) = \sigma_x^2$$

$$E\left[(x_t - \mu)(x_{t-s} - \mu)\right] = E\left[(x_{t-r} - \mu)(x_{t-r-s} - \mu)\right] = \gamma_s : \ \text{cov}(x_t, x_{t-s}) = \text{cov}(x_{t-r}, x_{t-r-s})$$

where $\mu$, $\sigma_x^2$ and all $\gamma_s$ are constants

Testing for nonstationarity is equivalent to testing for the existence of a unit root. Several statistical methods are constructed to test for unit roots. In this study, the familiar Augmented Dickey Fuller (ADF) method is applied.

4.1.1 Augmented Dickey-Fuller (ADF) Test

The Augmented Dickey-Fuller (ADF) test is an extension of the Dickey-Fuller (DF) test according to Dickey and Fuller (1979,1981). To eliminate the possible autocorrelation occurred in the original DF test which contains only one lag, extra lagged terms of the dependent variable are included in the model, resulting in the model to be applied in the ADF test. The ADF test can be written in the three following possible forms:

$$\Delta x_t = \gamma x_{t-1} + \sum_{i=1}^{p} \beta_i \Delta x_{t-i} + \varepsilon_t$$

$$\Delta x_t = \alpha + \gamma x_{t-1} + \sum_{i=1}^{p} \beta_i \Delta x_{t-i} + \varepsilon_t$$

$$\Delta x_t = \alpha + \gamma x_{t-1} + \lambda t + \sum_{i=1}^{p} \beta_i \Delta x_{t-i} + \varepsilon_t$$

where $\Delta$ is the difference operator, $\alpha$, $\gamma$, $\beta$ and $\lambda$ are coefficients to be estimated, $x_t$ is the variable whose unit roots are examined, $t$ is the time trend which is included to test for trend stationary of the variable and $\varepsilon$ is the white noise error term. The optimal lag length of the ADF regression is determined by the Schwartz Bayesian Criterion (SBC). The test involves testing the null hypothesis that $\gamma = 0$ (the series is nonstationary) against the alternative hypothesis that $\gamma < 0$ (the series is stationary). The critical values for each of the three above models are tabulated in MacKinnon (1996). If the null hypothesis can be rejected, the series $x_t$ is stationary at level or integrated of order zero, $x_t \sim \mathrm{I}(0)$. However, if the null hypothesis cannot be rejected, $x_t$ is nonstationary at level and is said to be integrated series. In an empirical model, the conventional OLS can be applied if all the variables are stationary. If not, further investigations are needed.

### 4.2 Cointegration Test

To examine the long-run equilibrium relationship between exchange rates and export prices, the Johansen and Juselius (1991, 1995) multivariate cointegration approach, which is based on estimating an error correction formulation of a VAR model is employed. Suppose that the variables in the pass-through equation are individually $I$ (1) and follow a vector autoregressive (VAR) of order p:

$$Y_t = \mu + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + ..... + \Phi_p Y_{t-p} + e_t \qquad (4.2.1)$$

where $Y_t$ is an $n \times 1$ vector of variables

$\mu$ is an $n \times 1$ constant vector

$\Phi_i$ is an $n \times n$ matrix of unknown parameters to be estimated, $i = 1, 2, ...., p$

$e_t$ is an $n \times 1$ independent and identically distributed vector of the error terms.

The above system can be reparameterized in the error correction format, that is

$$\Delta Y_t = \mu + \Gamma_1 \Delta Y_{t-1} + .... + \Gamma_{p-1} \Delta Y_{t-p+1} + \Pi Y_{t-p} + e_t \qquad (4.2.2)$$

where

$$\Gamma_i = (\Phi_1 + \Phi_2 .... + \Phi_i) - I_n, \text{ for } i = 1, 2, ..., p\text{-}1, \qquad (4.2.3)$$

$$\Pi = \Phi_1 + \Phi_2 + .... + \Phi_p - I_n \qquad (4.2.4)$$

where $\Pi$ contains the information on possible cointegrating relations between the $n$ elements of $Y_t$.

In equation (4.2), there are three possible cases. First, the matrix $\Pi$ can be the null matrix, which implies that the rank of $\Pi$ (r) equals 0. Then there is no long-run relationship among the variables and the system is reduced to a vector autoregressive (VAR) model in first differences. Second, the matrix $\Pi$ can have full rank *n*. In this case, the vector process $Y_t$ is stationary and the cointegrating relationship is undefined. Between these two extremes is the third case where matrix $\Pi$ has rank deficiency or $0 < rank\ (\Pi) = r < n$. This implies that there exists $r$ cointegrating vectors that make the linear combinations of $Y_t$ become stationary or cointegrated. $\Pi$ can be decomposed as $\Pi = \alpha\beta'$, where $\alpha$ and $\beta$ are $(n \times r)$

parameter matrices. The matrix $\beta$ contains the $r$ cointegrating vectors, while the matrix $\alpha$ represents the error-correction parameters.

The Johansen maximum likelihood cointegration testing method aims to test the rank of the matrix $\Pi$ in (4.2) using the reduced rank regression technique based on canonical correlations. There are two types of the tests for the rank of $\Pi$.

First is the Trace test. The null hypothesis that there are at most $r$ cointegrating vectors is tested against the alternative hypothesis that there exists $r$ or more cointegrating vectors. The Trace test (which is a likelihood ratio test) is defined as:

$$Trace\ (r) \ = \ -T \sum_{i=r+1}^{n} \log(1-\hat{\lambda}_i) \qquad (4.2.5)$$

where $T$ is the number of observations.

Another useful test is given by testing the significance of the estimated eigenvalues themselves, or

$$\lambda_{\max} = -T \log(1-\hat{\lambda}_{r+1}) \qquad (4.2.6)$$

which can be used to test the null hypothesis of $r$ against $r$+1 cointegrating vectors.

In this study, the Trace test is used to find the number of cointegrating vectors[1].

4.2.1 Lag length selection

As already mentioned, the cointegration test is based on a VAR model which involves a number of lagged variables in the system. Determining the lag length of VAR is crucial since the cointegration test is really sensitive to lag length. Too short lag length may not capture full dynamics of variables, while too long lag length reduces the degree of freedom. Generally, there is no consensus on what criterion should be placed on. In this study, the Likelihood Ratio test (LR test) and Akaike Information Criterion (AIC) is employed. Ideally, the selected lag length should also make the VECM have Gaussian error terms (i.e. standard normal error

[1] Monte Carlo results reported by Cheung and Lai (1993) indicate that the trace test is more robust than the $\lambda_{\max}$ test to possible non-normality of the residuals.

terms that do not suffer from non-normality, autocorrelation, heteroskedasticity, etc.) Therefore, diagnostic checks regarding these properties are conducted.

4.2.2 Diagnostic Test

In performing VECM residual diagnostic test, three possible properties are tested, i.e. autocorrelation, heteroskedasticity and non-normality. The methods used in diagnostic test are

*Lagrange-Multiplier (LM) test* is used to test for VECM residual autocorrelation.

*White Heteroskedasticity test (no cross terms)* is used to test for heteroskedasticity.

*Jarque-Bera residual normality test* is used to test for VECM residual normality.

4.2.3 Hypothesis Testing about the Cointegrating Vectors

One of the most interesting aspects of the Johansen procedure is that it allows for testing restricted forms of the cointegrating vectors. In terms of equation (3.2.11), i.e., $pxd_t = \delta + \gamma c_t - \beta_1 erd_t - \beta_2 erd_t^D + \lambda pc_t + u_t$, the restriction of interest is: $\beta_2 = 0$. Equivalently, this can be written as

Restricted model: $pxd_t = \delta + \gamma c_t - \beta_1 erd_t + \lambda pc_t + u_t$

Unrestricted model: $pxd_t = \delta + \gamma c_t - \beta_1 erd_t - \beta_2 erd_t^D + \lambda pc_t + u_t$

The objective is to test for the significant influence of the additional variable, $erd_t^D$ on the model so that we know whether to include this variable in the model. If we fail to reject the null hypothesis that $\beta_2 = 0$, then $\beta_2$ should be excluded from the model. Exchange rate pass-through is said to be symmetric. But if we can reject the null hypothesis that $\beta_2 = 0$, then $\beta_2$ should be included in the model. Exhange rate pass-through is said to be asymmetric.

The key insight to this hypothesis test is that *if there are r cointegrating vectors, only these r linear combinations of the variables are stationary.* Thus, the test statistics involve comparing the number of cointegrating vectors under the null and alternative hypotheses. Let $\hat{\lambda}_1, \hat{\lambda}_2, \ldots.., \hat{\lambda}_n$ and $\hat{\lambda}_{1^*}, \hat{\lambda}_{2^*}, \ldots.., \hat{\lambda}_{n^*}$ denote the ordered characteristic roots of the unrestricted and restricted models, respectively. To test restrictions, form the test statistic:

$$T\sum_{i=1}^{r}\left[\ln(1-\hat{\lambda}_{i^*})-\ln(1-\hat{\lambda}_i)\right] \qquad (4.2.7)$$

Asymptotically, this statistic has a $\chi^2$ distribution with degrees of freedom equal to the number of restrictions. Small values of $\hat{\lambda}_{i^*}$ relatively $\hat{\lambda}_i$ (for $i \le r$) imply a reduced number of cointegrating vectors. Hence, the restriction embedded in the null hypothesis is binding if the calculated value of the test statistic exceeds that in a $\chi^2$ table. If not, the null hypothesis is not binding. In the latter case, we cannot reject the null hypotheis that $\beta_2 = 0$. Thus, we can exclude $\beta_2$ from the cointegration space.

## 4.3 Vector Error Correction Model (VECM)

A vector error correction model (VECM) is a restricted VAR designed for use with nonstationary series that are known to be cointegrated. The VECM has cointegration relations built into the specification so that it restricts the long-run behavior of the endogenous variables to converge to their cointegrating relationships while allowing for short-run adjustment dynamics. The cointegration term is known as the *error correction term* since the deviation from long-run equilibrium is corrected gradually through a series of partial short-run adjustments. Consider the following equation.

$$y_t = \alpha_t + \beta x_t + z_t \qquad (4.3.1)$$
$$z_t = y_t - (\alpha_t + \beta x_t) \qquad (4.3.2)$$

“Granger Representation Theorem” (Engle and Granger, 1987) says that if $x_t$ and $y_t$ are cointegrated, we can find short-run adjustment dynamics pattern in the form of “Error-Correction Mechanisms” which can be written as follows.

$$\Delta x_t = \phi_1 z_{t-1} + \{\text{lagged } (\Delta x_t, \Delta y_t)\} + \varepsilon_{1t} \qquad (4.3.3)$$
$$\Delta y_t = \phi_1 z_{t-1} + \{\text{lagged } (\Delta x_t, \Delta y_t)\} + \varepsilon_{2t} \qquad (4.3.4)$$

where $z_{t-1}$ is the error-correction (EC) term, $\varepsilon_{1t}$ and $\varepsilon_{2t}$ are white noise and $\phi_1$, $\phi_2$ are non-zero coefficients.

From the Vector Error Correction Model (VECM), we can estimate both short-run and long-run export price adjustment coefficients and know the speed of adjustment of short-run disequilibrium toward long-run equilibrium.

The VECM in this study takes the following form:

$$\Delta pxd_t = \phi_0 + \phi_1 ECT_{t-1} + \sum_{i=1}^{k} \phi_2 \Delta pxd_{t-i} + \sum_{i=1}^{k} \phi_3 \Delta c_{t-i} + \sum_{i=1}^{k} \phi_4 \Delta erd_{t-i} + \sum_{i=1}^{k} \phi_5 \Delta pc_{t-i} + \varepsilon_t \qquad (4.3.5)$$

where the $ECT_{t-1}$ is the error-correction term derived from the long-run cointegrating relationship. The estimated coefficient of $ECT_{t-1}$, $\phi_1$, measures the speed of adjustment of short-run disequilibrium toward long-run equilibrium while $\phi_2$ ,…., $\phi_5$ measure the short-run relationship between the variables.

## 4.4 Definitions and Sources of Data

The data are monthly and all of them cover the seven year flexible exchange rate period from January 2000 to December 2006. They are defined as the following.

*Dollar export price (PXD)* is the price of Thailand's exports to the world market denominated in US dollar. The data is in terms of export price index in US dollar terms obtained from the Bureau of Trade and Economic Indices, Ministry of Commerce.[2] The export products included in each industry are categorized by export trade classification of the Ministry of Commerce. All data are seasonally adjusted.[3]

*Cost of production (C)* is specific to each export industry. Due to the scarcity of this kind of data, the widely used proxy, the producer price index (PPI) of the corresponding industry is utilized. PPI series are obtained from CEIC Database.

*Exchange rate (EXD)* is expressed as units of Thai Baht (THB) per unit of US dollar (USD). The exchange rate series is the official rate (monthly average) obtained from the International Monetary Fund, *International Financial Statistics (IFS)* CD-ROM.

*Competitor's price (PC)* is the price of competing products in the world market expressed in terms of US dollar. It is proxied by export share weighted average of the export price index of the corresponding products (exports of the same 4-digit and 6-digit harmonized code) exported from four major suppliers in the world market. All data are seasonally adjusted. Data is from author's calculation based on data taken from World Trade Atlas (WTA) and Global Trade Atlas (GTA).[4]

---

[2] The export price indexes are true price indexes, rather than the widely used unit value series which have limitations as price proxies, especially for manufactured goods (Lipsey et al., 1991). The MOC collects export prices from commercial invoices of surveyed firms and used them to construct the export price indexes. The unit value series, on the other hand, can be obtained by dividing export value by export quantity of each harmonized code and use them to construct the so called unit value index.

[3] Data are seasonally adjusted by X-12 method constructed by the US Bureau of Census.

[4] See Appendix A for the method of construction of this variable.

Because the export price index of motor cars and garments are not available. Therefore, this study uses the export price index of *vehicles and parts and accessories thereof* and *textiles* to be proxies for them. This is because motor cars and garments account for a large proportion in total exports of *vehicles and parts and accessories thereof* and *textile*s, respectively.

As in many previous studies, this study utilizes producer price index as a proxy for cost of production (*c*) of each manufactured export product. PPI is the price that the producers perceive from selling their products at the factory. Certainly, PPI does not exactly represent the true cost of production because it already incorporates the markup. However, it is the price that the producers obtain at the factory which excludes transport costs and value added taxes (VAT). Thus, so long as the true costs of production are unavailable, PPI is believed to represent the true cost at a certain level. The PPI series used as a proxy for cost of production are listed in table 4.1.

Table 4.1

Thailand’s PPI series used as proxy for cost of production (*c*)

| **Export Industry** | **Name of PPI series** |
|---|---|
| *Rubber products* | Rubber product |
| *Canned fish & seafood* | Processed Food: Fish & Aquatics Animals: Processed Fish |
| *Iron & steels* | Basic Metals: Iron, Steel & Ferro Alloys |
| *Furniture and parts* | Other Manufactured Good: Furnitures |
| *Motor cars, parts and accessories* | Motor Vehicles and Bodies (MV) |
| *Garments* | Textiles |
| *Plastic products* | Plastic Product |
| *Chemical products* | Chemicals, Chemical Product & Synthetic Fibres |