Chapter 3

Methodology

3.1 Population and Sample of the Study

The population of the study is the ten countries of the ASEAN group. The sample set of the study of the relationship between stock market development and economic growth includes the countries with stock exchanges.

The sample set of the study on the relationship between stock market development and income inequality includes only five countries—excluding Vietnam—since the data of the inequality index of Vietnam are not available.

3.2 Models

All models used in the study are panel models. An important benefit of using panel analysis is that it has the ability to investigate the dynamic changes from the short time-series. However, a limitation of using panel model is that it cannot include the policy dummy as an independent variable since the value of policy dummy cannot be easily derived from the data of all the countries.

There are two main groups of panel models used in the study. The first group is composed of those illustrating the relationships—both direct and indirect—between stock market development and economic growth. The second group shows the relationship between stock market development and income inequality. All models are formed linearly. Given the unavailability of data for some variables in certain countries during certain periods, all models are unbalanced panel models.

3.2.1 Models showing relationships between economic growth and stock market variables

The study aims to find out both the direct and indirect effects of the stock market on economic growth. The direct relationship between stock market variables—market capitalization ratio (MCR) and turnover ratio (TR)—and economic growth—growth of nominal GDP (GGDP, %)—is represented by *Model 1*. (Please refer to *3.3.2* for variable's descriptions.) *Model 2* illustrates the indirect impact of stock market variables on economic growth. Both models are based on models employed by Mohtadi and Agarwal (2004).

Model 1: Direct relationship

In this model, the dependent variable is GGDP. Independent variables include MCR and TR as well as other macroeconomic variables: growth of gross fixed capital formation (GK, %), growth of household final consumption expenditure (GC, %), ratio of foreign direct investment and GDP (FDI/GDP), growth of exports of goods and services (GEX, %), financing via international capital markets as % of GDP (INTERFIN, %), growth of value added in agricultural sector (GVAAG, %), growth of value added in manufacturing sector (GVAMANU, %) and growth of value added in service sector (GVASER, %).

$$GGDP_{it} = \alpha_1 + \alpha_2 MCR_{it} + \alpha_3 TR_{it} + \alpha_4 GK_{it} + \alpha_5 GC_{it} + \alpha_6 FDI/GDP_{it} + \alpha_7 GEX_{it} + \alpha_8 INTERFIN_{it} + \alpha_9 GVAAG_{it} + \alpha_{10} GVAMANU_{it} + \alpha_{11} GVASER_{it} + {}_{1\,it} \qquad (3.1)$$

where α_1 is an intercept term and $\square_{1\,it}$ is a residual term.

Model 2: Indirect relationship

In this model, stock market variables first affect some certain macroeconomic variables. Then these macroeconomic variables further affect economic growth. Therefore, there are two stages in *Model 2*. In the first stage GK, FDI/GDP and INTERFIN are each modeled with MCR and TR as their independent variables.

$$GK_{it} = \alpha_{12} + \alpha_{13} MCR_{it} + \alpha_{14} TR_{it} + {}_{2\,it} \qquad (3.2)$$

$$FDI/GDP_{it} = \alpha_{15} + +\alpha_{16} MCR_{it} + \alpha_{17} TR_{it} + {}_{3\,it} \qquad (3.3)$$

$$INTERFIN_{it} = \alpha_{18} + \alpha_{19} MCR_{it} + \alpha_{20} TR_{it} + {}_{4\,it} \qquad (3.4)$$

where α_{12}, α_{15} and α_{18} are intercept terms, and $\square_{2\,it}, \square_{3\,it}$ and $\square_{4\,it}$ are residual terms.

In the second stage, GGDP is modeled with fitted values of GK, FDI/GDP and INTERFIN—FGK, FFDI/GDP and FINTERFIN—as well as other independent variables included in *Model 1* except MCR and TR.

$$GGDP_{it} = \alpha_{21} + \alpha_{22}FGK_{it} + \alpha_{23}FFDI/GDP_{it} + \alpha_{24}FINTERFIN_{it} + \alpha_{25}GEX_{it} + \alpha_{26}GC_{it} + \alpha_{27}GVAAG_{it} + \alpha_{28}GVAMANU_{it} + \alpha_{29}GVASER_{it} + {}_{5\,it} \quad (3.5)$$

where α_{21} is intercept terms, and $\square_{5\,it}$ are residual terms.

3.2.2 Models showing relationship between income inequality and stock market variables

Model 3

On the side of income inequality, instead of using the well-known Gini index which suffers a few inconsistencies as the dependent variable, the less-known but superior estimated household inequality index (EHII) is chosen to represent the inequality. The inconsistencies of Gini index include low observation frequency, unexplained jumps in the series and the uses of mixed data types: gross versus net income data, household versus individual income data, and income versus expenditure data (Gimet and Lagoarde-Segot, 2010).

EHII was developed by Galbraith and Kum (2003) and updated by Daymon and Gimet (2009). It combines Gini index with a more precise Theil-index based measure of industrial sector's dispersion of pay from UTIP-UNIDO database.

$$GINIDS = \alpha + \beta T + \gamma X + \quad (3.6)$$

where $GINIDS$ is Denninger and Squire (1996) Gini coefficient, T is the Thiel-index based measure of industrial sector's wage inequality, and X is a vector of conditioning variables e.g. dummies reflecting data source and the ratio of manufacturing employment to total population.

$$EHII = \hat{\alpha} + \hat{\beta}T + \hat{\gamma}X \qquad (3.7)$$

where $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$ are deterministic terms from *Eq. 5*. Gimet and Lagoarde-Segot, 2010, who also employ EHII in their study, state that EHII is the most precise and extensive international income distribution.

In *Model 3*, EHII is the dependent variable with MCR, TR, growth of GDP per capita (GGDPCAP), GVAAG, GVAMANU and GVASERV as independent variables.

$$EHII_{it} = \alpha_{30} + \alpha_{31}MCR_{it} + \alpha_{32}TR_{it} + \alpha_{33}GGDPCAP_{it} + \alpha_{34}GVAAG_{it} + \alpha_{35}GVAMANU_{it} + \alpha_{36}GVASER_{it} + {}_{6\,it} \qquad (3.8)$$

where α_{15} is an intercept term and ${}_{6\,it}$ is a residual term.

3.3 Data

3.3.1 Data collection

All data are secondary data obtained from various sources. Macroeconomic variables and stock market variables are from World Development Indicators and the Global Development Finance (WDI_GDF) database. The income inequality index—EHII—is from University of Texas' Inequality Project.

3.3.2 Data description

Table 3.1 shows that descriptive statistics of the data used in the study. *Figure 3.1-Figure 3.14* show graphs of each variable during 1988-2009.

Table 3.1 Descriptive statistics of variables

Variables	Obs.	Mean	Std. Dev.	Min.	Max.
1.Growth of nominal gross domestic product (%)					
GGDP	83	4.88	4.13	-13.13	11.17
2. Growth of gross domestic product per capita, purchasing power parity (%)					
GGDPCAP	56	3.94	4.59	-14.32	10.63
3.Ratio of foreign direct investment (BoP) and gross domestic product					
FDI/GDP	83	0.02	0.03	-0.04	0.11
4.Growth of trade as % of GDP (%)					
GTRADE	83	1.55	12.00	-34.56	71.78
5.Growth of exports of goods and services (BoP) (%)					
GEX	83	10.04	11.92	-18.97	32.31
6.Financing via international capital markets as % of gross domestic product (%)					
INTERFIN	83	3.50	2.62	0.09	9.28
7.Growth of gross fixed capital formation (%)					
GK	83	8.57	18.09	-60.26	37.78
8.Growth of household final consumption expenditure (%)					
GC	83	5.18	5.01	-12.09	17.48
9.Growth of nominal value added in agricultural sector (%)					
GVAAG	83	7.27	15.60	-50.29	59.09
10.Growth of nominal value added in manufacturing sector (%)					
GVAMANU	83	9.53	14.19	-58.73	52.55
11.Growth of nominal value added in service sector (%)					
GVASER	83	9.18	13.62	-59.00	48.03
12.Market capitalization ratio (Market capitalization over GDP)					
MCR	83	0.69	0.66	0.01	3.29
13.Turnover ratio (Value of shares traded over market capitalization)					
TR	83	0.49	0.32	0.08	1.64
14.Estimated household inequality index					
EHII	56	41.89	4.88	34.26	53.93

Source: from computation.

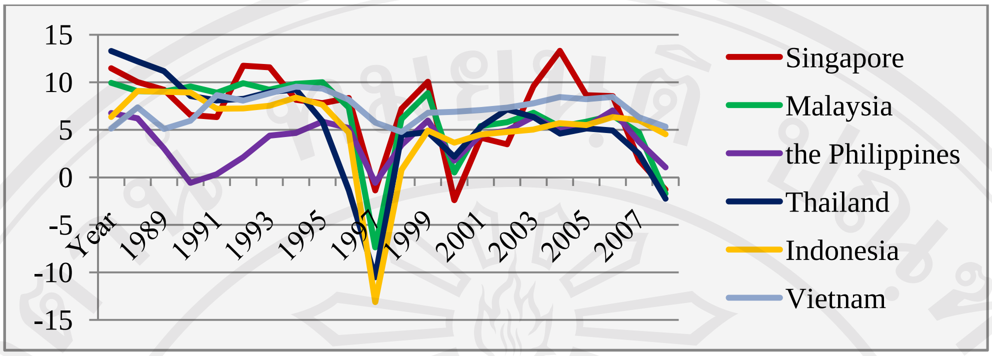

Figure 3.1 GGDP (%) of Singapore, Malaysia, the Philippines, Thailand, Indonesia and Vietnam

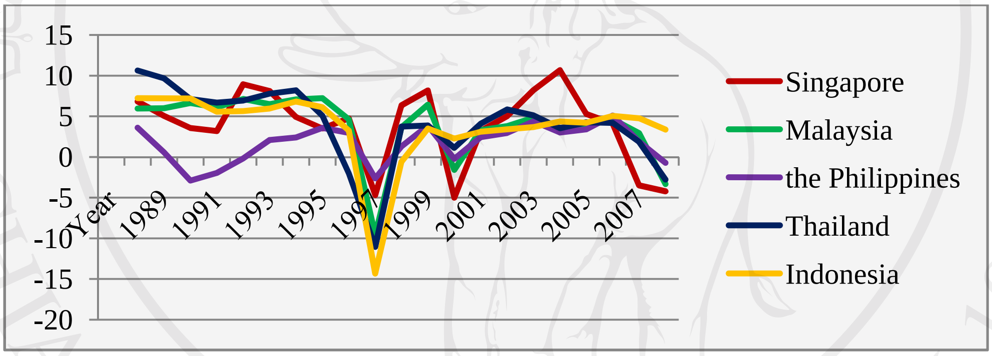

Figure 3.2 GGDPCAP (%) of Singapore, Malaysia, the Philippines, Thailand and Indonesia

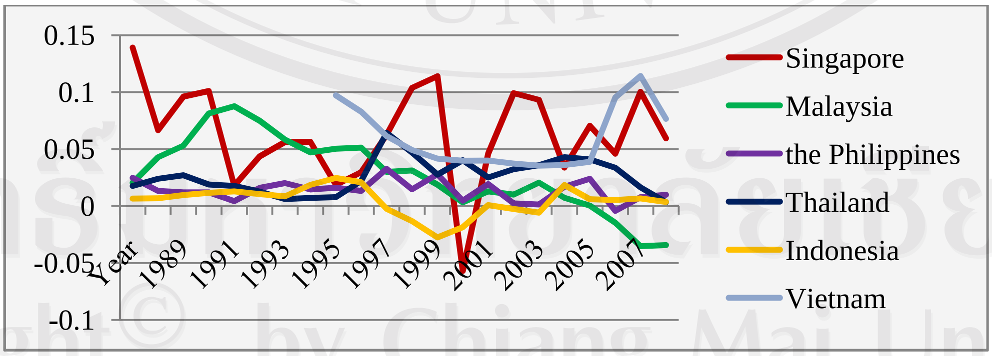

Figure 3.3 FDI/GDP of Singapore, Malaysia, the Philippines, Thailand, Indonesia and Vietnam

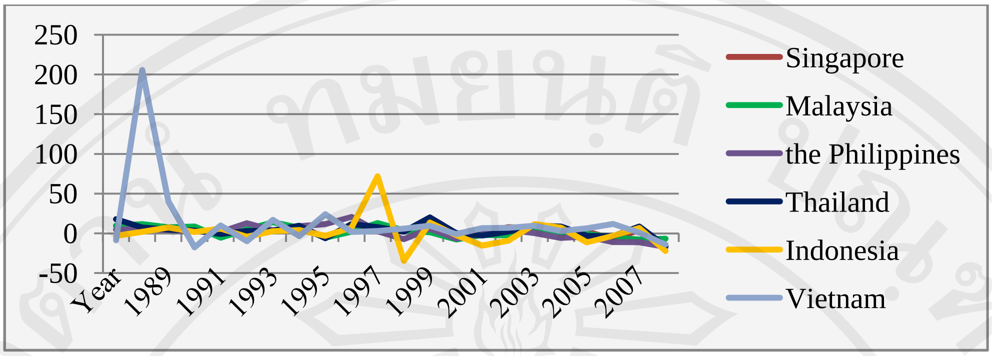

Figure 3.4 GTRADE (%) of Singapore, Malaysia, the Philippines, Thailand, Indonesia and Vietnam

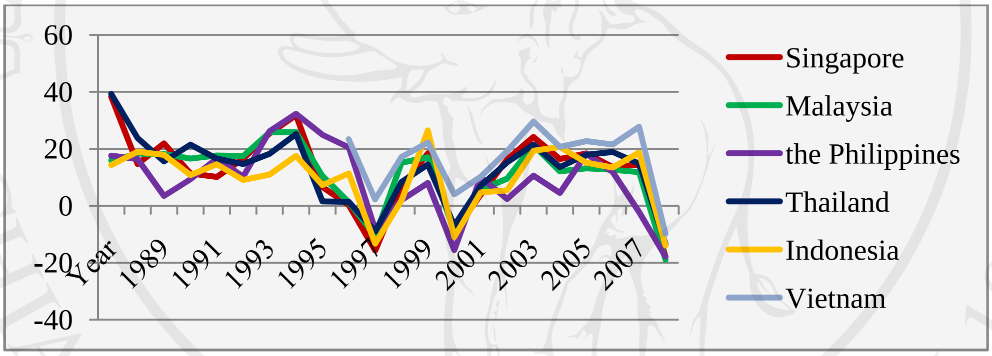

Figure 3.5 GEX (%) of Singapore, Malaysia, the Philippines, Thailand, Indonesia and Vietnam

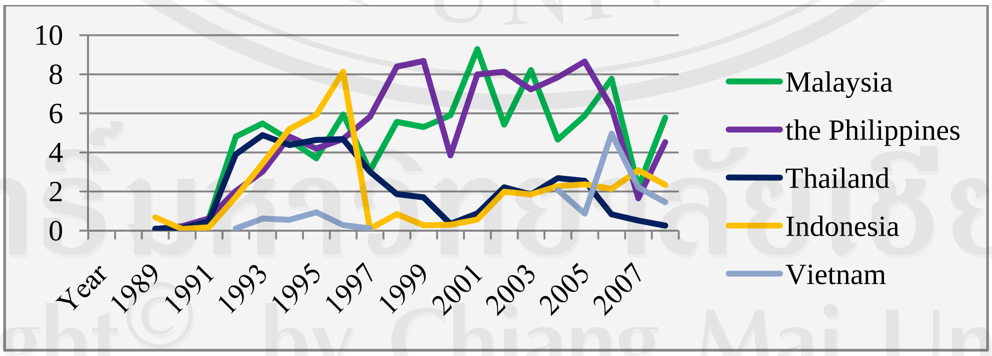

Figure 3.6 INTERFIN (%) of Malaysia, the Philippines, Thailand, Indonesia and Vietnam

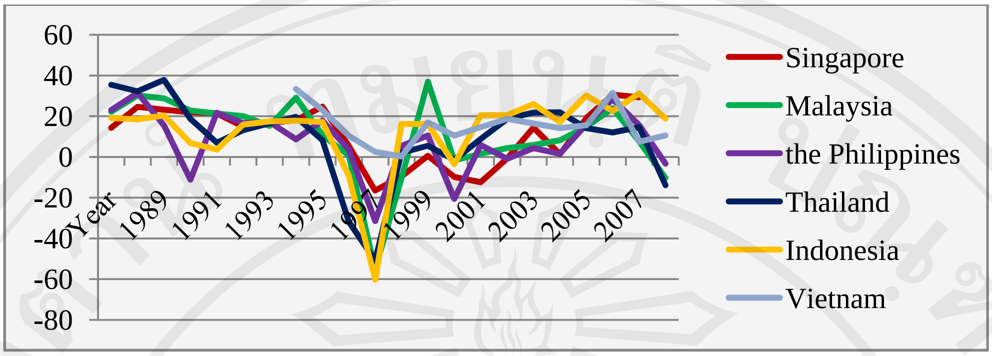

Figure 3.7 GK (%) of Singapore, Malaysia, the Philippines, Thailand, Indonesia and Vietnam

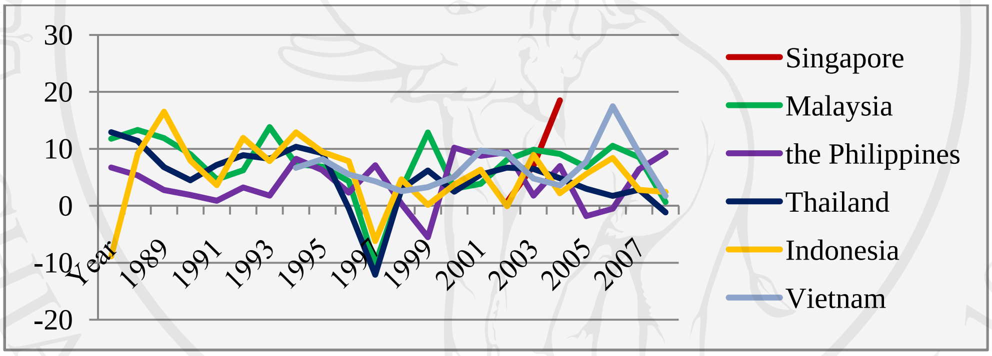

Figure 3.8 GC (%) of Singapore, Malaysia, the Philippines, Thailand, Indonesia and Vietnam

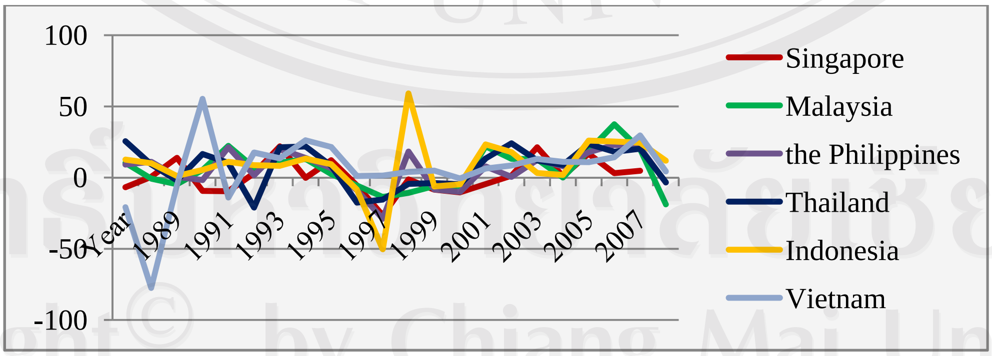

Figure 3.9 GVAAG (%) of Singapore, Malaysia, the Philippines, Thailand, Indonesia and Vietnam

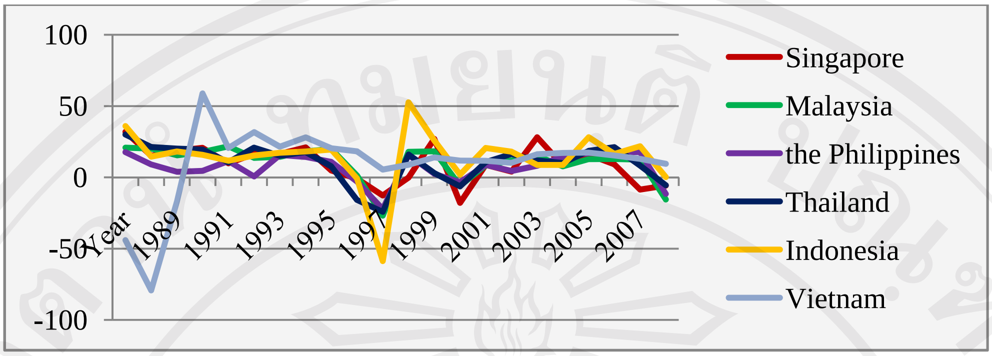

Figure 3.10 GVAMANU (%) of Singapore, Malaysia, the Philippines, Thailand, Indonesia and Vietnam

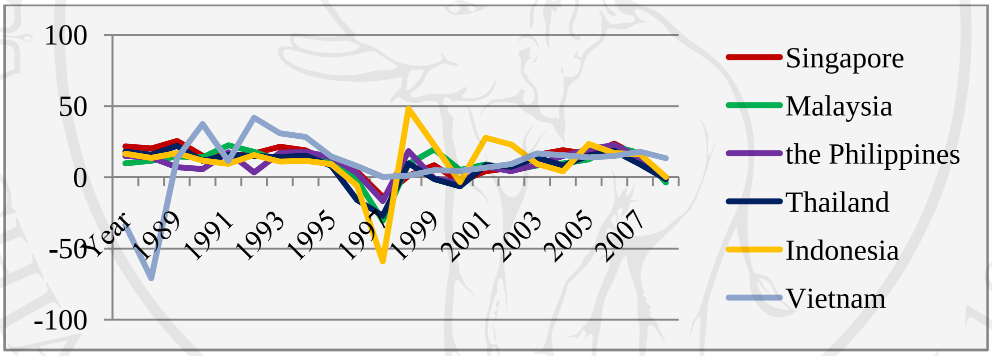

Figure 3.11 GVASER (%) of Singapore, Malaysia, the Philippines, Thailand, Indonesia and Vietnam

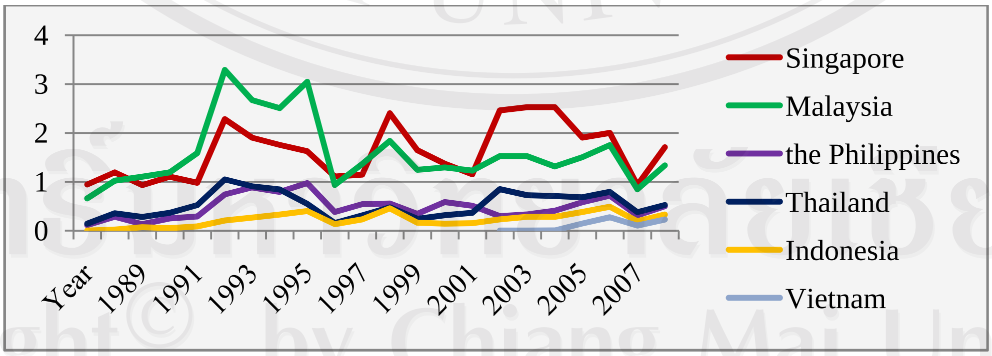

Figure 3.12 MCR of Singapore, Malaysia, the Philippines, Thailand, Indonesia and Vietnam

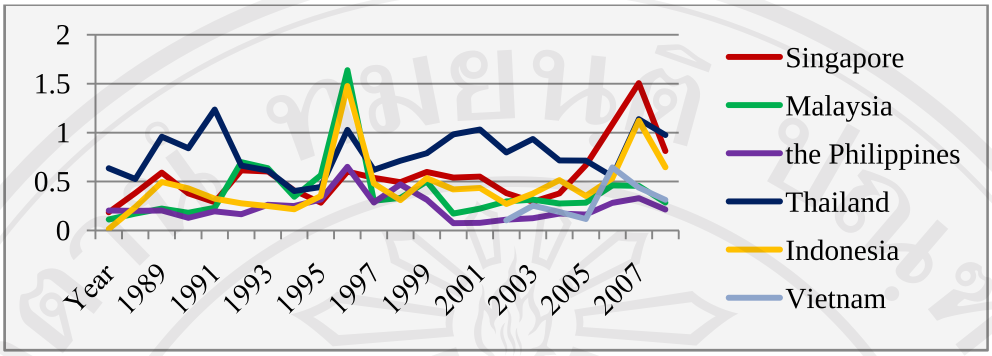

Figure 3.13 TR of Singapore, Malaysia, the Philippines, Thailand, Indonesia and Vietnam

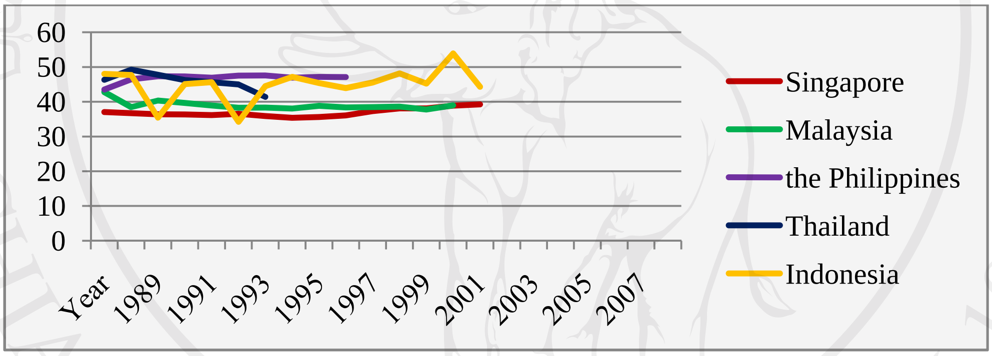

Figure 3.14 EHII of Singapore, Malaysia, the Philippines, Thailand and Indonesia

3.4 Analysis Methods

Employing data of the ASEAN economies, panel analysis is employed on all models (see *3.3.1*). Prior to all analysis, unit root tests are employed to test stationarity of all variables used in the study (see *3.3.2*). After the stationarities in all variables are confirmed, in line with Mohtadi and Agarwal (2004) whose models are the basis of the first two models, least-square regression and two-stage least-square regression (see *3.3.3*) are employed on *Model 1* and *Model 2* respectively. Then *Model 1* is analyzed with Granger causality test (see *3.3.4*) to test for causality between stock market variables and economic growth. For *Model 3*, since the value of the dependent variable—EHII—lies in the limited range within 0 and 100, Tobit (censored regression) analysis is employed (see *3.3.5*).

3.4.1 Panel analysis (summarized from Yaffee, 2003**)**

The method of panel data analysis is of studying a particular subject within multiple sites—cross sections—which are periodically observed over a defined time span—time series. Panel analysis, with enough cross sections, allows the dynamics of change to be observed from even short time series. Employing panel analysis can enhance the quality and quantity of data that would not be possible under either time series or cross sections alone (Gujarati, 2003).

Generally, panel data sets include sequential blocks of cross-sectional data within each of which resides a time series.

There are two dimensions of the variables. First is a cross-sectional unit of observation—in this case is country *i*. Second is temporal reference or year *t*. The error term also has two dimensions: country and year. The data set is called a balanced panel if there are no missing values, and an unbalanced panel otherwise. A general linear panel model can be written as followed:

$$y_{it} = X'_{it}\beta_{it} + \varepsilon_{it} \qquad (3.9)$$

or

$$y_{it} = \alpha_i + X'_{it}\beta_{it} + \varepsilon_{it} \qquad (3.10)$$

where y_{it} is vector 1×1 of dependent variable, α_i is intercept term, X_{it} is vector $k \times 1$of independent variables, β_{it} is vector $k \times 1$of coefficient terms, and ε_{it} is residual term.

3.4.2 Panel unit root tests

Among various methods of panel unit root tests, Im, Pesaran and Shin (IPS) tests (Im, Pesaran and Shin, 2003) and Fisher type tests (Maddala and Wu, 1999, and Choi, 2001) are the two tests available for unbalanced panel data. Nevertheless, IPS tests cannot deal with variables that has less than 10-period values per cross section (some countries in the study has less than 10-period EHII). Therefore, Fisher-Augmented Dickey-Fuller and Fisher-Philips-Perron are employed due to their abilities to deal with all data presented in the study.

Fisher type test: this type of test has been proposed by Maddala and Wu (1999), and by Choi (2001). Based on Fisher's (1932) results, it is the tests that combine the p-values from individual unit root tests.

Define π_i as the p-value from individual unit root test for cross-section i. The null hypothesis is that all N cross-sections have unit root. From this hypothesis, the asymptotic result is:

$$-2\sum_{i=1}^{N} \log(\pi_i) \to x_{2N}^2 \qquad (3.11)$$

Choi further demonstrates that:

$$Z = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \Phi^{-1}(\pi_i) \rightarrow N(0,1) \qquad (3.12)$$

where Φ^{-1} is the inverse of the standard normal cumulative distribution function (Eviews 7 Help Topic).

3.4.3 Panel linear least-squared regression (LLSR)

This type of regression is the most widely used modeling method due to its effectiveness and completeness. It has the ability to adapt to broad types of situations and to perform effectively even with a small data set.

The data that can be used in the LLSR needs to have function in which each independent variable is multiplied by an unknown parameter. Moreover, there has to be one unknown parameter with no corresponding independent variable in the function. This is either intercept or error term of the function. Finally, independent variables—each multiplied by a parameter—and the isolated parameter are summed to produce the final function value. With these properties, the model becomes a linear-in-parameter model.

LLSR estimates the unknown parameters by minimizing sum of squared deviations between the data and the model. In the minimization process, the overdetermined system of equations is formed by the data to a system of *P* equations in *P* unknown. Then, estimated parameters are obtained by solving this new system (NIST/SEMATECH, 2011).

3.4.4 Panel Granger causality test

Consider the following heterogeneous autoregressive model:

$$y_{it} = \alpha_i + \sum_{k=1}^{K} \gamma_i^{(k)} y_{i\,t-k} + \sum_{k=1}^{K} \beta_i^{(k)} x_{i\,t-k} + \square_{\boldsymbol{it}} \tag{3.13}$$

$$\beta_i = (\beta_i^{(1)}, \dots, \beta_i^{(K)})' \tag{3.14}$$

where $\alpha_i$is the fixed individual effects, the autoregressive parameters $\gamma_i^{(k)}$ and the coefficients $\beta_i^{(k)}$ differ across countries. This model is a fixed coefficients model with fixed individual effects. For each country, $\square_{it}$ are $i.i.d\ (0, \sigma_{\square,i}^2)$ and are distributed independently across cross sections. The null hypothesis is:

$$H_0: \beta_i = 0 \;\; \forall i = 1, \dots, N \tag{3.15}$$

The null hypothesis is therefore there is a causality relationship from x to y for at least one country (Hurlin and Venet, 2008).

3.4.5 Panel Tobit model

Tobit model or censored regression model can be described as:

$$y_{it}^* = x_{it}\beta + y_{it-1}^*\lambda + \square_{it} \tag{3.16}$$

$$y_{it} = max\{y_{it}^*, 0\} \tag{3.17}$$

$$\square_{it} = d_i + u_{it}, \quad t = 1, \dots, T \quad i = 1, \dots, N \tag{3.18}$$

where d_i is This model is characterized by lagged latent dependent variables, u_{it} is an idiosyncratic error varying across time and individuals (Chang, 2002).