CHAPTER 3

Methodology

The study initially aims to investigate the factors matter Cambodia's bilateral trade with its major trading partners in ASEAN by the gravity model. Furthermore, it aims to predict the Cambodia's trade potential and performance with its major trading partners in ASEAN. To analyze these two main objectives, this chapter will explain the methodology of this study into 5 main sections. The first section describes the nature of variables and data sources of each variable (section 3.1). The second part will define the specification gravity model which is adopted from the traditional gravity model to use in this study (section 4.2). Section 4.3 will explain expected result of each independent variables base on the theoretical foundation and empirical studies. Section 4.4 will discuss about the econometrical application of the gravity model estimation. This section demonstrates the two estimation techniques such as Pooled OLS and Random Effects Model. Additional, it also expresses the Breusch-Pagan LM test to decide whether Pooled OLS or Random Effects Model is preferred. Especially, there are some other diagnostic tests to verify whether no significant problems occur in the model. The last section is well defined the forecasting of potential trade and performance (section 4.5).

3.1 Description of Variables and Data source

The factors influencing bilateral trade between Cambodia and its trading partners can be analyzed by the gravity framework. This research engages with a panel dataset of 6 cross-sections of Cambodia's country partners in ASEAN over the entire period of 21 year, from 1995 to 2015. The variables using in this study is described below:

Total bilateral trade (TT_{ijt}): the dependent variable which is sum of yearly bilateral exports and imports between Cambodia and each of her trading partners. Data on bilateral exports and imports were obtained from the CIEC Data Manager in faculty of

Economics, Chiang Mai University.

Per capita GDP ($PCGDP_{it}$ and $PCGDP_{jt}$): is an explanatory variable which presents for the income level and purchasing ability of Cambodia and trading partners. The yearly data on per capita GDP of each country in this study was obtained from the CIEC Data Manager in faculty of Economics, Chiang Mai University.

Bilateral Distance ($Dist_{ij}$): The empirical studies naturally use distance between countries' capital city. Hence, the bilateral distance or geographical distance between Cambodia and its trading partners is measured in kilometers between Phnom Penh capital city (Cambodia) to the capital city of each country partner. The data was obtained from Centre d'Etudes Prospectives et d'Informations Internationales (CEPII). The bilateral trade between Cambodia and its major trading in ASEAN is shown in table 3.1.

Table 3.1 The bilateral distance between Cambodia and it trading partners

Countries	Capital city	Bilateral distance (Kilometres)
Indonesia	Jakarta	1982.272
Malaysia	Kuala Lumpur	1005.151
Singapore	Singapore	1152.419
Thailand	Bangkok	535.9692
Vietnam	Hanoi	1054.802
Philippines	Manila	1771.795

Source: Centre d'Etudes Prospectives et d'Informations Internationales (CEPII)

Real Bilateral Exchange Rate ($RBER_{ijt}$): The real bilateral exchange rate is real exchange rate between the Cambodia riel and the currency of the three bordering countries. RBER was calculated as the nominal bilateral exchange rate between the riel and each trading partner's currency (E_{ijt}) multiplied by the ratio of foreign price index (CPI_j) to riel's price index (CPI_{it}). The formula of real exchange rate can be written:

$$\mathrm{RBER}_{ijt} = \frac{\mathrm{E}_{ijt}\mathrm{CPI}_{jt}}{\mathrm{CPI}_{it}} \qquad (3)$$

Temporarily, the availability of data nominal exchange rate is in national currency of

each country against US dollar. Hence, it is obligatory to calculate the nominal bilateral exchange rate between the riel and the national currency of partners by the triangular arbitrage technique. The yearly data of E_{ijt}, CPI_j, and CPI_{it} are compiled from CEIC Data Manager at Chiang Mai University.

Three dummy variables are included in the model:

1. ***Common border-sharing*** is a dummy variable takes value 1 when Cambodia and trading partners have common border-sharing, 0 otherwise.

2. ***AFTA*** is a dummy variable for number of years since Cambodia and neighbor countries become a member of ASEAN Free Trade Area, defined as being equal to 1 for 2000-2015 and 0 otherwise;

3. ***crisis*** is dummy variable for a number of years of the Asian crisis and global financial and economic crisis, defined as being equal to 1 for 1997, 1998, 2008-2010 and 0 otherwise. (Hor & Thaiprasert, 2015).

Table 3.2 Variables summary (1990-2015)

Variables	**Data sources**
Total Bilateral Trade (TT_{ijt})	Sum of bilateral exports and imports) CEIC Data Manager at Chiang Mai University
Per capita GDP ($PCGDP_{it}$ and $PCGDP_{jt}$)	Asian Development Bank (ADB), Key Indicators for Asia and the Pacific 2015.
Distance ($Dist_{ij}$)	Centre d'Etudes Prospectives et d'Informations Internationales (CEPII)
Real Bilateral Exchange Rate ($RBER_{ijt}$)	Calculating by using this formula: $\mathrm{RBER}_{ijt} = \frac{E_{ijt}CPI_{jt}}{CPI_{it}}$
Common border-sharing ($Border_{ij}$)	Takes value 1 when Cambodia and trading partners have common border-sharing, 0 otherwise
AFTA	Take value 1 for 2000-2014 and 0 otherwise
crisis	A dummy variable for a number of years of Asian crisis and global financial and economic crisis, defined as being equal to 1 for 1997, 1998, 2008-2010 and 0 otherwise.

3.2 Gravity Model Specification

Specification of the Gravity Model in the trade situation of the Kingdom of Cambodia with its trading partners in ASEAN, the basic bilateral gravity model Tibergen (1962) who expressed gravity model of trade which was adopted from the law of gravity of physician Isaac Newton. GDP and distances are introduced as the two powerful determinants effect bilateral trade between two countries. The bilateral gravity equation is presented as below:

$$T_{ij} = \alpha \frac{GDP_i \cdot GDP_j}{D_{ij}^2} \quad (4)$$

where: T_{ij} is the total bilateral trade flow from original country i to country j

GDP_i and GDP_j are the economic size (gross domestic product) of two countries (i and j)

D_{ij} is the distance between two countries (i and j) and α is a constant term

An alternative basic gravity trade equation is expressed in the log-linear form by many empirical studies. The equation 5 is a linear equation in logarithmic form of gravity model (Greene, 2013):

$$\ln T_{ij} = \beta_0 + \beta_1 \ln GDP_{it} + \beta_2 \ln GDP_{jt} + \beta_3 \ln D_{ij} + \varepsilon_{ijt} \quad (5)$$

The augmented gravity model using in this paper contain necessarily variables such as per capita GDP of Cambodia and its major partners in ASEAN ($PCGDP_{ij}$), real bilateral exchange rate between Cambodia and its major partners in ASEAN ($RBER_{ij}$), geographical distance between Cambodia and its major partners ($Dist_{ij}$), common border ($border_{ij}$), ASEAN free trade agreement (AFTA), Asian financial crisis and global financial crisis (crisis). The gravity model of Cambodia's bilateral trade can be mathematically written as following:

$$\ln(TT_{ijt}) = \beta_0 + \beta_1 \ln(PCGDP_{it}) + \beta_2 \ln(PCGDP_{jt}) + \beta_3 \ln(RBER_{ijt}) + \beta_4 \ln(Dist_{ij}) + \beta_5 Border_{ij} + \beta_6(AFTA) + \beta_7(crisis) + u_{ijt} \quad (6)$$

Where: $ln(TT_{ijt})$: percentage change of total bilateral trade between Cambodia and its partners j at time t

$ln(PCGDP_{it})$ and $ln(PCGDP_{jt})$ are the percentage change of per capital GDP of Cambodia and trading partner j at time t

$ln(RBER_{ijt})$ is the percentage change of real bilateral exchange rate between Cambodia and trading partner j at time t

$ln(Dist_{ij})$ is the percentage change of bilateral distance between Cambodia and trading partner j

$Border_{ij}$ is the dummy variable, take value 1 when both Cambodia and trading partner j share common border, 0 otherwise.

$AFTA$ is the dummy variable, take value 1 when both Cambodia and trading partner j are belong to AFTA, 0 otherwise

Crisis is a dummy take value 1 in 1997-1998 and 2008-2009, and 0 otherwise

3.3 The expected result or Hypothesis

The per capita GDP indicate the purchasing power and income level. In this study, it works as a proxy of economic size or production capacity. The increase income level indicates the increase in production capacity which tends to increase the imports and exports. Hence, the increasing of per capita GDP of Cambodia and its trading partners will increase the bilateral trade of Cambodia with her trading partners. The coefficient of per capita GDP is expected to be positive. ***Hypothesis 1:*** $\boldsymbol{\beta_1 and\ \beta_2 > 0}$

The real bilateral exchange rate indicate the relative price of foreign goods in terms of domestic goods (BONUEDI, 2013). The international competitiveness of producing goods is measured by real bilateral exchange rate. Once the real bilateral exchange rate is appreciated, it means that Cambodian Riel in term of partners' currency is more expensive. The domestic goods seem to be more expensive as well. Hence, it will increase the imports, while the exports will get hurt. On the other hand, if the real

bilateral exchange rate is depreciated, it means Cambodian Riel in term of partners' currency is getting cheaper. It leads the domestic good is cheaper too. Thus, it will increase the exports rather than imports. In short, the real bilateral exchange rate coefficient can be positive and negative in the study. ***Hypothesis 2:*** $\boldsymbol{\beta_3 > 0}$ ***or*** $\boldsymbol{\beta_3 < 0}$

The bilateral distance or geographical distance between Cambodia and its trading partners is a proxy serves for the trade cost. Hence the bilateral distance seems to have negative relationship with bilateral trade. The coefficient of bilateral distance is negative. ***Hypothesis 3:*** $\boldsymbol{\beta_4 > 0}$

Two neighbor countries have a common border-sharing seems to trade more with other. Hence, Cambodia's bilateral trade increases if the Cambodia and trading partners have common border-sharing. ***Hypothesis 4:*** $\boldsymbol{\beta_5 > 0}$

Cambodia became a member of ASEAN Free Trade Area in 1992. The purpose is to remove the trade barriers between Cambodia and ASEAN countries by eliminating tariff to be 0 tariff. Cambodia expected Common Effective Preferential Tariff (CEPT) Scheme in 2010. Even though, AFTA is expected to be a tool to promote trade in ASEAN region by removing tariff, but the empirical analysis of many previous studies present various results. Some of them illustrates the trade recreation by AFTA. On the other hand, AFTA will reduce the exports in the regional trade which is called diverted trade (Heng, 2014). Hence, AFTA is expected to boost Cambodia bilateral trade. ***Hypothesis 5:*** $\boldsymbol{\beta_6 > 0}$

The Asian Financial Crisis occurred in Thailand in 1997 and spread to other countries such as Malaysia, Philippines, Indonesia, and South Korea cited in Heng (2014). On the other hand, the Global Financial Crisis during 2008-2009 caused China's, Japan's, and United States' income dropped, and influenced on Southeast Asia's trade. In conclusion, Cambodia's bilateral trade will be negatively impacted by crisis. ***Hypothesis 6:*** $\boldsymbol{\beta_7 < 0}$

Table 3.3 The expected sign of all independent variables

Parameters	Variables	sign
B_1	$PCGDP_c$	(+)
B_2	$PCGDP_j$	(+)
B_3	$RBER_{cj}$	(-) or (+)
B_4	$Dist_{cj}$	(-)
B_5	$Border_{ij}$	(+)
B_6	AFTA	(+)
B_7	crisis	(-)

3.4 Econometric application with Gravity Model

According to above literature review, the recent empirical study used cross-section data to estimate the gravity models. The panel data is considered on the pooling time-series of cross-sections which present the a panel of countries followed over time (Baltagi, 2008). Moreover, there are several advantages using panel data introduced by Baltagi (2008). Firstly, panel data sets give the ability to control for individual heterogeneity efficiently. Additionally, it provides a better recognition and estimation effects that are simply not noticeable in cross-sections or time-series data. Panel data also gives more informative data, more reliable estimations, and more knowledgeable behavioral models of testing with less limited assumptions. Repeated cross-sections can show how the proportions change overtime which are better fitted to study the dynamics of change. Conversely, there are limitations of using panel data. Firstly, the poolability assumption is needed, although there are formal tests to estimate its validity. Secondly, the complicated analysis is caused by the potential cross-sectional dependence. The balance panels are required in some tests and methods; and another challenge is cross country data consistency (Baltagi, 2008; Hsiao, 2003).

The panel data framework will be preceded by using 2 different estimation techniques such as Pooled OLS and Random Effects Model. These two models are explained as follow:

3.4.1 Pooled Ordinary Least Squares (OLS)

The pooled OLS estimator is the simplest estimator which ignores the panel structure of the data. The pooled OLS model can be written as following:

$$Y_{it} = X_{it}\beta + \alpha + u_{it} \quad (7)$$

where: Y_{it}: the observation on dependent variable for cross- country i in time t

X_{it}: (1×K) vector of explanatory variables observed for country i in time t

β: (K ×1) vector of parameters

u_{it} is an error or disturbance term specific to country i in period t

The intercept (α) and all the coefficients (β) are not changeable according to the pooled OLS approach. Additionally, α and β are identical for all individuals across time, and the error term is independent and identically distributed $u_{it} \sim iid(0, \sigma_u^2)$ for all *i* and *t*. It means that the errors are homogenous across individuals and times. According to Gujarati (2004) the highly restricted assumptions is caused the pooled regression ignores the individuality of each country and distorts the true picture of the relationship between the dependent variables (Y) and explanatory variables (X) across the countries.

3.4.2 Random Effects Model

Random effects model is worked contrarily to fixed effects model, it assumes that the intercept and a random component ε_{it} can capture the individual effects. The random component is independently and identically distributed over individuals, so it is not associated with the regressors and part of the error term. The form of the REM can be expressed as follow:

$$Y_{it} = X_{it}\beta + \mu + \alpha_i + \varepsilon_{it}, \qquad \varepsilon_{it} \sim iid(0, \sigma_u^2); \qquad \alpha_i \sim iid(0, \sigma_\alpha^2) \quad (8)$$

where $\alpha_i + \varepsilon_{it}$ is an error term containing an individual specific component α_i, and a time series and cross section error component ε_{it}. The REM is estimated by using the General Least Square estimator which is a weighted average of between and within estimator. The advantage of random effects model is to allow time-invariant variables to be included among the regressors.

Finally, the Breusch-Pagan Lagrange Multiplier (LM) test is used to decide whether Random Effects Model or Pooled OLS Model is preferred. The null hypothesis in the Breusch-Pagan LM test is that variances across entities is zero, while the alternative hypothesis is variances across entities is different from zero. The command of Breusch-Pagan LM test for Random Effects Model in Stata 13 is xttest0.

3.4.3 Other Diagnostic Tests

After the Breusch-Pagan LM test, Random Effect Model is preferred. Hence, we do other tests to relax the model above. Firstly, the multicollinearity test is observed by using correlation matrix approach and Variance Inflation Factor (VIF). It aims to test for correlation between variables. Second test is Modified Wald Test for GroupWise Heteroscedasticity to diagnose the heteroscedasticity between groups. Thirdly, the serial correlation tests apply to the panel data with long time series over than 20 years. The serial correlation problem is causes the standard errors of the coefficients to be smaller than the actual value. Lastly, the cross-sectional dependence can lead to bias in tests results; hence the Pesaran CD's and Friedman's tests are used to detect cross-section dependence.

If the models above present the significant problems such as Heteroscedasticity, Multicollinearity, autocorrelation, or cross-sectional dependence, it will show the weakness of model because those problems will lead to bias results. In order to eliminate those spurious problems, Torres-Reyna (2007) recommended to use the Feasible Generalized Least Square (FGLS).

3.5 Predicting the Cambodia's potential trade and performance

The bilateral trade potential has been predicted by some authors using the out-of-sample approach and the in-sample approach. In this study, the out-of-sample approach is employed to forecast the potential and performance of Cambodia's bilateral trade with its major trading partners. The estimation coefficient βs from the regression model above will be used to predict the potential trade with the mathematical form as below:

$$\ln(TT_{ijt}) = \beta_0 + \beta_1 \ln(PCGDP_{it}) + \beta_2 \ln(PCGDP_{jt}) + \beta_3 \ln(RBER_{ijt}) + \beta_4 \ln(Dist_{ij}) + \beta_5 Border_{ij} + \beta_6(AFTA) + \beta_7(crisis) + u_{ijt} \quad (9)$$

The obtained parameters (β) estimates from the gravity model under the panel data framework were used to predict the trade potential between the Cambodia and trading partners (Brülhart & Kelly, 1999; Hamilton & Winters, 1992; Wang & Winters, 1992). According to M. Lubinga, Ogundeji, and Jordaan (2014), the negative and positive value of potential trade were defined differently. A negative value of potential trade indicates that un-exhausted trade potential is presented; therefore, there is a room for extending the trade. Additionally, a positive potential trade value specifies that there is no un-exhausted trade potential which means that the room for trade expansion is very narrow or full.

The trade performance of Cambodia's bilateral trade with her trading partners is predicted by using the Relative differences index (Rd) and Absolute difference (Ad). There are several papers employed the same technique to forecasting the potential trade such as Batra (2004), Chen et al. (2008), M. H. Lubinga (2009), and M. Lubinga et al. (2014).

Firstly, the mean predicted trade value and mean actual trade value are used to calculate the Relative difference index (Rd %) as following:

$$\mathrm{Rd}_{ijt}(\%) = \frac{(\mathrm{A\_Trade}_{ijt} - \mathrm{P\_Trade}_{ijt})}{(\mathrm{A\_Trade}_{ijt} + \mathrm{P\_Trade}_{ijt})} * 100 \qquad (10)$$

where Rd_{ijt} is the relative differences index in Cambodia's trade with its trading partners j measure as percentage

$\mathrm{A\_Trade}_{ijt}$ is the mean actual trade and $\mathrm{P\_Trade}_{ijt}$ is mean predicted trade

The Rd % can change between -1 and 1, which describes the good and bad trade performs between Cambodia and its trading partners as well as the future trade destination. It assumes that a country has a good trade performance and cooperation with its partners when the Rd % is positive. The negative Rd % indicates that a country has a bad trade performance. To present the trade performance between Cambodia and its trading partners in ASEAN, we calculate the Rd % follow out-off-sample approach. The out-off-sample approach is observed at three different periods. First is to observe the Rd % during the whole period of 21 years of study which is from 1995 to 2015.

Secondly, we determine the trade performance between them during the period of 1999-2015. Fourthly, we estimate the trade performance during the period of 2004-2015. Finally, we find the trade performance from 2010 to 2015.

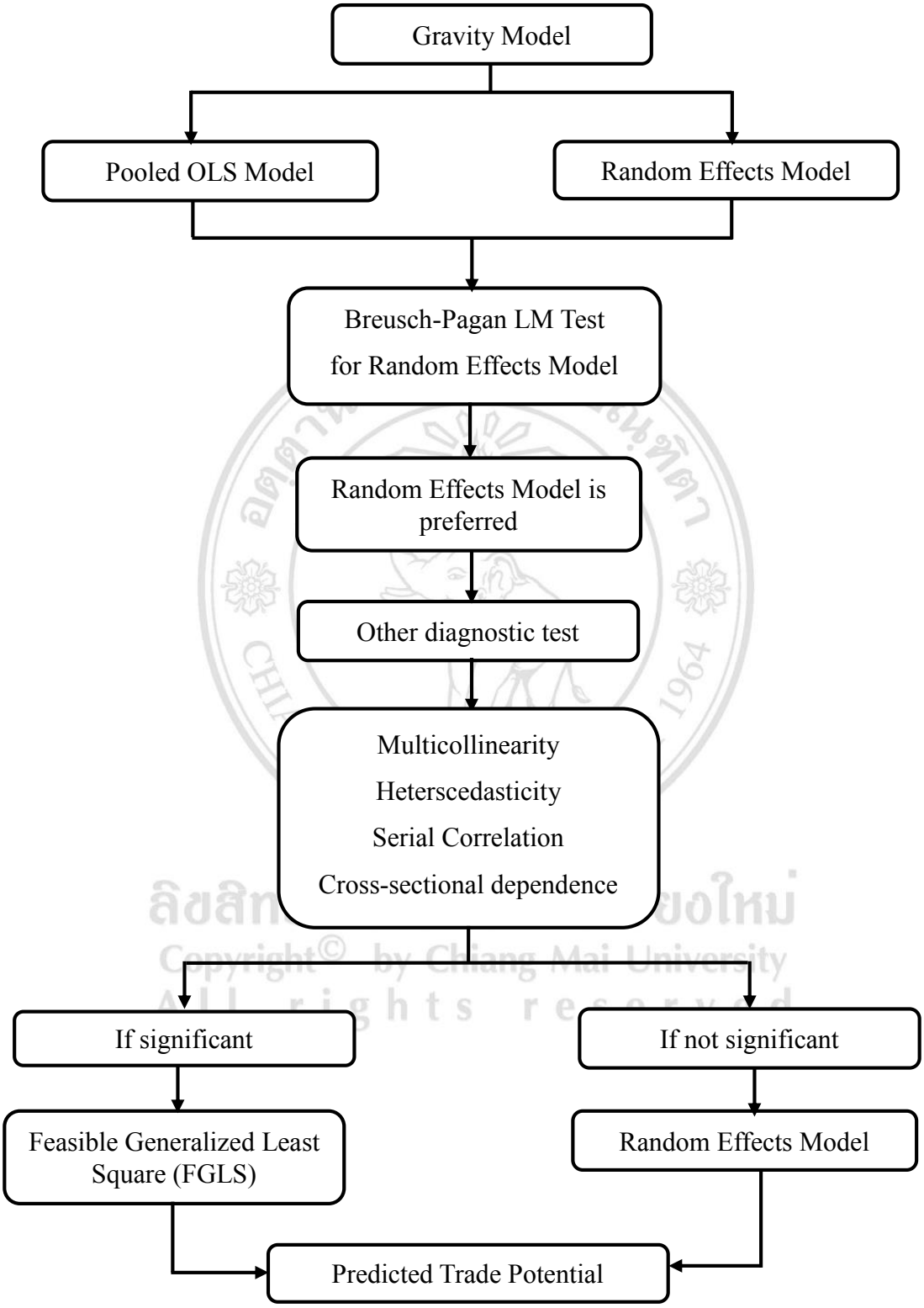
Secondly, Absolute difference (Ad) index is also used to investigate the trade performance. Absolute difference index is determined by the subtraction between the actual trade (A_Trade) and the simulated trade (P_Trade).

$$\mathrm{Ad}_{ijt} = \mathrm{A\_Trade}_{ijt} - \mathrm{P\_Trade}_{ijt} \quad (11)$$

where Ad_{ijt} is an Absolute difference index which sometimes is known as exhausted trade potential

The Ad index can be used to obtain the value of trade gain or trade lose between home country and its major trading partners, and can be used to investigate the direction of the future trading countries. The Absolute difference (Ad) is used to verify that Relative differences index (Rd %) is true.

3.6 Summary of Methodology Framework



Source: Created by Author

Figure 3.1 Methodology Framework