ARBITRAGE OPPORTUNITIES DISCOVERY IN THAILAND'S SPOT AND FUTURES MARKET : PAIR TRADING STRATEGY FROM THRESHOLD CO-INTEGRATION MODEL

SURASAK CHOEDPASUPORN

A THEMATIC PAPER SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF MANAGEMENT COLLEGE OF MANAGEMENT MAHIDOL UNIVERSITY 2014

Independent Study

entitled

ARBITRAGE OPPORTUNITIES DISCOVERY IN THAILAND'S SPOT AND FUTURES MARKET: PAIR TRADING STRATEGY FROM THRESHOLD CO-INTEGRATION MODEL

was submitted to the College of Management, Mahidol University for the degree of Master of Management

on

4 November, 2014

……………………………………

Mr. Surasak Choedpasuporn
Candidate

……………………………………

Piyapas Tharavanij,
Ph.D.
Advisor

……………………………………

Assoc. Prof. Tatre Jantarakolica,
Ph.D.
Chairperson

……………………………………

Assoc. Prof. Annop Tanlamai, Ph.D.
Dean
College of Management, Mahidol University

……………………………………

Kaipichit Ruengsrichaiya,
Ph.D.
Committee Member

ACKNOWLEDGEMENTS

First of all, I would like to express my deepest gratitude to the chairperson of IS committee and co-advisor of this study, Assoc. Prof. Dr. Tatre Jantarakolica, along with the advisor, Dr.Piyapas Tharavanij and another committee member, Dr.Kaipichit Ruengsrichaiya for providing an encouragement, a lot of insightful information and invaluable suggestions throughout the study. Completion of this study might be virtually impossible without their support. They warmly welcomed and gave me an opportunity for this study, even though I am not a Financial Management student that they directly supervise. Aside from this study, I also learnt many useful lessons from them, especially in Finance and Economic topics.

In addition, I would like to thank Asst. Prof. Dr. Phallapa Petison and Asst. Prof. Dr. Parisa Rungruang and other lecturers in CMMU for teaching me very useful lessons. I would also like thank faculty staffs for their support.

Moreover, I would also like to thank my friends at this college for their friendships, encouragement and taking care each other in our time together.

I also would like to thank the author of "Optimal Positioning in Thailand's Spot and Futures Markets: Arbitrage Signaling from Threshold Cointegration Model", Khemarat Songyoo. He selflessly supported and gave a lot of useful recommendations and information for this study.

Last but not least, I heartily thank my beloved family for always supporting me. If there was no support from them, I might not be able to achieve any goal including this study.

Surasak Choedpasuporn

ARBITRAGE OPPORTUNITIES DISCOVERY IN THAILAND'S SPOT AND FUTURES MARKET : PAIR TRADING STRATEGY FROM THRESHOLD CO-INTEGRATION MODEL

SURASAK CHOEDPASUPORN 5550221

M.M.

INDEPENDENT STUDY ADVISORY COMMITTEE : PIYAPAS THARAVANIJ, Ph.D., ASSOC. PROF. TATRE JANTARAKOLICA, Ph.D., KAIPICHIT RUENGSRICHAIYA, Ph.D.

ABSTRACT

This study examines the 5-minutes intraday price relationship between pairs of assets in Thailand's Stock Spot (SET) and Futures (TFEX) Markets. Three pairs of series of the same underlying asset (SET50, KTB, TRUE) which trade between 2nd July, 2014 to 29th August, 2014 are studied. The study finds long-run relationship and short-run dynamic of the prices of pairs. Considering the existence of the transaction cost in practical trading, the price relationship is estimated following the Threshold Vector Error Correction Model (TVECM). The TVECM pair trading strategy is formulated using the estimated parameters. Applying the formulated strategy, the arbitrage opportunities are found. The performance of the TVECM pair trading strategy is superior to the traditional pair trading strategy.

KEY WORDS: Threshold Co-integration / Pair Trading / Market Neutral / Statistical Arbitrage / Spot and Future

36 pages

CONTENTS

CONTENTS (cont.)

LIST OF TABLES

LIST OF FIGURES

CHAPTER I
INTRODUCTION

Equity Investments is considered as a high risky activity. The research of Nestorovski and Naumoski (2013) found that the volatility of the economy and the risk of equity investing are correlated in the same direction. Current economic conditions in Thailand and global economic conditions fluctuate. In particular, some countries in the European Union, experiencing no ability to repay debt. Including the United States, which has debt problems as well. As a result, global economy fluctuate and dynamic without a clear direction.

In Thailand, apart from global economic factors, domestic factors such as natural disasters and political instability also affect the volatility of the Thai Economy. According to the earlier statement of Nestorovski et al.(2013) , this will lead to the risk of investing in equities in the end. Thus, under the present circumstances, the risk management of investment, is very important to investors. The Stock Exchange of Thailand as a regulator and promoter of investment in the Thailand’s Stock Market, has initiated a market in derivatives, called Thailand Futures Exchange (TFEX), in the year 2006 with many important objectives. One of the objectives is to provide investors a tool for efficient risk management such as Futures and Options.

Aside from risk management, investors may use derivatives in many way. One approach to take advantage of derivatives in both risk management and profit speculation is Pair Trading Strategy. In some countries, Pair Trading Strategy is widely used, especially in fund equity risk (Hedge Fund) (Caldeira & Moura, 2012). Pair Trading Strategy (a.k.a Market Neutral Strategy) is a possible way that investors can expect high returns with low risk. Another strength of Pair Trading is to eliminate feelings, judgment and ability of investor out of investing decision making, then replace them with a clear principle. (Gatev, Goetzman, Rouwenhorst, 2006). Principles of Pair Trading strategy is to invest in two assets, which the prices can be expected to closely related, at the same time. When prices of the pair diverge, open a

Short Position in a higher priced asset, and simultaneously open a Long Position in the lower priced asset with equal value. Over time, the prices will converge, then close all Positions. The difference in price at open and close positions will become profit (Vidyamurthy, 2004). Pair Trading Strategy can be applied to many types of assets, including stocks, derivatives and commodity products. Key success factors of strategy implementation are pair selection and position timing. Vidyamurthy (2004) proposed a way for pair selection by using the concept of Cointegration. The Cointegration is a process of time series, which is introduced by Granger (1981) as a tool to analyze the relationships of couples in long-term. In short-run, the relationship is analyzed using Error Correction Model (a.k.a. ECM) (Enger & Granger, 1987). In reality, with existence of the transaction cost such as commission cost, the ECM might not be suitable to describe the relationship of the pair. There is an extend model, Threshold Vector Error Correction Model (a.k.a. TVECM), which is considered a difference of adjustment process in different regimes. The TVECM can also be applied to generate trading signal for position timing in Pair Trading Strategy (Songyoo, 2013).

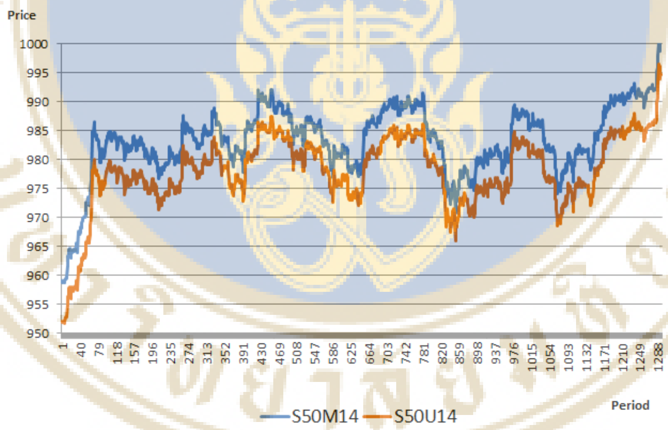

Figure 1.1 Prices of S50M14 and S50U14 (Futures of SET50 Index)

The previous study of Songyoo (2013) focused in pair of Spot and its Future. This paper aims to study further of the TVECM pair trading strategy in broader dimension. Figure 1.1 shows intraday prices of S50M14 and S50U14 which share the same underlying asset of SET50 Index. The illustration shows that their prices are highly correlated. The author considers this type of pair might be found the arbitrage opportunity as well. Our study will study on this type of pair.

The performance of the TVECM pair trading strategy is measured and compared to the performance of traditional pair trading strategy.

This paper is separated into 6 chapters as Introduction, Literature Review, Methodology, Empirical Results, Conclusion, References.

CHAPTER II
LITERATURE REVIEW

2.1 Theories

2.1.1 Efficient Market Hypothesis & Cost of Carry Model

Efficient Market Hypothesis (EMH) was presented by Fama (1965). Fama explained that the capital market is efficient or the current price is already reflected by the stock information related to them. Therefore, the current price is a reasonable price. Fama (1970) has further divided the Market into three forms.

2.1.1.1 Weak Form is considered that the current price already reflected by the price information in the past. But the investors can utilize public information, and insider information to make an abnormal profit.

2.1.1.2 Semi-strong Form is considered that the current price already reflected by the past price and public information. Only investors with insider information can make an abnormal profit.

2.1.1.3 Strong Form is considered that the current price already reflected by the past price, public and insider information. No one can make an abnormal profit.

We can conclude that in any form of market, the past price cannot be used to make an abnormal profit. Additional, if the market is fully efficient, all economic agent will have the same information. Then, the future price at time (t) should be expected to equal the spot price at the maturity date (T) of the future contract. Considering the cost of carrying model into this situation, the future price will be expected to equal to the spot price plus its carrying cost through time until the maturity date.

$$f_{t,T} = E_t(S_T) = S_t(\text{the cost of carrying asset over time})$$

If that is the case, the gap between future price and spot price should be constant and there is no arbitrage opportunity. But if the market is not fully efficient, the gap will not be constant and there is an arbitrage opportunity (Songyoo, 2013).

2.1.2 Mispricing, Arbitrage Opportunity & Pair Trading

If the market is not fully efficient and the gap between the future price and spot price is not constant, there will be a chance that the future price does not equal to the spot price plus its carrying cost or they are mispricing. For example, the future price ft,Tis higher than the spot price plus carrying cost.

$$f_{t,T} - S_t(\text{the cost of carrying asset over time}) > 0$$

In this scenario, there will be an arbitrage opportunity to short sell the expensive one (ft,T) and buy the cheap one (St). The selling force will lower the price of the expensive one, whereas the buying force will increase the price of the cheap one. As a result, the difference between both prices will be diminished until it disappears (Songyoo, 2013).

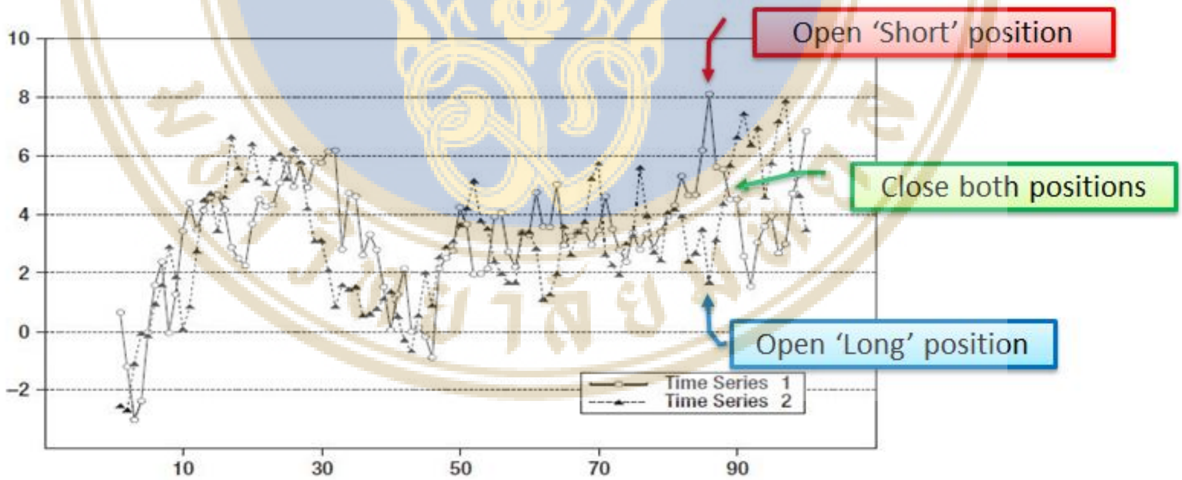

Figure 2.1 Demonstration of the pair trading strategy

The Pair Trading Strategy comes into play when the investors found the existence of an arbitrage opportunity. As we known that the future price and the spot price will move together but they might diverge in some chances. Opening of 'Short Selling' position in the expensive one, and in the same time opening of 'Long Buying' position in the cheap one, then hold the positions until the prices converge, will create a profit with a minimal risk (Vidyamurthy, 2004).

2.2 Empirical Research

Thongthip (2010) applied Threshold Autoregressive Model (TAR) along with cost of carry model to explain the lead-lag and long-run relationship between SET50 Index and its future which were traded between October 2008 to September 2009. The result shows that the prices of pair move together and confirm that long-run relationship exists between both market. Anyway, lead-lag relationship does not found in daily data, but it found at intraday data of 5-minute data. Kaewmongkolsri (2011) studied KTB and its future which were traded between July to December 2010 with Vector Error Correction Model and found long-run and short-run relationship. Intraday price data is recommended to use for study, since the relationship do not last for more than half an hour. This study also confirms long-run relationship of pair prices at 10-minute data. Songyoo (2013) also applied the Threshold Cointegration (TVECM) to explain the relationship between spot and future market. This study also confirms that long-run relationship exists between two markets for SET50 Index, KTB equity and their futures at intraday 10-minute data. In this study, the author also formulated a pair trading strategy by applying estimated threshold as a trigger point for positioning signal. The simulated portfolio using price data traded between September - November 2011 can make a positive return and confirm the existence of an arbitrage opportunity.

CHAPTER III

METHODOLOGY

3.1 Unit Roots

The stationary property of the data is one of primary factors to be studied. The data with a stationary process will have a steady state of mean and variance as time passes. In the other hand, if the process is non-stationary, the process is said to has a unit root. A common way to test a unit root is performing Augmented Dickey-Fuller Test (a.k.a. ADF Test) (Dickey & Fuller, 1981). The ADF Test can be performed by using the following equation.

$$\Delta x_t = \mu_1 + \gamma x_{t-1} + \mu_2 t + \sum_{i=1}^{\infty} \beta_i \Delta x_{t-i} + \varepsilon_t$$

Where X_t represents the series of data to be tested a unit root. For this study, X_t is a series of log futures price or log spot price.

The test hypothesis is H_0 : = 0 and H_a : <> 0.If the null hypothesis is rejected, the series is stationary and has no unit root. The order of integration is at level or I(0). If the null hypothesis is failed to be rejected, the series is non-stationary and has a unit root. In this case, we need to test the series at its first difference. If the series is stationary at its first difference, the series is said to has integration of order 1 or I(1).

3.2 Co-integration & Error Correction Model

Granger (1981) has proposed a long-term relationship between the 2 variables by explaining that when 2 variables have the same Order of Integration, and found a linear combination of both variables that produces another variable which has a lower Order of Integration, then the 2 variables are considered to have a long-term relationship or said to have a Cointegration property. However, when 2 variables are related in the long term, the two variables may deviate apart in the short term. To maintain the long-relationship, there must be a mechanism to adjust the deviation of the two variables to return to their long-term equilibrium. Such a mechanism has been proposed as the Error Correction Model (a.k.a. ECM) (Engle & Granger, 1987). For the Error Correction Model of CI(1,1) is formulated as the following model.

$$\Delta x_t = \alpha_0 + \alpha_x z_{t-1} + \sum \alpha_{1j} \Delta x_{t-j} + \sum \alpha_{2j} \Delta y_{t-j} + \epsilon_{xt}$$

$$\Delta y_t = b_0 + b_y z_{t-1} + \sum b_{1j} \Delta y_{t-j} + \sum b_{2j} \Delta y_{t-j} + \epsilon_{yt}$$

Where $z_{t-1} = x_{t-1} - \beta y_{t-1}$

z_{t-1} is called 'Error Correction Term' (a.k.a. ECT). The ECT is the adjusting part that maintains both variables to return or converge to their equilibrium. β is the cointegration coefficient.

Johansen (1991) has proposed a Maximum-Likelihood Estimation method to test the ECM as an extend version of Vector Autoregressive Model (VAR) called Vector Error Correction Model (VECM). The test follows this hypothesis.

$H_0 : rank(\Pi) = 0$ and $H_a : rank(\Pi) \neq 0$

Where Π is the cointegration matrix.

Using the Trace test or Maximum Eigenvalues test, if the null hypothesis is failed to be rejected, then there is no cointegration, if the null hypothesis is rejected, then there is cointegration.

3.3 Threshold Vector Error Correction Model

Balke & Fomby (1997) have suggested the possibility that the relationship between the two variables may not adjust as a simple linear process or may not happen all the time, but it may occur when variables are deviations from equilibrium up to a certain point (Threshold value or γ). For example, economic agents may not take any action, if they expect their returns do not more than the costs occurred. If the adjustment process is following this feature, it will have a Threshold Cointegration property. We can consider Threshold Cointegration property to have many regimes or bands. Each regime will have its adjustment behavior of its own. For Three regimes, the equation is as follows, For regime 1, when $(x_{t-1} - \beta y_{t-1}) \leq \gamma_a$

$$\Delta x_t = a_0 + \alpha_{1x} z_{t-1} + \sum a_{11j} \Delta x_{t-j} + \sum a_{12j} \Delta y_{t-j} + \epsilon_{xt}$$

$$\Delta y_t = b_0 + \alpha_{1y} z_{t-1} + \sum b_{11j} \Delta x_{t-j} + \sum b_{12j} \Delta y_{t-j} + \epsilon_{yt}$$

For regime 2, when $\gamma_a < (x_{t-1} - \beta y_{t-1}) \leq \gamma_b$

$$\Delta x_t = a_0 + \alpha_{2x} z_{t-1} + \sum a_{21j} \Delta x_{t-j} + \sum a_{22j} \Delta y_{t-j} + \epsilon_{xt}$$

$$\Delta y_t = b_0 + \alpha_{2y} z_{t-1} + \sum b_{21j} \Delta x_{t-j} + \sum b_{22j} \Delta y_{t-j} + \epsilon_{yt}$$

For regime 3, when $\gamma_b < (x_{t-1} - \beta y_{t-1})$

$$\Delta x_t = a_0 + \alpha_{3x} z_{t-1} + \sum a_{31j} \Delta x_{t-j} + \sum a_{32j} \Delta y_{t-j} + \epsilon_{xt}$$

$$\Delta y_t = b_0 + \alpha_{3y} z_{t-1} + \sum b_{31j} \Delta x_{t-j} + \sum b_{32j} \Delta y_{t-j} + \epsilon_{yt}$$

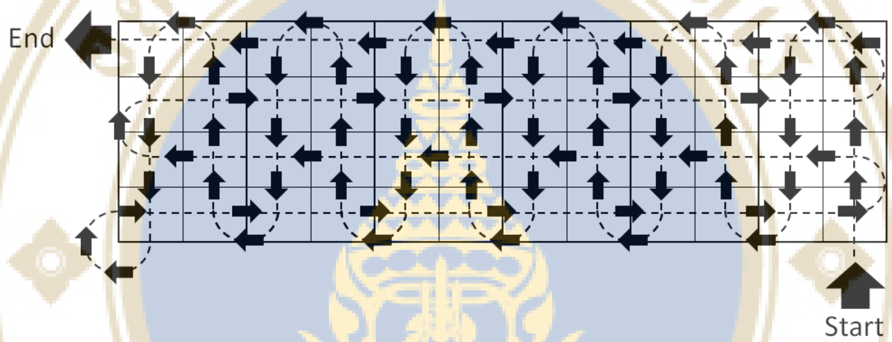

Figure 3.1 Illustration of Grid Search

Balke et. al (1997) presented a method to test the Threshold Cointegration as follows. The test is divided into two steps. The first step is to test a Cointegration property of the Time-series. If the Cointegration exists, then test the next step by testing for a Threshold or Nonlinear property of the Time-series. However, this guideline is available only when we know the Cointegration Vector (β). For this reason, Hansen & Seo (2002) have proposed a MLE method using the Grid Search method illustrated as Figure 2.1. This method will generate all possible pairs of the β (Cointegrating Vector) and γ (Threshold value) within a scope and constraints, then test every pair to find the optimal β (Cointegrating Vector) and γ (Threshold value) by using the AIC and SBIC selection criteria. In addition, Hansen et. al (2002) also proposed the SupLM Test called 'Hansen-Seo test' to test the null hypothesis of linear cointegration (no threshold behavior) versus the alternative hypothesis of threshold cointegration.

3.4 Trading rule for pair trading strategy

From the TVECM, we known that the adjustment process will separated into many regimes. The regimes will be decided from the threshold values. Applying this concept into Pair Trading Strategy, we can use the threshold values and regimes as a signaling tool. From the study of Songyoo (2013), we found that using 3-regimes TVECM, most of observation was found to fall into the Regime 2. This regime is also called “No-arbitrage band”. If the observation fall into Regime 1 or Regime 3, the gap of mispricing will strong enough to gain a profit. The previous study suggested a pair trading rule as these steps. At first, open Long/Short positions when the observation is out of Regime 2. Then, if observation returns to the Regime 2, close the positions. The trading rule is illustrated as Figure 3.2.

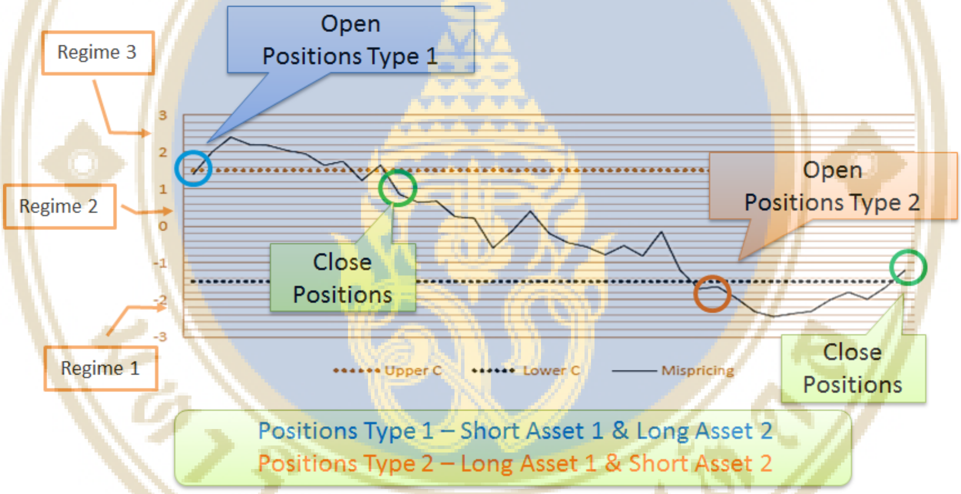

Figure 3.2 Demonstration of TVECM pair trading rule proposed by Songyoo (2013)

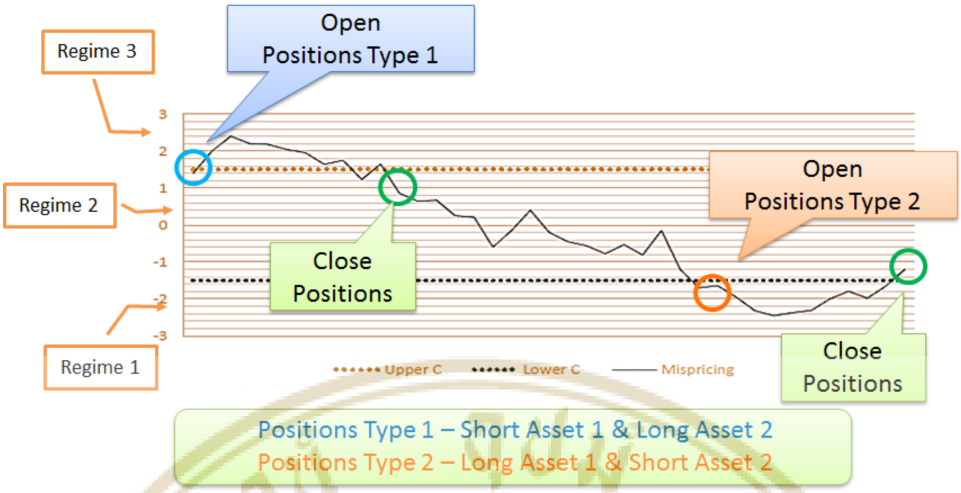

Figure 3.3 Demonstration of Adjusted TVECM pair trading rule

Since the transaction cost is considered to highly affect the performance of trading rule, we adjusts the trading rule to reduce trading over minor gap of mispricing by skipping of position closing when the observation returns to the Regime 2, and instead, close the position only when observation shifts across regime 1 to regime 3 or the opposite way. The adjusted trading rule is illustrated as Figure 3.3.

3.5 Performance measurement of trading strategy

To be realistic, we perform the out-sample test by applying time-rolling in our measurement. The time-rolling procedure is described as following. First, setting initial training periods to estimate the parameters. This study set the initial period as 600 periods or 10 trading days. Second, execute trading rule by using the estimated parameters for next periods. This study uses these parameters for 300 periods or 5 trading days. Third, after end of rule execution period from second step, move the training period forward same length as the execution period and repeat first and second steps until end of data. The time-rolling procedure is illustrated as Figure 3.4

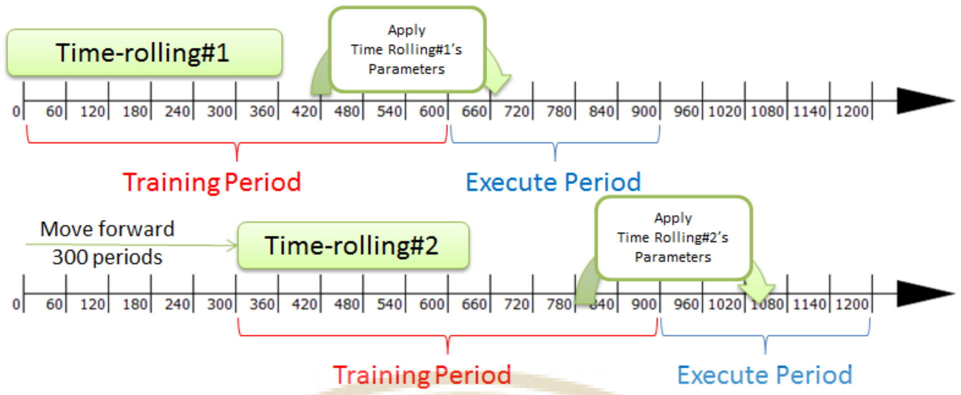

Figure 3.4 Demonstration of time-rolling procedure

Table 3.1 Transaction cost of trading for proprietary trade

Asset type	Transaction Fee	Trading size
Stock	0.0015 x 0.07 x [stock price] (equals to VAT of commission cost)	Multiplying of 100 units
SET50 Future	7 THB per contract	200 units per contract
Stock Future	35 THB per contract	1,000 units per contract

To calculate net profit, we consider using transaction cost of proprietary trade. The transaction cost is described as Table 3.1. Apply these transaction cost rate, we can calculate net profit as the performance of the trading strategy. Aside from absolute return from the net profit, we compare the performance with the traditional pair trading strategy. The traditional pair trading strategy applied moving averages and standard deviations as triggering signal. The position opening will occur when the observation deviates from the moving average more than 2 times of standard deviations. The position closing will occur when the observation converges.

CHAPTER IV
DATA

4.1 Data and pair selection criteria

Our research aims to study the long-run equilibrium relationship and the short-run dynamic between the prices of assets sharing the same underlying asset. The pair of asset with a long-run relation will be studied its potential for the arbitrage opportunities from the deviation of two prices. Our focus is the assets that are traded in Thailand Stock Spot (SET) and Futures (TFEX) markets. The data used in the study are obtained from the eFin Smart Portal software provided by www.efinancethai.com on 28th September, 2014. To prevent a no-trading price bias, a criteria based on liquidity of series will be a counter-measure. The selected series will be ones with less than 10% of missing volume trade. Previous research suggested to use price data that higher frequency than half an hour (Kaewmongkolsri, 2011). A recent study of Songyoo (2013) found that optimal frequency was 10-minute for that period. Anyway, we found that 5-minute frequency is more suitable for this research because the current liquidity is more than previous research.

In this research, the pair is formed by 2 types of series. For type 1, the pair is formed by a spot and its future. For type 2, the pair is formed by two futures from different contract months of the same underlying asset.

4.2 Data and pair selection result

The pairs of assets are selected under the criteria. Trading period ranges from 2nd July 2014 to 29 August 2014 including 40 trading days. At 5-minute price data, there are 2,439 observations. Anyway, we found that the liquidity of most Stock Futures are very low, as a result there are 3 pairs selected under the criteria. The selected pairs are as following.

1) S50U14 (SET50 Index Futures September 2014 Contract)
   and S50Z14 (SET50 Index Futures December 2014 Contract)
2) KTB and KTBU14 (KTB Futures September 2014 Contract)
3) TRUE and TRUEU14 (TRUE Futures September 2014 Contract)

CHAPTER V
EMPIRICAL RESULT

5.1 Unit root test and long-run relationship estimation

Before examining the long-run relationship of each pairs, order of integration of each series and their cointegration residuals should be assessed. To assess the order of integrations, the unit root test will be performed by applying Augmented Dickey Fuller Test (ADF).

Table 5.1 ADF Test result of lnS50U14 and lnS50Z14 series

Price Series	**ADF Statistics**	**Critical Value (5% conf)**	**Conclusion**
ln S50U14	1.2074	-1.95	*Non-stationary*
First Diff of ln S50U14	-14.9476	-1.95	*Stationary*
ln S50Z14	1.3428	-1.95	*Non-stationary*
First Diff of ln S50Z14	-14.8831	-1.95	*Stationary*
Cointegration Residuals	-2.901	-1.95	*Stationary*

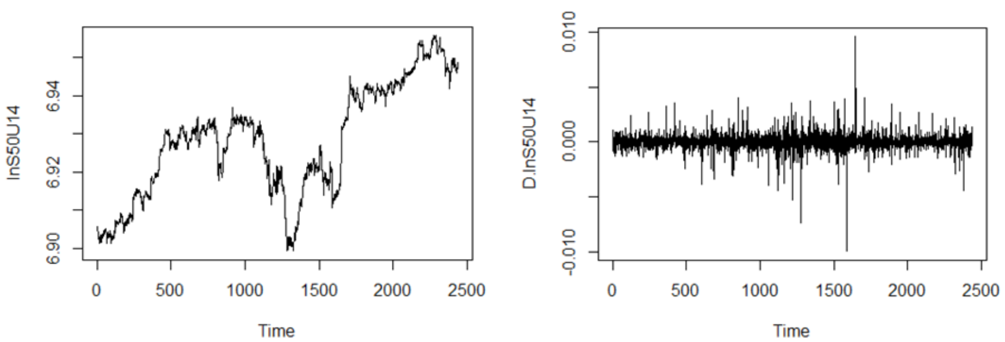

Figure 5.1 Plot of log S50U14 price series and its first difference

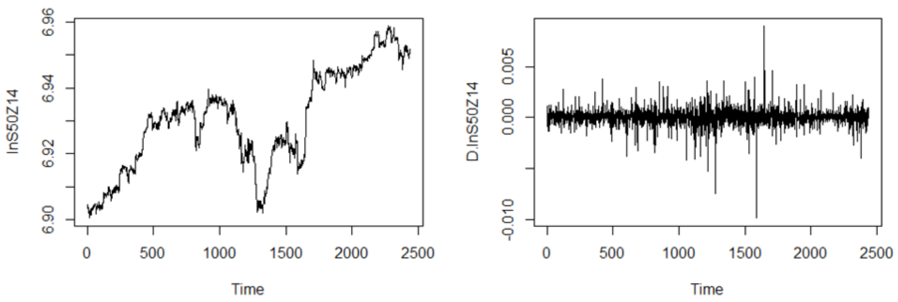

Figure 5.2 Plot of log S50Z14 price series and its first difference

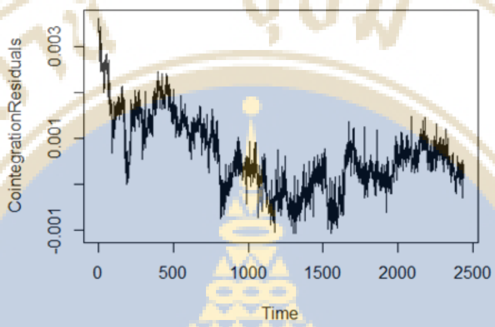

Figure 5.3 Plot of cointegrating residuals of both series

Table 5.1 summarizes the result of ADF Test of log of S50U14 price series and log of S50Z14 price series. Both log of S50U14 price series and log of S50Z14 price series are I(1) as illustrated in Figure 5.1 and Figure 5.2. The cointegration residuals of both series is I(0) or log of S50U14 price series and log of S50Z14 price series is cointegrated of order (1,1) as illustrated in Figure 5.3. The cointegrating equation of the long-run relationship for log of S50U14 price series and log of S50Z14 price series is as following equation.

$$lnS50U14 - 0.9731*lnS50Z14 - 0.1835 = residuals$$

5.2 Short-run dynamic estimation

To estimate short-run dynamic of pairs, we apply Johansen's MLE for Vector Error Correction Model (VECM). For pair of log S50U14 and log S50Z14, the optimal lags is selected by SBIC criteria as shown in Table 5.2 and the result of estimation shows as following equation.

Table 5.2 SBIC Criteria of each no. of lags

No. of Lags	SBIC Criteria
1	-75347.99
2	-75450.98
3	-75509.91
4*	-75511.85
5	-75498.51
Optimal No. of Lags = 4	

Table 5.3 Result of VECM for pair of log S50U14 and log S50Z14 at lags = 4

$\Delta lnS50U14$	**Coefficients**	**Standard Error**
Correction Term	-0.0097	0.0207
$\Delta lnS50Z14$	**Coefficients**	**Standard Error**
Correction Term	0.0350	0.0204

The error correction term is $y_{t-1} - 0.9731x_{t-1} - 0.1835$. For VECM, the coefficients of error correction term represent speed of adjustment. For this pair, speed of adjust of log S50U14 is 0.0097 with negative sign and speed of adjust of log S50Z14 is 0.0350 with positive sign. Considering magnitude of the adjustment speed, we can estimate that the convergence process will take at least 100 periods for log S50U14 series and 30 periods for log S50Z14 series. Since the speed of adjustment is quite slow, it can be an effect of transaction cost that might lead to Threshold behavior in the adjustment process.

5.3 Threshold Vector Error Correction Model Estimation

After we found long-run relationship behavior of the pair, we can further analyze the relation to assess existing of Threshold behavior. Following steps describes in chapter 3, we can estimate TVECM model for the pair series.

Table 5.4 SBIC criteria for each no. of lags for pair of log S50U14 and log S50Z14

No. of Lags	SBIC Criteria
1	-75285.54
2*	-75319.88
3	-75303.85
4	-75264.81
5	-75206.03
Optimal No. of Lags = 2	

Table 5.5 TVECM result under Three-regime for pair of log S50U14 and log S50Z14

Item	Values
Cointegrating Vector	(1,-0.9995833)
No. of Lags	2
Threshold Values	-0.0004032538 -0.0001141731
No. of Observations	2,439
Upper Regime *(78.45% of Obs)*	Coefficient ECT lnS50U14 = -0.010713508 Coefficient ECT lnS50Z14 = -0.003017899
Middle Regime *(16.54% of Obs)*	Coefficient ECT lnS50U14 = -0.4226353 Coefficient ECT lnS50Z14 = -0.1566429
Lower Regime *(5.01% of Obs)*	Coefficient ECT lnS50U14 = 0.1678026 Coefficient ECT lnS50Z14 = 0.7490027

For pair of log S50U14 and log S50Z14, the optimal lags is selected by SBIC criteria as shown in Table 5.4 and the result of TVECM estimation under Three-regime is shown in Table 5.5.

5.4 Time rolling test

Performance of trading rule is measured by calculation of net profit from portfolio simulation. Out-sample performance or time-rolling procedure is used to make the portfolio simulation more realistic. The simulation uses trading period from 2nd July 2014 to 29th August 2014. The first time rolling will be set training period from 2nd July 2014 to 16th July 2014 which consists of 10 trading days or 600 observations. The estimated parameters or threshold values will be used for next 5 trading days or 300 observations. The time-rolling is repeated using 5-day time rolling.

5.4.1 Hansen-Seo Test

Before performing the portfolio simulation, we should test that whether the pairs have threshold behavior or not. As discussed in chapter 3, Hansen et. al (2002) proposed 'Hansen-Seo test' to test existence of the threshold behavior. We perform Hansen-Seo test for every time rolling of each pair. The result is shown in Table5.6.

Table 5.6 Hansen-Seo test result

Time Rolling	P-Value		
	Pair 1 S50U14-S50Z14	**Pair 2** KTB-KTBU14	**Pair 3** TRUE-TRUEU14
1) Training Period : 1-600	0.06	**0.00	**0.00
2) Training Period : 301-900	0.53	**0.00	**0.00
3) Training Period : 601-1200	N/A	**0.00	**0.00
4) Training Period : 901-1500	0.10	0.42	**0.00
5) Training Period : 1201-1800	0.97	**0.00	**0.00
6) Training Period : 1501-2100	*0.03	0.17	**0.00
7) Training Period : 1801-2400	0.13	**0.00	*0.03

Note : * p-value < 0.05, **p-value < 0.01

Time rolling no.3) of Pair 1 cannot be estimated any threshold parameter, in this case we skip this time-rolling. For overall result of the test, the result shows that in some time rollings, the null hypothesis of linear cointegration (no threshold behavior) cannot be rejected. Anyway, we still can use the estimated threshold parameters to use as signaling point in pair trading strategy.

5.4.2 Trading Rule Performance Measurement

We simulate portfolio for each trading rule. The performance of each trading rule for each pair is shown in Table 5.7.

Table 5.7 Performance result of each trading rule

Trading Rule		**Pair 1 (S50U14 - S50Z14)** ***THB 200 / index point***	**Pair 2 (KTB - KTBU14)** ***1,000 shares / contract***	**Pair 3 (TRUE - TRUEU14)** ***1,000 shares / contract***
Trading Rule 1 (Original TVECM)	No. of Transactions	132	192	160
	Gross Profit	3,820	7660	6076
	Transaction Cost	1,848	7182	5774
	Net Profit	*1,972	478	302
Trading Rule 2 (Adjusted TVECM)	No. of Transactions	82	124	128
	Gross Profit	2,940	5280	5585
	Transaction Cost	1,148	4641	4620
	Net Profit	1,792	*639	*965
Traditional Pair Trading	No. of Transactions	36	40	46
	Gross Profit	1,280	1,790	1,535
	Transaction Cost	504	1,496	1,659
	Net Profit	776	294	(-124)

For pair 1 of S50U14 - S50Z14, the original TVECM pair trading strategy generates the best result of 1,972 THB of net profit for trading 1 contract at a time.

For pair 2 of KTB - KTBU14, the adjusted TVECM pair trading strategy generates the best result of 639 THB of net profit for trading 1 contract at a time.

For pair 3 of TRUE - TRUEU14, the adjusted TVECM pair trading strategy generates the best result of 965 THB of net profit for trading 1 contract at a time.

To study further, we shortened the length of training period and execute period which will make the pair trading rule response to price data faster. We adjusted starting time to make total execute periods equal to trading result above and let them comparable. In this part, the first time rolling will be set training period from 9th July 2014 to 16th July 2014 which consists of 5 trading days or 300 observations. The estimated parameters or threshold values will be used for next trading day or 60 observations. The time-rolling is repeated using 1-day time rolling. The performance of each trading rule for each pair is shown in Table 5.8.

Table 5.8 Result of each trading rule with different training and execute period

Trading Rule		**Pair 1 (S50U14 - S50Z14)** ***THB 200 / index point***		**Pair 2 (KTB - KTBU14)** ***1,000 shares / contract***		**Pair 3 (TRUE - TRUEU14)** ***1,000 shares / contract***	
		Training : 600 Execute : 300	Training : 300 Execute : 60	Training : 600 Execute : 300	Training : 300 Execute : 60	Training : 600 Execute : 300	Training : 300 Execute : 60
Trading Rule 1 (Original TVECM)	No. of Transactions	132	152	192	226	160	196
	Gross Profit	3,820	5,100	7660	9,890	6076	6,852
	Transaction Cost	1,848	2,128	7182	8,453	5774	7,073
	Net Profit	*1,972	*2,972	478	1,437	302	(-221)
Trading Rule 2 (Adjusted TVECM)	No. of Transactions	82	86	124	178	128	136
	Gross Profit	2,940	3,100	5280	8,750	5585	5,820
	Transaction Cost	1,148	1,204	4641	6,658	4620	4,908
	Net Profit	1,792	1,896	*639	*2,092	*965	*912
Traditional Pair Trading	No. of Transactions	36	60	40	46	46	34
	Gross Profit	1,280	1,820	1,790	1,770	1,535	1,257
	Transaction Cost	504	840	1,496	1,720	1,659	1,226
	Net Profit	776	980	294	50	(-124)	31

Result for length of training period = 300 and length of execute period = 60 is as follows.

For pair 1 of S50U14 - S50Z14, the original TVECM pair trading strategy also generates the best result of 2,972 THB of net profit for trading 1 contract at a time.

For pair 2 of KTB - KTBU14, the adjusted TVECM pair trading strategy also generates the best result of 2,092 THB of net profit for trading 1 contract at a time.

For pair 3 of TRUE - TRUEU14, the adjusted TVECM pair trading strategy also generates the best result of 912 THB of net profit for trading 1 contract at a time.

CHAPTER VI
CONCLUSION

This study examines the long-run relationship, short-run dynamic and threshold cointegration behavior of pairs of assets in Thailand's Stock Spot and Futures Market. The arbitrage opportunities among the markets are assessed from performing a portfolio simulation of a statistical arbitrage strategy called, "Pair Trading Strategy" (a.k.a "Market Neutral Strategy"). Three pairs of assets, "S50U14&S50Z14" – "KTB&KTBU14" – "TRUE&TRUEU14", are selected to studied using 5-minute price data between 2nd July, 2014 to 29th August, 2014 which include 40 trading days or 2,439 observations for each pair.

The result shows that each pair has long-run relationship. With existence of transaction cost (e.g. Commission Cost, value-added tax), threshold behavior is considered to be existing. Threshold Vector Error Correction Model (TVECM) is applied to estimate the thresholds parameter.

A previous study of Songyoo (2013) proposed a pair trading strategy which applies thresholds parameter from TVECM. This study adjusts the strategy and measures the performance by net profit of the simulated portfolios and comparing results to the traditional pair trading strategy which use standard deviation as trigger point.

The result shows that arbitrage opportunities exist in the markets for the proprietary trader using the pair trading strategy applying TVECM's threshold parameters as signal trigger. The performance of the original TVECM pair trading strategy and another adjusted version are superior to the traditional pair trading strategy. Difference in length of training period and execute period make the result of each strategy vary. Minus return found in original TVECM strategy, whereas the adjusted TVECM strategy still create a positive return. This would be a sign of more robustness in adjusted TVECM strategy. Anyway, the limited numbers of studied pairs is insufficient to decide the best strategy.

This study limits the size of portfolio to trade only one contract at a time. To trade more than one contract, the liquidity of the asset will be a major issue to be concerned. Anyway, we estimate a maximum potential return for each pair by calculation of average trading volume per period and then multiply it with return for one contract. As a result, we have maximum potential return of each pair in descending order as "S50U14&S50Z14" (THB 216,956) , "TRUE&TRUEU14" (THB 203,615) and "KTB&KTBU14" (THB 138,072).

REFERENCES

Balke, N. S., & Fomby, T. B. (1997). Threshold cointegration. International economic review, 627-645.

Dickey, D. A., & Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. Econometrica: Journal of the Econometric Society, 1057-1072.

Engle, R. F., & Granger, C. W. (1987). Co-integration and error correction: representation, estimation, and testing. Econometrica: journal of the Econometric Society, 251-276.

Fama, E. F. (1965). The behavior of stock-market prices. Journal of business, 34-105.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work*. The journal of Finance, 25(2), 383-417.

Gatev, E., Goetzmann, W. N., & Rouwenhorst, K. G. (2006). Pairs trading: Performance of a relative-value arbitrage rule. Review of Financial Studies, 19(3), 797-827.

Granger, C. W. (1981). Some properties of time series data and their use in econometric model specification. Journal of econometrics, 16(1), 121-130.

Hansen, B. E., & Seo, B. (2002). Testing for two-regime threshold cointegration in vector error-correction models. Journal of econometrics, 110(2), 293-318.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. Econometrica: Journal of the Econometric Society, 1551-1580.

Kaewmongkolsri, C. (2011). Lead-lag Relationship and Price Discovery in KTB Spot and KTB Futures Markets. Faculty of Commerce and Accountancy, Thammasat University.

Nestorovski, M., Naumoski, A. (2013). Economic Crisis and the Equity Risk Premium. 9th International ASECU Conference on "Systemic Economic Crisis: Current Issues and Perspectives".

REFERENCES (cont.)

Songyoo, K. (2013). Optimal Positioning in Thailand's Spot and Futures Markets: Arbitrage Signaling from Threshold Cointegration Model (Dissertation, Thammasat University).

Thongthip, S. (2010). Lead-lag Relationship and Mispricing in SET50 Index Cash and Futures Markets (Doctoral dissertation, Faculty of Economics, Thammasat University).

Vidyamurthy, G. (2004). Pairs Trading: quantitative methods and analysis (Vol. 217). John Wiley & Sons.

APPENDICES

APPENDIX A
EMPIRICAL RESULT OF PAIR 2 : KTB & KTBU14

A.1 Unit root test and long-run relationship estimation

Before examining the long-run relationship of each pairs, order of integration of each series and their cointegration residuals should be assessed. To assess the order of integrations, the unit root test will be performed by applying Augmented Dickey Fuller Test (ADF).

Table A.1 ADF Test result of lnKTB and lnKTBU14 series

Price Series	ADF Statistics	Critical Value (5% conf)	Conclusion
ln KTB	1.2048	-1.95	*Non-stationary*
First Diff of ln KTB	-16.8014	-1.95	*Stationary*
ln KTBU14	1.1618	-1.95	*Non-stationary*
First Diff of ln KTBU14	-15.7308	-1.95	*Stationary*
Cointegration Residuals	-5.1521	-1.95	*Stationary*

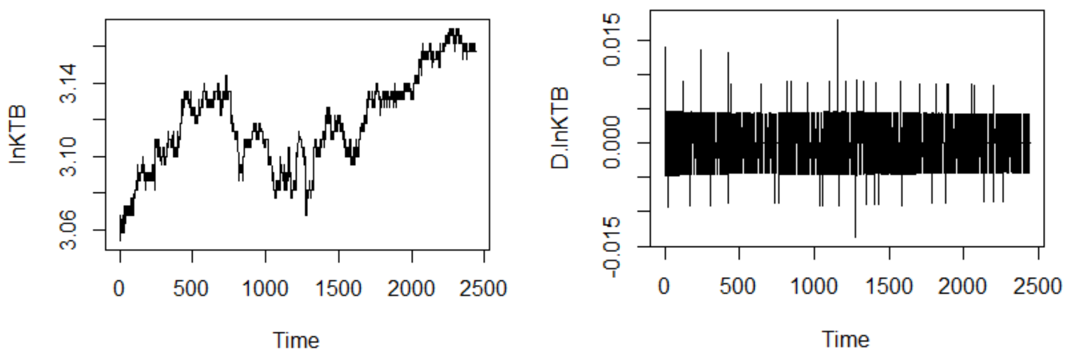

Figure A.1 Plot of log KTB price series and its first difference

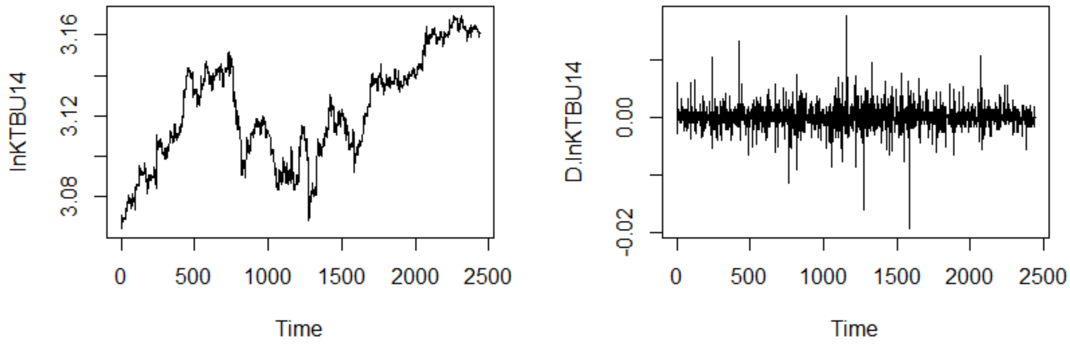

Figure A.2 Plot of log KTBU14 price series and its first difference

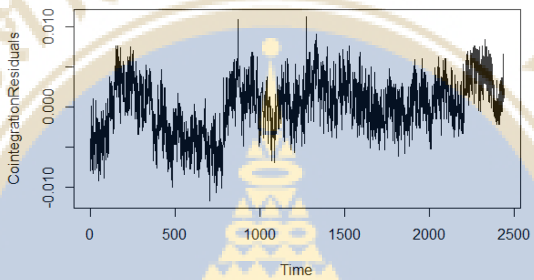

Figure A.3 Plot of cointegrating residuals of both series

Table A.1 summarizes the result of ADF Test of log of KTB price series and log of KTBU14 price series. Both log of KTB price series and log of KTBU14 price series are I(1) as illustrated in Figure A.1 and Figure A.2. The cointegration residuals of both series is I(0) or log of KTB price series and log of KTBU14 price series is cointegrated of order (1,1) as illustrated in Figure A.3. The cointegrating equation of the long-run relationship for log of KTB price series and log of KTBU14 price series is as following equation.

$$lnKTB - 0.9732*lnKTBU14 - 0.0798 = residuals$$

A.2 Short-run dynamic estimation

To estimate short-run dynamic of pairs, we apply Johansen's MLE for Vector Error Correction Model (VECM). For pair of log KTB and log KTBU14, the optimal lags is selected by SBIC criteria as shown in Table A.2 and the result of estimation shows as following equation.

Table A.2 SBIC Criteria of each no. of lags

No. of Lags	SBIC Criteria
1	-60146.33
2*	-60195.98
3	-60190.13
4	-60169.46
5	-60147.73
Optimal No. of Lags = 2	

Table A.3 Result of VECM for pair of log KTB and log KTBU14 at lags = 2

$\Delta lnKTB$	Coefficients	Standard Error
Correction Term	-0.1439	0.0181
$\Delta lnKTBU14$	**Coefficients**	**Standard Error**
Correction Term	0.0304	0.0119

A.3 Threshold Vector Error Correction Model Estimation

After we found long-run relationship behavior of the pair, we can further analyze the relation to assess existing of Threshold behavior. Following steps describes in chapter 3, we can estimate TVECM model for the pair series.

Table A.4 SBIC criteria for each no. of lags for pair of log KTB and log KTBU14

No. of Lags	SBIC Criteria
1*	-60069.65
2	-60053.86
3	-59988.43
4	-59893.8
5	-59816.43
Optimal No. of Lags = 1	

Table A.5 TVECM result under Three-regime for pair of log KTB and log KTBU14

Item	Values
Cointegrating Vector	(1,-0.9982666)
No. of Lags	1
Threshold Values	0.003186499 0.004160424
No. of Observations	2,439
Upper Regime *(24.66% of Obs)*	Coefficient ECT lnKTB = -0.3336562 Coefficient ECT lnKTBU14 = 0.1379370
Middle Regime *(11.08% of Obs)*	Coefficient ECT lnKTB = 0.5694124 Coefficient ECT lnKTBU14 = 0.7640926
Lower Regime *(64.26% of Obs)*	Coefficient ECT lnKTB = -0.09145585 Coefficient ECT lnKTBU14 = 0.01749865

For pair of log KTB and log KTBU14, the optimal lags is selected by SBIC criteria as shown in Table A.4 and the result of TVECM estimation under Three-regime is shown in Table A.5.

APPENDIX B
EMPIRICAL RESULT OF PAIR 3 : TRUE & TRUEU14

B.1 Unit root test and long-run relationship estimation

Before examining the long-run relationship of each pairs, order of integration of each series and their cointegration residuals should be assessed. To assess the order of integrations, the unit root test will be performed by applying Augmented Dickey Fuller Test (ADF).

Table B.1 ADF Test result of lnTRUE and lnTRUEU14 series

Price Series	**ADF Statistics**	**Critical Value (5% conf)**	**Conclusion**
ln TRUE	1.2871	-1.95	*Non-stationary*
First Diff of ln TRUE	-14.5677	-1.95	*Stationary*
ln TRUEU14	1.3269	-1.95	*Non-stationary*
First Diff of ln TRUEU14	-13.7488	-1.95	*Stationary*
Cointegration Residuals	-4.6128	-1.95	*Stationary*

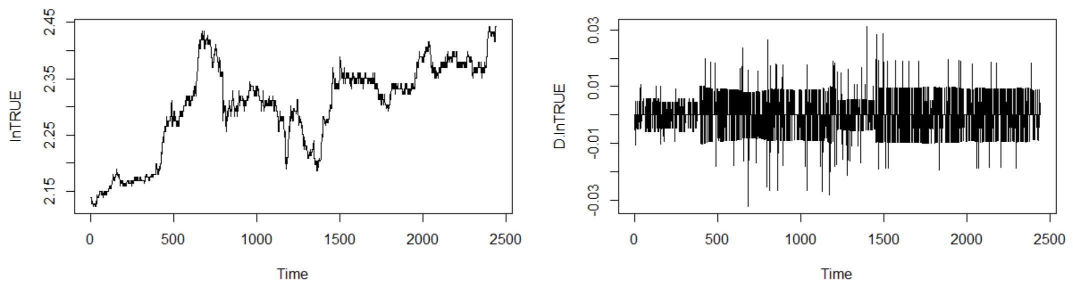

Figure B.1 Plot of log TRUE price series and its first difference

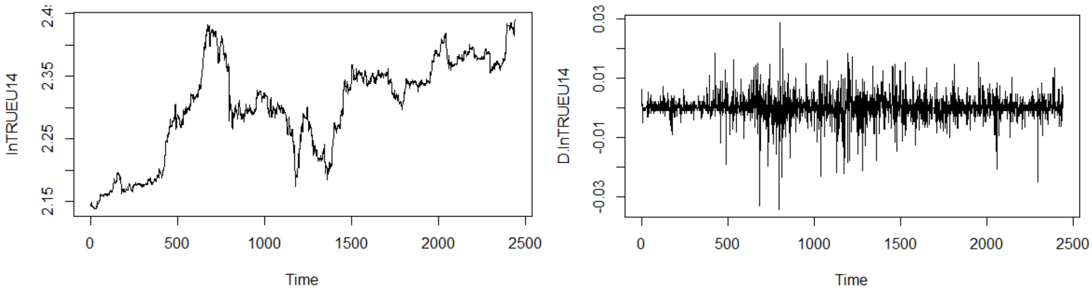

Figure B.2 Plot of log TRUEU14 price series and its first difference

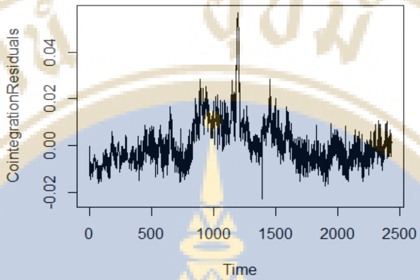

Figure B.3 Plot of cointegrating residuals of both series

Table B.1 summarizes the result of ADF Test of log of TRUE price series and log of TRUEU14 price series. Both log of TRUE price series and log of TRUEU14 price series are I(1) as illustrated in Figure B.1 and Figure B.2. The cointegration residuals of both series is I(0) or log of TRUE price series and log of TRUEU14 price series is cointegrated of order (1,1) as illustrated in Figure B.3. The cointegrating equation of the long-run relationship for log of TRUE price series and log of TRUEU14 price series is as following equation.

$$lnTRUE - 1.0133*lnTRUEU14 + 0.03115 = residuals$$

B.2 Short-run dynamic estimation

To estimate short-run dynamic of pairs, we apply Johansen's MLE for Vector Error Correction Model (VECM). For pair of log TRUE and log TRUEU14, the optimal lags is selected by SBIC criteria as shown in Table B.2 and the result of estimation shows as following equation.

Table B.2 SBIC Criteria of each no. of lags

No. of Lags	SBIC Criteria
1	-52871.78
2*	-52970.13
3	-52947.77
4	-52926.33
5	-52882.86
Optimal No. of Lags = 2	

Table B.3 Result of VECM for pair of log TRUE and log TRUEU14 at lags = 2

$\Delta lnTRUE$	**Coefficients**	**Standard Error**
Correction Term	-0.0556	0.0137
$\Delta lnTRUEU14$	**Coefficients**	**Standard Error**
Correction Term	0.0229	0.0096

B.3 Threshold Vector Error Correction Model Estimation

After we found long-run relationship behavior of the pair, we can further analyze the relation to assess existing of Threshold behavior. Following steps describes in chapter 3, we can estimate TVECM model for the pair series.

Table B.4 SBIC criteria for each no. of lags for pair of log TRUE and log TRUEU14

No. of Lags	SBIC Criteria
1	-52829.88
2*	-52871.50
3	-52796.82
4	-52730.27
5	-52674.12
Optimal No. of Lags = 2	

Table B.5 TVECM result under Three-regime for pair of log TRUE and log TRUEU14

Item	Values
Cointegrating Vector	(1,-1.001244)
No. of Lags	2
Threshold Values	-0.007136592 0.017959877
No. of Observations	2,439
Upper Regime *(1.93% of Obs)*	Coefficient ECT lnTRUE = 0.2421379 Coefficient ECT lnTRUEU14 = 0.1656139
Middle Regime *(59.32% of Obs)*	Coefficient ECT lnTRUE = -0.01867759 Coefficient ECT lnTRUEU14 = 0.01332959
Lower Regime *(38.75% of Obs)*	Coefficient ECT lnTRUE = -0.19547302 Coefficient ECT lnTRUEU14 = 0.04773023

For pair of log TRUE and log TRUEU14, the optimal lags is selected by SBIC criteria as shown in Table B.4 and the result of TVECM estimation under Three-regime is shown in Table B.5.